İ S T A T İ S T İ Ğ E G İ R İ Ş

DERS NOTLARI

Prof.Dr. Ömer Cevdet Bilgin

Atatürk Üniversitesi | Fen Fakültesi | İstatistik Bölümü

Son güncelleme | 18 Eylül 2024

“GÖRÜNEN HAKİKAT OLSAYDI, BİLİME İHTİYACIMIZ OLMAZDI!”

İÇİNDEKİLER

1. **Temel Kavramlar**
2. **Görsel Betimleme**
3. **Sayısal Betimleme**
4. **İlişkileri Betimleme – Bir Nitel ve Bir Nicel Değişken**
5. **İlişkileri Betimleme – İki Nitel Değişken**
6. **İlişkileri Betimleme – İki Nicel Değişken**
7. **Veri Toplama Metotları**
8. **Nispi Frekanslardan Olasılıklara**
9. **Rastgele Değişkenler ve Kesikli Olasılık Dağılımları**
10. **Beklenen Değerler**
11. **Normal Dağılım**
12. **Örnekleme Dağılımları**
13. **Güven Aralıkları**
14. **Anlamlılık Testleri – Aritmetik-ortalama ve ve Orantı**
15. **Anlamlılık Testleri – Frekanslar ve En Küçük Kareler Doğrusunun Eğimi**
16. **Anlamlılık Testleri – İkiden Çok Aritmetik-ortalamayı Karşılaştırmak**

Tablo A. **Rastgele Sayılar**

Tablo B. **Standart Normal Olasılıklar**

Tablo C. **Student-*t* Dağılımına ait Kritik Değerler**

Tablo D. **Ki-kare Dağılımına ait Kritik Değerler**

Tablo E. ***F*-dağılımına ait Kritik Değerler**

ÖNSÖZ

Bu bir giriş dersi; seni bir istatistikçiye dönüştümek gibi bir kastı yok. Mamafih, dersin sonunda, basit hesaplamaları yapabilmen, istatistiksel yazılım paketlerinin çıktılarını anlayabilecek ve istatistik öğrenimini ilerletebilecek bir temele sahip olman ve bir istatistikçiyle etkilice iletişebilmen umulur. Bunların yanısıra, istatistiği, bilim hakkında ve kanıt hakkında düşünmenin bir yolu olarak görmen umulur; sadece biribiriyle bağlantısız bir takım araçlar olarak görmen değil. Konular bu derste aynı zamanda istatistiksel düşünme bağlamında ele alınır ki böyle bir anlayış veriye-dayalı ve risk-merkezli dünyamızda her gün daha önemli olmaktadır.

İstatistiksel testlerin ve metotların sayısının çok fazla olmasına karşın, istatistik esasen nispeten çok daha az sayıdaki kavramlara ve fikirlere dayanır. Bu dersin maksadı, bu genel kavramların ve fikirlerin anlaşılmasına odaklanmak, onlarla gerçek yaşam problemleri arasında bağ kurabilme ve gerçek yaşam problemlerinin çözümünde onları kullanabilme yeteneğini kazandırmaktır.

BÖLÜM 01 | TEMEL KAVRAMLAR

Giriş

Cevaplanması gereken bir soru veya açıklanması gereken bir şey varsa, önce mütenasip veriyi toplamak iyi bir yaklaşımdır.

İstatistik bir bilim dalı (ve bir bakıma, aynı zamanda, belki bir sanat dalı veya zanaat dalı) olarak, bir *fenomen* (veya bir *proses*) hakkında *çıkarım*da bulunmak maksadıyla, bu fenomene ait eldeki *veri*den yararlanmayı sağlayan çeşitli fikirleri, metotları ve teknikleri kapsar. Verinin toplanması, düzenlenmesi ve özetlenmesine **betimleyici istatistik** denir. İlgilenilen fenomene dair genel-geçer bir neticeye varmak üzere, sınırlı sayıdaki verinin kullanılmasına ise **çıkarımsal istatistik** denir.

Bunun yanısıra, bir **istatistik**, belirli sayıdaki *gözlem* sonuçlarından türetilen bir niceliktir. En önemli **betimleyici istatistikler** *aritmetik ortalama, ortanca, minimum, maksimum, yüzdelikler, gidim, standart sapma* ve daha başka *ortalama* ve *değişim* ölçüleridir. Bunlar ve benzeri istatistikler çoğunuz için elbette yeni değil, bir kısmını zaten yaşantınızda kullanmışlığınız, en azından duymuşluğunuz vardır muhakkak. *Veri kümelerini* görselleştirmek için kullanılabilen tablolar ve grafikler de, aynı zamanda, betimleyici istatistikler olarak telakki edilebilirler. Tablolar ve grafikler gözlemlerin nasıl *dağıldıklarını* bize göstermelerinin yanısıra, fenomenlerin biribirleriyle olan *ilişkilerini* de görebilmemizi sağlar. **Çıkarımsal istatistikler**, betimleyici istatistikleri girdi olarak kullanır ve belirli sayıdaki gözlem sonuçlarından türetilen bu istatistiklerin *hangi olasılıklarla genelleştirilebileceğini* ifade eder. Mesela *güven aralığı* ve *P-değeri* çıkarımsal istatistiklere misaldir.

Ham verinin taşıdığı enformasyonu daha kolay ve daha hızlı anlaşılabilir ve iletilebilir bir forma dönüştürmek için *betimleyici istatistikleri* kullanırız. Daha küçük bir kısmına bakmak suretiyle, olanaklı tüm veri hakkında bir yargıda bulunmak için *çıkarımsal istatistikleri* kullanırız.

Ancak neredeyse yirminci yüzyılın başına kadar, *istatistik* terimi; asker, doğum, vergi ya da her neye sahipseniz onların sayısı manasına geliyordu. Bir çok durumda, *istatatistik* hala bir şeyin sayısı manasına gelmektedir; Türkiye Cumhuriyeti için en önemli istatistikler *Türkiye İstatistik Yıllığı*'nda yer almaktadır. Fakat söz konusu devletler olsa bile, bir fenomenin tezahürünü tüm olanaklı bireyler üzerinde gözleyebilmek külfetli bir iştir. Bilhassa, *örnek yoklamaları* ve *deneyler* vasıtasıyla topladıkları veriyi kullanarak doğal fenomenleri öğrenmeye ve açıklamaya çalışan bilimciler, kullanabilecekleri zaman, para ve işgücü kaynakları kısıtlı olduğu için, olanaklı tüm bireyler yerine yalnızca bireylerin bazılarından veri elde edebilmek durumunda kaldılar. Ve böylece, örnekten elde ettikleri betimleyici istatistiklerin *popülasyon parametrelerini* ne kadar isabetle *kestirebileceğini* bilmek gayesiyle, bu bilimciler 18. yüzyıldan başlayarak matematikçiler tarafından geliştirilmekte olan *olasılık teorisini* kullandılar.

Veri, hemen hemen her zaman bir *bellisizlik* barındırır. Bu bellisizlik, ölçülen bireyler arasındaki farklılıktan kaynaklanabileceği gibi ölçüm işleminin yarattığı farklılıktan da kaynaklanabilir. Veriden genelgeçer neticelerin çıkarılması, insanoğlunun dünyaya dair bilgisini artırmasında esastır ve tüm rasyonel bilimsel tahkikatın da temelidir. **İstatistiksel çıkarım**, verideki bellisizliğe rağmen, bunu yapabilmemiz için, bize metotlar ve araçlar temin eder. Bellisizliğin veriye nasıl dahil olduğunu açıklayan bir *olasılık modelinin* olanaklılığı, son yüzyılda insanoğlunun yaşadığı bilimsel dönüşümün en önemli unsurudur.

Nihayetinde, *istatistik bilimi, örnekleme* ile elde edilen veriyi kullanarak, bir örneğin çekildiği *popülasyon* hakkında *çıkarımda* bulunmak için çeşitli teknikleri uygulamayı ihtiva eder. Bu iki aşamalı sürecin birinci aşamasında, bir örneği toplar ve *ölçeriz*, sonra, bu örneği görsel ve sayısal olarak *betimleriz*; ikinci aşamada, örnekteki ölçümlere dayanarak *popülasyon hakkında neticeler çıkarırız*.

Örnek için kesin ve kati ifadeler kullanabiliriz. Tam aksine, popülasyondaki tüm bireyleri örneklemediğimizden dolayı, popülasyon için elimizdeki örnekten çıkaracağımız neticeler kesin ve

kati olabilmezler. Fakat, bunlar makul birer olasılıkla kabul edilebilirler. Yani, minimum maliyetle, örnekleme yoluyla, popülasyon hakkında hala nispeten kuvvetli neticelere varabiliriz.

İstatistik biliminin gücü, yalnızca, bir örneği toplamak, ölçmek ve bir populasyon hakkında çıkarımlarda bulunmak için bize bir metodoloji sağlamasında değildir. İstatistik bilimi, aynı zamanda, çıkarımlarımızın ne kadar kesin ve ne kadar kati olacağını ölçmek üzere bize bir ölçü sağlamasından dolayı güçlüdür.

Bilimcinin maksadı doğayı anlamaktır. Bunun için doğal fenomenleri *modellemeye* (yani açıklamaya) çabalar. Gözlemlerine uydurabildiği *modeller* sayesinde, bilimci doğa olayları hakkında yararlı *tahminlerde* bulunabilir. On dokuzuncu yüzyıldan başlayarak, bilimciler *deterministik (belirlemeci)* modellerin *rastgele* doğal fenomenleri açıklamaktaki eksikliklerini gidermek maksadıyla, onları mütenasip *olasılıksal* terimler ile yeniden inşa ettiler ve böylece *istatistiksel (stokastik/olasılıksal) modellere* sahip olduk.

Günümüzde, *istatistik bilimi, bellisizliğin* hüküm sürdüğü her yerde daha doğru kararlar verebilmek için uygulanmaktadır. Bilimsel araştırmaların neticelendirilmesinde, hükümet politikalarına karar verilmesinde ve büyük şirketlerin paraya yön vermesinde *veri betimlemeleri, olasılık kestirimleri, parametre kestirimleri* ve *anlamlılık testleri* vazgeçilmez olmuştur.

Bilimsel Metot

Bilim dünyayı anlamaya dairdir, fakat bir bilgi müktesebatından daha ziyade bir prosestir. **Bilimsel bilgi** bilimsel prosesten öğrendiğimiz şeydir, öyle ki bu proses veri toplamayı ihtiva eder ve bir sistematiğe sahiptir. Bu sistematiğe *bilimsel metot* adı verilir. **Bilimsel metot**, bilgi edinme prosesinin salt *a-priori*, yani deneyimden ârî, bir akıl yürütmeye, sezgiye veya ilhama dayanmasını değil, doğal dünyada yapılan *gözlemlere* (*ve deneylere*) dayanmasını gerektirir.

Modern bilim tipik olarak üç ana kola ayrılır. (1) **Doğa bilimleri**: Mesela fizik, kimya ve biyoloji gibi, fiziksel dünyayı öğrenmeye çalışanlar. (2) **Sosyal bilimler**: Mesela ekonomi, psikoloji ve sosyoloji gibi, bireyleri ve toplumları öğrenmeye-çalışanlar. (3) **Uygulamalı bilimler**: Mesela tıp, tarım ve mühendislik gibi bilimsel bilgiyi pratik maksatlarla kullananlar. Bunların yanı sıra, bilimsel metodu çok veya az benimseyen veya hiç benimsemeyen mesela matematik, lisan, hukuk, tarih, felsefe, güzel sanatlar ve ilahiyat gibi bilgi alanları yine mevcuttur.

Bilimsel araştırma, fenomenleri soruşturmak ve açıklamak için bilimsel metoda göre veri toplamaktır. Bilimsel araştırma her zaman bir soruyla başlar. Mesela "Bu derse kayıtlı öğrencilerin yaşı kaç?" sorusu gibi, "Erzurumda yaşayan insanlar günde kaç bardak çay içer?" sorusu gibi, "İstatistik dersinde başarı düzeyi ve cinsiyet arasında bir ilişki var mı?" sorusu gibi veya "Fasulye bitkisine çiçeklenme döneminde Türk halk müziği dinletmenin, verime etkisi var mı?" sorusu gibi. Daha sonra, bu süreç, hakkında soru sorduğumuz fenomenlere dair gözlemlerimize ve arkaplan bilgimize (literatür) dayanarak bir *hipotezin kurulmasıyla* devam eder. Bir **hipotez** bazen yalnızca herhangi bir fenomene veya çoğu zaman fenomenler arasındaki ilişkiler hakkında bir soruya cevap veren ve kesin olmayan genel-geçer bir ifadedir. Bilimsel araştırmacı kurduğu hipotezi *dizayn ettiği deney* (yani topladığı veri) ile test eder ve bir *neticeye varır*. Elbette ki bazı durumlarda bilimsel soruşturmacı, fenomenler arasındaki ilişkilerin mahiyetine dair önceden belirlenmiş beklentilere sahip olmayabilir, böyle bir durumda, soruşturma, bir hipotezin test edilmesi yerine, fenomenler arasındaki olanaklı ilişki biçimlerini aramaya dönüşür.

İstatistik ve Bilimsel Metot

Deneysel sonuçlara dayanarak bir hipotezin objektif değerlendirmesini yapmak zordur çünkü: (1) akla gelebilen tüm olayları gözlemek olanaklı değildir ve (2) neden-ve-etki yasaları genellikle tam olarak bilinmediği için, içkin değişkenlik mevcuttur. Bilimciler tikel vakalardan daha geniş genellemelere doğru akıl yürütmelilerdir. Bu bellisiz bir çıkarım prosesidir. Bu proses doğru olmayan hipotezleri çürütmemize olanak sağlar, fakat doğru olan hipotezleri ispatlamaya olanak sağlamaz. İstatistik bir bilim olarak tümevarımcı çıkarımların bellisizliğini değerlendirmeye yarayan teknikleri

yaratır, geliştirir ve uygular. İstatistik araştırmacıların veri yığınlarından manalı neticeler çıkarmasını sağlar. İstatistik bilimsel ölçüme uygulanabilen bir araçtır.

Hülasa, bir şeyi (ki bu şey fenomenler veya fenomenler arasındaki ilişkiler olabilir) öğrenmek istediğimiz için onu *gözleriz*. Yaptığımız gözlemler sonucunda *veri* toplamış oluruz. Bu veri, ancak öğrenmek istediğimiz şey hakkında *enformasyon* sağlıyor ise yararlıdır. Fenomenler ve fenomenler arasındaki ilişkiler hakkında sağlanan enformasyon insan zihninde, bu zihnin arkaplan derinliğine bağlı olarak yararlılığı değişen bilgiye dönüşür. Her halükarda, bir *istatistiksel* öğrenme-çalışması, aşağıdaki diyagramda gösterdiğim üzere, bu kavramsal adımları takibederek, ilerler.

Bu bölümde, buraya kadar, istatistik biliminin mahiyetinden, tarihçesinden ve bağlamından kısaca bahsettim. Bu mevzularda geçen istatistiksel terimleri ya **koyu** veya *italik* harflerle yazdım. Bu kelimelerin bazılarının ne manaya geldiğine dair belki bir fikrin olsa bile, kuvvetle muhtemelen pek çoğunun terminolojik tanımlarını bilmiyor olabilirsin. Ama bundan sonrasında, bunlar ve daha başka istatistiksel terimlerin tanımlarını ve altlarında yatan kavramları öğreneceksin. Şunu lütfen aklından çıkarma ki yeni bir bilim öğrenirken onun terminolojisini öğrenmek esansiyeldir. Bunu yeni bir dil öğrenmeye benzetebiliriz. Terminoloji, belli bir bilim, sanat veya faaliyet alanına mahsus spesiyal kelimelerdir (yani terimler). Bir bilim hakkında okuyabilmemiz ve konuşabilmemiz için, önce o bilimin *terminolojisini* öğrenmiş olmamız gerekir. Terimlerin kelimelerden farkı, kesinlikli tanımlara sahip olmalarıdır. Kelimelerin birden çok manaları olabilir ve çoğu zaman biribirlerinin yerine kullanılabilirler. Ama terimler asla.

Kavramlar terimlerle adlandırdığımız gerçekliklerdir. İyi bir öğrenme için, terimlerin tanımladıkları *kavramlara* odaklanmalısın. Bu derste üzerinde durduğum asıl husus, istatistiksel kavramları belli bir hiyerarşi dahilinde, biribiri üzerine bina ederek geliştirebilmektir. Kavramlar, yazılan metin içerisinde dikkatlice geliştirildiğinden, metni düşünerek okumalı ve açıklanan şeyi tam olarak anlamaya çaba göstermelisin. Bir misalin sorusunu ve cevabını okurken bile, misallendirilmek istenen altta yatan kavram hakkında düşünmelisin.

Fenomen

Bilimin uğraşı alanı doğadaki fenomenlerdir. İnsanoğlunun doğada duyularıyla (görme, işitme, tatma, koklama ve dokunma) deneyimlediği her şey bir **fenomendir**. Mesela dokunduğumuz her cisim, gördüğümüz her şekil, işittiğimiz her ses, aldığımız her koku veya tat bir fenomenin tezahürüdür. En nihayetinde insanın kendisi ve davranışları da doğada birer fenomendir. Burada doğal fenomenleri en genel haliyle telakki ediyoruz. Fiziksel, kimyasal ve biyolojik fenomenlerin yanı sıra psikolojik, sosyolojik, ekonomik, kültürel vb. fenomenleri de öğrenmeye-çalışıyoruz. Diğer bir deyişle canlı veya cansız, somut veya soyut tüm varlıklar, oluşlar ve bunların sahip oldukları özellikler, ve bunlar arasındaki ilişkiler, eğer şu veya bu şekilde gözlenebiliyor ise, birer fenomendir.

Bir fenomen mahiyeti itibarıyla *basit* veya *kompleks* olabilir. Basit fenomenlere misal olarak boy uzunluğunu, yıllık geliri, aracın hızını, hastalığın şiddetini, türü, ırkı, mesleği, basıncı, yerçekimini, yaşı ve cinsiyeti verebiliriz. Kompleks fenomenlere misal olarak iklim değişikliğini, sokalardaki suçu ve şiddeti, üniversiteye giriş sınavında öğrencilerin başarısını, depremi, rüzgarı, dalgaları, deja-vuyu, büyümeyi, çekirge istilasını ve sosyal medya içeriklerinin yayılmasını verebiliriz. Basit fenomenler doğrudan gözlenebilir. Fakat kompleks fenomenler sahip oldukları *özellikler* (veya *karakteristikler*) üzerinden gözlenirler.

Bilimsel terminolojide, fenomen öyle bir şeydir ki onun meydana gelişi veya varlığı gözlenebilir. Bu terminolojik manası ve kelimenin halk dilinde kullanıldığı üzere "sıradışı veya fark-yaratan bir şey" manası (mesela "sosyal medya fenomeni" gibi) elbette ki aynı değildir. (Fakat *sosyal medya*

fenomenliği de bir bilimsel fenomen olarak ele alınmak suretiyle, öğrenilmeye-çalışılabilir.)

Gözlem, Çıkım ve Veri

Bir fenomenin tezahürü (veya meydana gelmesi) sonucunda almış olduğu değerin kaydedilmesine **gözlem** denir. Genellikle, bir fenomenin belli bir zamanda ve belli koşullar altında meydana gelen sonucunu belirlemek üzere gözlem yapılır. Mesela *öğrenme* bir fenomendir. İstatistik dersini alan öğrencilerin vize sınavından aldıkları puanları kaydetmek bu fenomene ait birer gözlemdir. Her gözlem bir **çıkımla** sonuçlanır. Kaydedilen bir çıkım, artık **veridir**.

Gözlem ünitesi

Üzerinde gözlem yapılan kişi, nesne veya durumlara **gözlem ünitesi** denir. Meydana gelişi gözlenen fenomenin tipine göre gözlem üniteleri canlı veya cansız, somut veya soyut olabilir. Bir şeyin gözlem ünitesi olabilmesi için ilgilenilen özelliği bakımından nicelenebilir veya en azından nitelenebilir olması gereklidir. Mesela insan, tavşan, ağaç ve bakteri kolonisi gibi canlı gözlem ünitelerinin yanısıra konut, otomobil, rulman, ampul, kaya parçası, toprak numunesi, su numunesi ve 100 mililitrelik serum gibi cansız gözlem ünitelerini sayabiliriz. Diğer yandan işletme, aile, hane, doğum, grev, trafik kazası, müsabaka gibi durumlar da birer gözlem ünitesidir. Gözlem ünitesi aynı zamanda konusu olduğu öğrenme-çalışması bağlamında *birey, vaka, özne veya denek* olarak ta adlandırılabilir.

Ölçme

Öğrenmeye-çalıştığımız fenomeni gözlemeye başladığımız zaman, tezahürünü (yani meydana gelişini) nicel veya nitel olarak betimlerken buluruz kendimizi. Buna **ölçme** deriz. Ölçme, esasen bir tasnif (yani sınıflandırma) işlemidir. Fenomenin tekrar tekrar meydana gelişi (ya da gözlem ünitelerinin sahip olduğu özellik) bazen aynı sınıfta bazen farklı sınıflarda betimlenebilir.

Mesela insanda boy fenomenini gözlediğimizi farzedelim. Boy insanların bir özelliğidir. Ben, insanların boyunu "uzun" ve "kısa" diye iki ayrı sınıfta betimleyerek ölçebilirim. Bir başkası, "uzun," "orta" ve "kısa" diye üç ayrı sınıfta betimleyerek ölçebilir. Bir başkası ise, kendince belirlediği sayıdaki sınıflarda betimleyerek ölçebilir. Veya, genellikle yaptığımız gibi, sen, bir metre kullanarak istediğin hassasiyette (m, dm, cm, mm, µm, vb.) betimleyerek ölçebilirsin. Ölçme işlemini hasaslaştırdıkça, sonsuza giden sayıdaki sınıflardan birinde betimleyebilirsin, gözlediğin insanın boyunu.

Demek ki, bir fenomenin tezahürünü betimlerken, çeşitli sınıflandırma sistemlerinden tercih ettiğimiz birini kullanabiliyoruz. Bu çeşitli sınıflandırma sistemlerinin her birine **ölçek** deriz. Ölçekler misalimizde de gördüğümüz üzere kabadan inceye doğru bir gidişe sahiptir.

Ölçüm Düzeyleri

Bir ölçeğin tipini (yani ölçülen değerin niteliğini) o ölçümün hangi düzeyde yapıldığı belirler. Ölçme işlemi şu dört düzeyden herhangi birinde yapılabilir: *Adlandırma, sıralama, aralık ve oran*. İstatistik bilimini gerçek problemlere uygularken, verinin ölçüm düzeyini bilmek bize hangi istatistiksel prosedürü kullanacağımıza karar vermede yardımcı olur. Bu dersin ileriki bölümlerinde bu ölçüm düzeylerine atıfta bulunacağız elbette, ama buradaki önemli şey sağduyudur: Veriye, ölçüm düzeyi bakımından uygun olmayan hesaplamaları yapmamalı ve istatistiksel metotları kullanmamalısın. Mesela, vatandaşlık numaralarımızın ortalamasını hesaplamak çok manasız olurdu çünkü bu rakamlar kimlik belirtmek maksadıyla kullanılır ve herhangi bir şeyin miktarı veya sayısı değillerdir.

Adlandırma düzeyinde ölçülen veri, yalnızca ad, etiket veya kategorilerden ibarettir. Bu tip veri kendi arasında *manalı bir tertibe imkan vermez*. Yani, adlandırılmış veriyi büyükten küçüğe, iyiden kötüye, azdan çoğa vs. gibi herhangi bir yönde sıralayamayız. Mesela göz rengi, doğum yeri, cinsiyet, politik tercih ve yaprak şekli bu tiptendir. Adlı veri herhangi bir sıralı tertipten ya da sayısal manadan

yoksun olduğu için hesaplamalarda kullanılmamalıdır. Özellikle, bu tip veri kompüterde kodlanırken farklı kategorilere bazen 0, 1, 2, 3 ve 4 gibi sayılar atanabilir; fakat bu sayılar ne gerçekten hesapsal bir mana taşır ne de bunlardan hesaplanacak ortalama gibi ölçülerin bir manası vardır. Yine, kategori adlarını alfabetik bir sırayla tertipleyebiliriz fakat bu da Ahmet'in Mehmet'ten veya kırmızının maviden daha büyük veya daha küçük, daha iyi veya daha kötü, daha az veya daha çok vs. olduğu manasına gelmez.

Sıralama düzeyinde ölçülen veri, bir yönüyle *sıralanabilir*, yani kategorilerin içkin bir sıralı tertibine olanak sağlar fakat kategoriler arasındaki farklar ya belirlenebilir değildir ya da manasızdır. Mesela bir dersten alınan *A, B, C, D, E* gibi harf notları kendi arasında iyiden kötüye sıralanabilir. Ama, mesela *B*'yi *A*'dan çıkartamayız, yani öğrencilerin gerçek puanları arasındaki farkları belirleyemeyiz. İnsan duygu, davranış ve tutumlarını konu edinen öğrenme-çalışmalarında sıralı veri sıkça karşımıza çıkar. Sıralı veri nispi karşılaştırmalar için enformasyon sağlar, farkların miktarı için değil. Genel manada, sıralı veri üzerinde matematiksel hesaplama yapılmamalıdır fakat uygulamada bu kabülün zorlandığı durumlar da çoktur.

Aralık düzeyinde ölçülen veri, şu bakımdan sıralı veriden ayrılır ki herhangi iki veri değeri arasındaki fark belirlidir ve manalıdır. Yani ölçüm birimi tanımlıdır. Bununla beraber, bu düzeydeki veri *doğal bir sıfır noktasına sahip değildir*, yani ölçülen niceliğin nerede yok olduğu (veya başladığı) bilinemez. Mesela 36.3 ^{o}C ve 39.5 ^{o}C'lik vücut sıcaklıkları aralık düzeyinde ölçülmüş birer veridir. Bu değerler sıralanmıştır ve aralarındaki farkın 3.2 ^{o}C olduğunu belirleriz. Bununla birlikte, doğal bir başlangıç noktası yoktur. 0 ^{o}C, bir başlangıç noktası gibi görünse de keyfidir ve ısının tamamen yokluğunu göstermez. Aynı şekilde, 1453 ve 1923 yılları da aralık düzeyindeki veridir. Çünkü zaman miladi 0 yılında başlamamıştır ve 0 yılı, "zamanın olmadığı yeri" gösteren doğal bir sıfır başlangıç noktasından ziyade keyfidir.

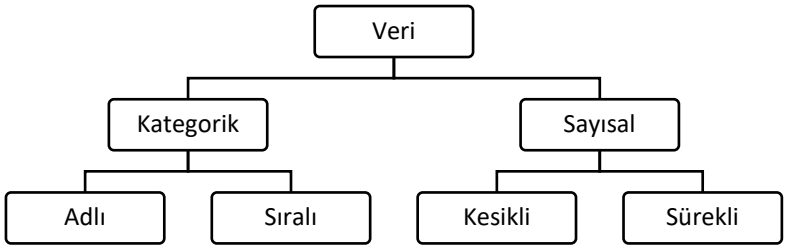
Oran düzeyinde ölçülen veriyi *aralık* düzeyindeki veriden ayıran ekstra özellik ölçülen niceliğin yokluğunu gösteren *doğal bir sıfır başlangıç noktasına sahip olmasıdır*. Bu düzeydeki veri için farklar ve oranlar belirli ve manalıdır. Mesela, bir aracın katettiği km cinsinden mesafe oran düzeyinde bir ölçümdür. Çünkü, 0 km hiç mesafe alınmadığını, 400 ise 200 km'den iki kat uzak mesafeyi gösterir. Diğer yandan ders kitaplarının fiyatlarına baktığımızda, 0 TL bedavayı ve 100 TL de 20 TL'nin dört katı fiyatı gösterir. Bu düzeydeki ölçmenin oran düzeyi diye adlandırılmasının sebebi sıfır başlangıç noktasının gerçekten mevcut olmasının oranları manalı kılmasıdır. Mesela 300 km'lik uzaklık 150 km'lik uzaklığın iki katıdır fakat 40 ^{o}C'lik sıcaklık 20 ^{o}C'lik sıcaklığın iki katı değildir çünkü *derece Celsius* sıcaklık oran düzeyinde değildir. Nicelikler için, birinin diğerinin iki katı olduğunu söylerken, gerçekten "iki katı" olup olmadığına dikkat etmek gerekir.

Sayısal veri ve kategorik veri

Ölçme işleminin mahiyetine göre yani gözlem amacıyla kullanılan ölçeğin hassasiyetine göre verinin iki tipi vardır: Sayısal veri ve kategorik veri. *Sayısal veri* doğal bir sayısal ölçekte kaydedilen ölçümleri kapsar. Mesela bebeğin doğum ağırlığı (gr), ailedeki çocuk sayısı, hastanın kan basıncı (mmHg) gibi. *Kategorik veri*, doğal bir sayısal ölçekte ölçülemeyen ama biribiriyle çakışmayan belirli kategorilerden birine düşen ölçümlerdir. Mesela öğrencinin cinsiyeti, doğumdaki ağrı düzeyi, seçmenin politik parti tercihi gibi.

Kesikli veri ve sürekli veri

Sayısal veri kendi arasında kesikli ve sürekli olmak üzere ikiye ayrılabilir. Değerler tam sayılar ile ifade edildiğinde, *kesikli sayısal veri* ortaya çıkar. Yani olanaklı sonuçlar sadece $0, 1, 2, 3, \cdots$ değerlerini alabilir. Değerler sadece tam sayılar ile değil kesirli sayılar ile de ifade edilebildiğinde, *sürekli sayısal veri* ortaya çıkar. Bu ölçümler sürekli bir ölçek üzerindeki sonsuz olanaklı sonuçlardan birine tekabül edebilir. Mesela dakikada nabız sayısı, ağızdaki çürük diş sayısı ve derinin belirli bir bölgesindeki leke sayısı kesikli sayısal değerler alırken, hava sıcaklığı (^{o}C), kandaki potasyum miktarı (mg/l) ve bel çevresi (cm) sürekli sayısal değerler alır.



Ölçüm düzeylerine göre veri tipleri.

Veri ve Enformasyon

Bilimsel bilgi edinme süreci verinin toplanması ile başlar ki bu safha istatistiksel analizde anlamlı bir rol oynar. Uygulanan ölçüm prosedürünün mahiyetine göre, öğrenilmeye-çalışılan bir fenomenin tezahürüne ait tekrarlanan gözlemlerin kaydedilmesi ile bir *veri kümesi* elde edilir. Esasen **veri** için güzel bir tanım şöyle yapılabilir: Organize edilmemiş (yani ham) halde bulunan rakam, sözcük veya sembollerin bir topluluğudur. Veri kelimesi, tekil veya çoğul manada kullanılabilir. Tekil kullanıldığında, tek bir *gözlem değerini*, çoğul kullanıldığında ise *gözlem değerleri kümesini* ifade eder. Veri, yalnızca bir değerler kümesi olarak, sırf kendi başına manasızdır.

Bir veri kümesi, bilimsel araştırmacının elde etmek istediği *enformasyonu* taşır. Nasıl ki ağaçlar, kendilerinden kağıdın elde edildiği ham materyal ise, benzerce veri de kendisinden enformasyonun elde edildiği ham materyal olarak görülebilir. Toplanan veri mütenasipçe organize edilip, işlendikten sonra, taşıdığı enformasyon açığa çıkarılır. **Enformasyon** için güzel bir tanım şöyle yapılabilir: Veri işlemenin sonucudur ki veriyi bir bağlama oturtur ve manalandırır. Veri enformasyona dönüştürüldüğü zaman, fuzuli detayların külfetinden arındırılarak daha kolay anlaşılabilir ve iletilebilir bir forma kavuşturulmuş olur. Yani, enformasyon mânâlı rafine edilmiş veridir ki başkalarına da yararlı olur.

Kendi başlarına ham gözlemler *veridir*, *enformasyon* veya *bilgi* değillerdir.

İstatistik bilimi, toplanan veriden enformasyon sağlama ve bu enformasyonu yeniden işleme prensiplerinin ve prosedürlerinin bir toplamıdır. Bu prensipleri ve prosedürleri bilmek, herşeyin ötesinde, günlük yaşantımızda bellisizlikle yüz yüze geldiğimiz zaman, zekice kararlar vermemizde bize yardımcı olur.

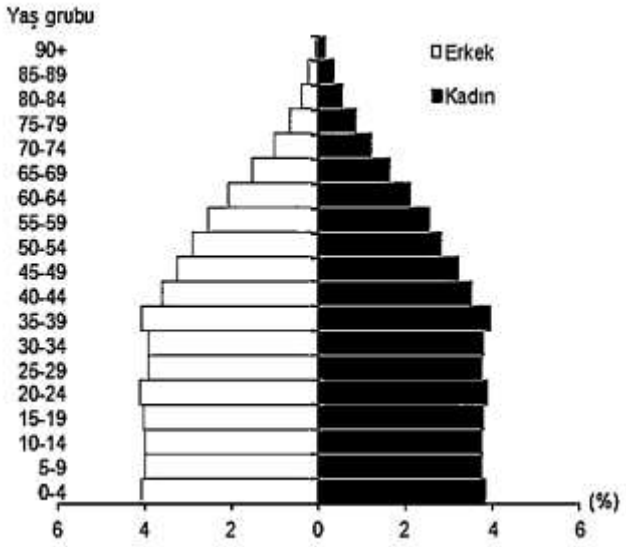
Aşağıda yer alan tabloda bir ilkokuldaki 20 öğrencinin ağırlıklarına ait toplanmış veriyi ve bu verinin taşıdığı enformasyonun bir betimlemesini görüyorsun.

Veri	**Enformasyon**
21 kg, 23 kg, 24 kg, 24 kg, 25 kg, 26 kg, 26 kg, 26 kg, 26 kg, 27 kg, 27 kg, 27 kg, 27 kg, 28 kg, 28 kg, 28 kg, 29 kg, 29 kg, 30 kg, 30 kg	5 bireyin ağırlığı 21 kg'dan 25 kg'a gidiyor, 15 bireyin ağırlığı 26 kg'dan 30 kg'a gidiyor

İstatistiksel enformasyonu sunmanın pek çok yolu vardır, mesela veriyi tabloya veya grafiğe dökmek bunlardandır. Aşağıda Türkiye İstatistik Kurumu (TÜİK) tarafından yayımlanmış istatistiksel enformasyonu görüyorsun ki soldaki figür 2018 yılı Türk nüfusunun toplamı, kadınlar ve erkekler için yaş dağılımını gösteren bir tablo ve sağdaki figür aynı veriye ait nüfus piramidi grafiğidir. Şimdi, bu enformasyonu sağlayan verinin ham halini düşünmeye bir çaba göster. Acaba gözlemler hangi gözlem üniteleri üzerinde yapıldı ve gözlem sonuçları nasıl kaydedildi?

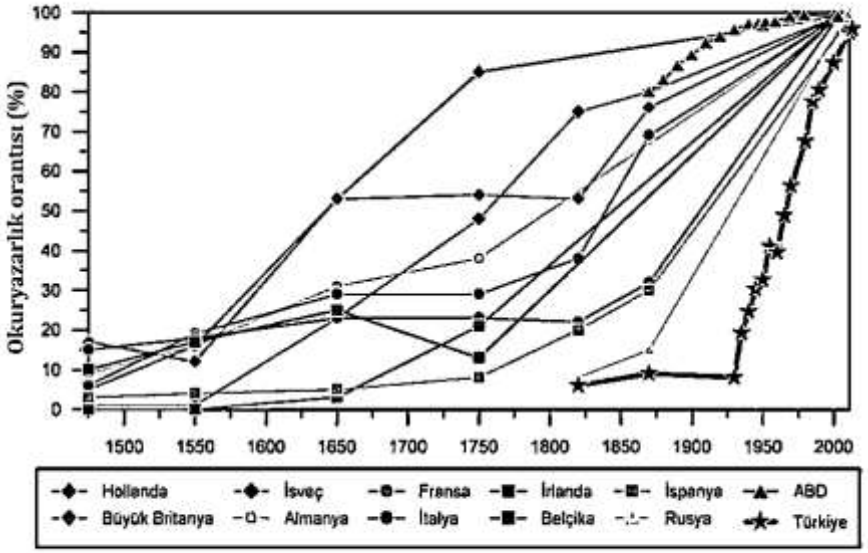
Yaş grubu Age group	Toplam	Erkek	Kadın
Toplam-Total	81 867 223	41 059 075	40 808 147
0-4	6 544 781	3 357 981	3 186 800
5-9	6 336 787	3 253 192	3 083 595
10-14	6 322 223	3 244 584	3 077 639
15-19	6 402 806	3 288 016	3 114 790
20-24	6 523 846	3 333 076	3 190 769
25-29	6 256 609	3 176 593	3 080 016
30-34	6 319 097	3 197 285	3 121 812
35-39	6 563 350	3 309 834	3 253 516
40-44	5 836 212	2 947 638	2 888 574
45-49	5 302 832	2 665 168	2 637 665
50-54	4 693 555	2 367 229	2 326 325
55-59	4 163 996	2 072 110	2 091 886
60-64	3 437 773	1 687 722	1 750 051
65-69	2 604 978	1 242 584	1 362 395
70-74	1 849 910	831 911	1 017 999
75-79	1 257 817	537 620	720 197
80-84	790 992	317 440	473 552
85-89	484 644	181 949	302 695
90-94	146 412	41 254	105 158
95-99	23 611	5 249	18 362
100+	4 990	638	4 352

TÜİK, Nüfus Projeksiyonları, 2018-2080
TurkStat, Population Projections, 2018-2080
Tablodaki rakamlar, yuvarlamadan dolayı toplamı vermeyebilir.
Figures in table may not add up to totals due to rounding.



Aşağıdaki grafik çeşitli ülke halklarının okuryazarlık orantılarının değişimini yıllar boyunca gösteriyor. Bu daha sofistike grafiğin taşıdığı istatistiksel enformasyonu yine kolayca okuyabiliyoruz. Elbette, bu enformasyonu sağlayan verinin ham hali hiç te bu kadar anlaşılabilir olmamalı.

Ülkelerin Okuryazarlık Orantılarının Yıllar boyunca Değişimi



Kaynak: http://www.ourworldindata.org/data/education-knowledge/literacy/
The Cambridge Economic History of Modern Europe: Volume 1, 1700-1870.
İstatistik Göstergeler 1923-2011, Türkiye İstatistik Kurumu

Ham gözlem sonuçlarını kullanılabilir ve paylaşılabilir enformasyona dönüştürmek için veriye dair şu soruların cevaplanması gerekir:

- En müşterek, tipik veya beklenen gözlem hangisidir?
- Gözlemler hangi sınırlar arasındadır?
- Veri nasıl görünüyor?

Veriye dair yukarıda sorduğumuz soruları cevaplamak için kullanabileceğimiz istatistiksel metotların seçimi elimizdeki verinin tipine bağlıdır.

Günlük yaşantımızda, her birimiz veriyle ve veriden çıkarılmış neticelerle, farkında olsak ta olmasak

ta, yüz yüze geliyoruz. Gazeteyi, radyoyu veya televizyonu açtığımız zaman, tüketici fiyatlarına veya enflasyona dair en son istatistiksel enformasyona sahip oluyoruz, işsizlik oranlarını ve bunun çeşitli toplum kesimlerinin yoksulluğu ile ilişkisi hakkındaki bir takım rakamlar veya grafikler önümüze koyuluyor. Yine, borsadaki gelişmelere, boşanma oranlarına, spor liglerindeki son duruma ait raporlar bize ulaşıyor. Dijital medyayı kullanarak, geçmiş yıllara ait enformasyona da çok çabucak ulaşabiliyoruz. Mesela, ülkelerin demografik, ekonomik veya sosyal istatistikler gibi. Bu tip istatistiksel enformasyon ülkelerin istatistik kurumları tarafından bizlere sunulmaktadır. Türkiye İstatistik Kurumu (TÜİK) yayınları ülkemize ait istatistiklerin başlıca resmi kaynağıdır. Bunun dışında, istatistik üreten pek çok yerli, yabancı ve uluslararası resmi ve gayriresmi kaynaklar mevcuttur.

Belirli bir veri kümesinden çıkarılan neticeler çoğu zaman hangi medya organında yayımlandığına bağlıdır ve çoğu zaman spesifik lobi gruplarının tesiri altındadır. Böyle olunca, yurttaşların, kendilerine yararlı neticelere ulaşabilmeleri için, maruz bırakıldıkları enformasyonu filtreden geçirebilmeye ve eleştirel olarak tüketebilmeye ihtiyaçları vardır.

Değişken

Herhangi bir veri kümesi bir grup birey hakkında enformasyon taşır. Bu enformasyon değişkenlerde taşınır. Bir veri kümesi bir veya birden fazla değişkene ait veriyi barındırıyor olabilir. Bir **değişken** bireyden bireye farklılık gösteren ve gözlenebilen (veya ölçülebilen) bir özelliktir. Mesela bir kişinin doğum yeri bir değişkendir. Diğer bir misal bir kişinin yıllık geliridir. Bir çam ağacının boyu hakeza bir değişkendir. Bu değişkenler farklı gözlem üniteleri (yani bireyler) için aynı veya farklı değerler alabilir. Eğer topladığımız veri *değişkenlik* göstermeseydi, tek bir gözlem sonucunda, bir özellik hakkında *neticeye* varmak yeterli olurdu. Ne var ki çoğu özellik bireyden bireye değişir. Buna ilaveten aynı birey için zamandan zamana da değişebilir.

Misal 1 Eğer bir yoklama yapmak suretiyle kişilere cinsiyetlerini, yaşlarını ve ağırlıklarını soracak olursan, bu değişkenlerin bu kişilerde hangi değerleri aldığını belirlemiş olursun. Bu sorulara kişilerin yanıtları senin veri kümeni oluşturur.

Doğal fenomenleri öğrenmeye-çalışan araştırmacılar bu fenomenlerin bazılarını yalnızca bir değişken ile ifade edebildikleri gibi, daha kompleks fenomenleri birden çok sayıda değişken ile ifade edebilirler, öyle ki bu değişkenler o kompleks fenomenin basitleştirilmiş parçalarıdır. Bilhassa bilimsel araştırmacılar kompleks fenomenleri daha iyi anlayabilmek için onları daha kolay ölçülebilir (veya gözlenebilir) parçalara bölmeyi tercih ederler.

Misal 2 İklim değişikliği kompleks bir fenomendir. Bu fenomene ait şu değişkenler mülahaza edilebilir: (1) Deniz seviyesi (2) Sıcaklık (3) Karbon salınımı miktarı (4) Yağış miktarı.

Misal 3 Üniversite giriş sınavında öğrencilerin başarısı bir kompleks fenomendir. Bu fenomene ait şu değişkenleri mülahaza edebiliriz: (1) Sınav puanı (2) Ders-çalışma süresi (3) Öğretmen başına düşen öğrenci sayısı (4) Sınıftaki öğrenci sayısı (5) Öğretmenlerin öğretimsel düzeyi (6) Öğretim tipi (7) Ev ve okul arası mesafe (8) Ebeveynin öğrenciye ayırdığı öğretim desteği süresi.

Değişkenler, kompleks fenomenlerin daha kolay ölçülebilen basitleştirilmiş parçalarıdır. Öğrenmeye-çalıştığımız herhangi bir kompleks fenomene ait gözlem sonuçlarını belirlenen değişkenlerin altında kaydederiz. Basit fenomenlerin kendileri zaten doğrudan ölçülebilen birer değişkendir.

Nicel Değişken ve Nitel Değişken

Nicel değişken, sayısal değerler alır. Tam sayılar veya kesirli sayılar olabilir. Bu sayısal değerler üzerinde aritmetik işlemler yapılabilir. Bu türden değişkenlere *sayısal değerli değişkenler* de deriz. Sayarak veya ölçerek gözleriz.

Nitel değişken, bir bireyi iki ya da daha çok kategoriden birine sokar. Sözel değerler alır. Bu türden

değişkenlere *kategorik değerli değişkenler* de deriz. Kategorize ederek gözleriz.

Yaş (yıl), kandaki alyuvar sayısı, ve kan basıncı (mmHg) gibi nicel değişkenler sayısal değerler alır ki bunlar üzerinde aritmetik işlemler yapabiliriz. Cinsiyet, doğum yeri ve harf notu gibi nitel değişkenler bireyleri basitçe kategorilere sokar ki kategorileri saymak ve/veya sıralamak dışında bir aritmetik işleme tabii tutamayız.

Bir bilimsel araştırmada değişkenleri anlamak, onları tanımlamak ve onları kontrol altına almak veya ölçmek çok önemlidir.

Birden Fazla Değişkene Ait Veri

İstatistik bilimini kullanarak cevaplamaya çalıştığımız önemli soruların çoğu, farklı değişkenler arasındaki ilişkilere dairdir. Mesela tütün içiciliği ve sağlık sorunları arasındaki ilişki, üzerinde en çok tartışılan ve bilimsel araştırma yapılan bir merak konusudur ki uzun süredir, tütün içiciliğinin akciğer kanseri ile *ilişkili* olduğuna inanılmaktadır.

İlişkileri öğrenmeye-çalışabilmek için, birden fazla değişkene ait veri gerekir. Bazı durumlarda ilişkiler iki değişken arasındadır, diğer bazı durumlarda bu değişkenlerin sayısı daha fazla olabilir. Böyle olunca, bir gözlem ünitesi üzerinde birden fazla değişkeni ölçmek icabeder.

Misal 4 Bir medikal bilimsel araştırma grubu rastgele seçtikleri 100 kişinin vücut yağı yüzdelerini ve kan basınçlarını ölçtüler. Bu durumda, iki sürekli sayısal değişken mevcut: vücut yağı yüzdesi ve kan basıncı. Öğrenme-çalışmasına katılan her bireyden her iki değişken için veri toplandı. Yani, her birey için iki veri değeri mevcut ve medikal bilimsel araştırma grubunun kullanacağı veri, bu değer çiftlerinden oluştu. Araştırmacıların çiftler olarak veri değerlerini kaydettikleri tablonun kısmi bir görünümü şöyle idi:

Gözlem ünitesi	Vücut yağı (%)	Kan basıncı (mmHg)
Birey 1	18	139
Birey 2	12	129
Birey 3	6	111
Birey 4	21	138

En temel düzeyde, birden fazla değişkeni öğrenmeye-çalıştığımız zaman, değişkenler arasında bir *çağrışım* ararız. Bu basitçe, bir değişkenin belirli değerlerinin diğer bir değişkenin belirli değerleri ile birlikte vuku bulması (veya ortaya çıkması) manasına gelir.

Mesela, yukarıdaki öğrenme-çalışması, yüksek vücut yağı yüzdesine sahip kişilerin yüksek kan basıncına sahip olmaya yatkın olduğunu göstermeyi amaçlamış olabilir. Eğer durum bunu gösterirse, iki değişkenin birbiriyle çağrışımlı olduğunu söyleyebiliriz. Ancak, bu, iki değişkenin ilişkili olabileceği yollardan yalnızca biridir. Mesela, tütün içmenin akciğer kanserine *neden* olduğu gibi, değişkenler arasındaki ilişkilerin diğer farklı yollarını öğrenmeye-çalışmak için, *kontrol altına alınmış* deneyler dizayn edebiliriz. Bu deneysel sonuçlar şu gibi daha sofistike soruları cevaplamayı gerektirebilir:

- Hangi değişkenler en münasebetli?
- İki deneyin çıkımları arasında farklar var mı?
- Bu farklar gerçek mi yoksa verideki parazitin bir sonucu mu?

Yanıt Değişkeni ve Açıklayıcı Değişken

Bir istatistiksel deneyin sonucunda gözlediğimiz bir değişkenin değerlerindeki değişikliğe bir başka değişkenin neden olduğunu veya en azından söz konusu değişikliği açıkladığını düşünüyorsak, bu değişkenlerden ilkine **yanıt değişkeni** ve ikincisine **açıklayıcı değişken** adı verilir. Pek çok istatistiksel deneyin maksadı, yanıt değişkeninin aldığı değerlere açıklayıcı değişkenin etkisini değerlendirmektir. Daha karmaşık ilişkilerde, birden çok sayıda yanıt değişkenleri veya açıklayıcı değişkenler bulunabilir.

Veri Kaynakları

Veri kümeleri başlıca iki kaynaktan sağlanabilir. Bir veri kaynağı birincil veya ikincil olabilir. Bu ikisi arasındaki ayrım veriyi kullanan kişinin aynı zamanda veriyi toplayan kişi olup olmadığına bağlıdır. Belirli soruları cevaplamak ve bir takım neticelere varmak için veri kullanan kişiye *veri analizcisi* denir. Veriyi toplayan kişiye *veri toplayıcısı* denir. Eğer veri analizcisi aynı zamanda veri toplayıcısı ise, o zaman veri *birincil kaynaktan* sağlanmış olur. Eğer veri analizcisi bir başkası tarafından toplanmış veriyi kullanıyor ise, o zaman veri *ikincil kaynaktan* sağlanmış olur.

Çoğu durumda, veri en kolay ikincil kaynaklar yoluyla elde edilir. Devlet kurumları ve başka kurumlar çoğu zaman kendi verilerini kamuya açarlar. Bu taktirde, bir veri analizcisi ihtiyaç duyduğu veriyi bu kurumların kaynaklarını kullanmak suretiyle toplayabilir, elbette bunu rapor etmesi koşuluyla.

Öte yandan, ikincil veri kaynaklarının mevcudiyetinin olanaklı olmadığı pek çok durum vardır. Popülasyonların yapısı zamanla ve farklı koşulların tesiriyle değişebilmektedir. Bilhassa bilimcilerin üzerinde çalıştıkları mevzulara dair *hipotezlerini test etmek* için birincil kaynaktan veri toplamaları gereklidir. Değişkenler arasındaki neden-ve-etki ilişkilerini açığa çıkarmak istedikleri zaman, bu veriyi *dizayn-edilmiş deneyleri* yürüterek elde edebilirler ancak.

Popülasyon ve Örnek

Popülasyon, öğrenme-çalışmamıza konu olan özelliğe sahip tüm bireylerin kümesidir. Bireyler; insanlar, hayvanlar, bitkiler, nesneler ve durumlar olabilir. Yani daha başka bir ifadeyle söylersek, popülasyon denildiği zaman bir değişkenin alabileceği tüm olanaklı sonuçların kümesi tanımlanmış olmaktadır. İstatistiksel manada, bir veri kümesidir, bir popülasyon. Ve bu veri kümesi bilimsel araştırmacının sağlamak istediği *enformasyonu* taşır.

İstatistiksel bir popülasyonu tanımlamanın başat koşulu popülasyonu oluşturan bireylerin durum, zaman veya mekan boyutunda spesifik bir tanımlarının yapılmış olması gerekliliğidir. Mesela Türkiye'de yaşayan tüm insanlar, belirli bir hastalığı taşıyan tüm bireyler, Doğu Anadolu Bölgesi'nin tüm köyleri, Erzurum'da yetişen kavak ağaçları ve yirmi birinci yüzyılda meydana gelen tüm tren kazaları birer popülasyondur.

Bir popülasyonu meydana getiren tüm bireylerin kümesi, çoğu durumda çok geniş ya da sonsuz genişlikte olabileceğinden, bilimsel araştırmacı bu kümenin sadece bir alt kümesini *örneklemekle* yetinmek durumunda kalacaktır. **Örnek**, popülasyonun herhangi bir alt kümesidir.

Sonraki bölümlerde öğreneceğin üzere, veri toplamak için değişik metotlara başvurabiliriz. Hangi veri toplama metodunu kullanmış olursak olalım, elimizdeki veri, bir popülasyonun bir *örneği* olacaktır. Eğer bu örnekten sağladığımız *enformasyonu*, örneği çektiğimiz popülasyona genellemek niyetindeysek, örneğin, çekildiği popülasyonu *temsil* kabiliyetine sahip olması zorunludur.

Hedef Popülasyon

Bir veri toplama prosesi, veri toplamanın niçin ve kimden yapılacağını betimleyen ereklerin tesbitiyle başlar. Veri toplayıcısının erekleri, toplanmasına ihtiyaç duyulan veri hakkında pek çok şey söyler. Hedef popülasyonun belirlenmesinde bu erekler rol oynar. Mesela, farzedelim ki okuduğun üniversitede mezuniyet partisi düzenleme komitesi başkanısın ve bu partinin planlanmasında kullanılmak üzere öğrencilerden enformasyon temin etmek istiyorsun. Bu genel amaç çerçevesinde, şunlar gibi daha rafine bazı erekler tespit edebilirsin: (1) Partinin başarılı olmasını sağlayacak faktörleri belirlemek için öğrencilerden enformasyon devşirmek, ki burada "başarı" kriterleri olanaklı en çok sayıda öğrencinin partiye katılması ve katılanların beklentilerinin karşılanmasıdır. (2) Düzenleme komitesi tarafından kullanılacak yararlı verinin elde edilmesi.

Bir veri toplama planı; hedef popülasyonu, veri gereksinimlerini ve ölçülecek değişkenleri açıkça betimlemek suretiyle; tespit edilen ereklere nasıl ulaşılacağını gösterecektir. "Hangi öğrenciler?"

sorusunu sormak hedef popülasyonun tanımlanmasına yardımcı olacaktır. Mezuniyet partisi düzenleme komitesi başkanı olarak, muhtemelen bu yıl mezun olacak öğrencilere sorularını sormak isteyeceksin. Yani, bu yıl itibarıyla okudukları bölümlerin son sınıfında olan öğrenciler. Fakat, bazı son sınıf öğrencileri çeşitli nedenlerle bu yıl mezun olamayacağı için, onlara danışmaya gerek yok. Yine, mezuniyetini uzatmış ve bu yıl mezun olacak daha eski öğrenciler olacağı için, onlara da danışmak gerek. Buna göre, hedef popülasyonunu şöyle tanımlayabilirsin: "Üniversitenin bu yıl itibarıyla mezun olacak öğrencileri."

Temsilci Örnek

Bir *temsilci örnek,* hedef popülasyonun sahip olduğu tipik özellikleri aynıyla yansıtır. Mesela seçimlerden önce yapılacak bir halkoyu yoklamasını düşünelim. Varsayalım ki halkoyu yoklaması Türkiyede'ki 55 milyon kayıtlı seçmenden iktidardaki partiyi destekleyenlerin yüzdesini kestirmeyi hedefliyor. İktidardaki partinin üyesi seçmenlerin oluşturacağı bir örnekten toplanacak yoklama verisinden yapılacak kestirime dayanacaksa, bu halkoyu yoklaması akıl işi olmaktan uzaktır elbette. Böylesi bir kestirim, neredeyse hiç kuşkusuz yükseğe *yanlı* olacaktır ve nihayetinde pek güvenilir olmayacaktır. Bir temsilci örneğin, her yönüyle popülasyonun küçük bir benzeri (veya modeli) olması gereksinimi, seçim prosedürünü ziyadesiyle zorlaştırmaktadır.

Rastgele Örnek

Temsilci örnek gereksinimini karşılamanın en müşterek yolu bir *rastgele örnek* çekmektir. Bir *rastgele örnek*, populasyondan aynı hacimdeki her alt kümenin seçilme şansının eşit olduğunu bize garanti eden örnektir. Eğer yoklama, popülasyondaki 55 milyon seçmenden 900'ünü örnekliyorsa ve her 900 seçmenlik alt kümenin seçilme şansı eşitse bu bir rastgele örnektir. Rastgele bir örnek çekmenin en basit yolu popülasyonu oluşturan tüm bireyleri bir torbaya koyup, karıştırıp, şansa bağlı olarak birer birer çekmektir. Elbette hemen hemen hiç bir popülasyon torbaya sığmayacağına göre şansa bağlı çekim işlemini, popülasyonu oluşturan bireylerin bir listesi üzerinde yapmak gerektiği açıktır.

Şunu iyi anlamalısın ki bir popülasyondan bir örneği rastgele çekmek, sana bir temsilci örnek garanti etmez; çektiğin örnekten popülasyona genelleştirme yaparken, bir olasılık tayin edebilmene olanak sağlar.

Örnekleme Ünitesi

Popülasyondan seçilerek örneğe girebilecek her gözlem ünitesi bir *örnekleme ünitesi* olabileceği gibi gözlem ünitelerinin bir gurubu da bir örnekleme ünitesi olabilir. Hangi örnekleme ünitesini kullanacağınızı belirlerken çoğu zaman öğrenme-çalışmanızın ölçmek istediği şeyi ve örneklemek istediğiniz popülasyon hakkında elde etmek istediğiniz veriyi göz önüne alacaksınız. Mesela sosyoekonomik bir öğrenme-çalışmasında bir örnekleme ünitesi bir kişi olabileceği gibi bir hane, bir apartman, bir sınıf, bir işletme veya bir köy olabilir. Sonrasında bu örnekleme ünitesinin kapsadığı gözlem üniteleri üzerinde ölçüm yapılır. Türkiye İstatistik Kurumu, vatandaşlarının mesela aylık kazancını öğrenmek için ülkedeki hanelerin bir örneğini çeker. Böyle olunca, hedef popülasyondaki haneler örnekleme üniteleri olarak hizmet eder ve örneklenen hanelerde ikamet eden kişiler de gözlem üniteleridir. Bir entomoloji öğrenme-çalışmasında, bir dönümlük parselleri örnekleyebiliriz ve ondan sonra, örneklenen parsellerde yer alan bireysel bitkilerin üzerindeki böceklerin sayısını ölçebiliriz. Bu halde, örnekleme ünitesi parseldir ve bireysel bitkiler gözlem üniteleridir.

Örnekleme Çerçevesi

İstatistik-bilimi açısından bakıldığında, popülasyonu belirleyen, bilimsel araştırmanın amacıdır. Bilimsel araştırmacı örnekten elde edeceği enformasyonu nereye kadar genellemek istiyorsa, hedeflediği popülasyonun hacmi de o kadar genişleyecektir. Bilimsel araştırmacının hedeflediği populasyon, bilimsel araştırmanın örnekleme çerçevesini de belirler.

Örnekleme çerçevesi, örneğin fiilen içinden çekileceği örnekleme ünitelerinin bir listesidir. İdeal olanı bu çerçevenin popülasyondaki her örnekleme ünitesini listeliyor olmasıdır fakat pratikte bu o kadar da kolay değildir. Örnekleme çerçevesinin doğru tespit edilebilmesi, popülasyonun örnekteki temsili ile alakalı olduğundan, rastgele örnekten populasyona yapılacak çıkarımı etkiler.

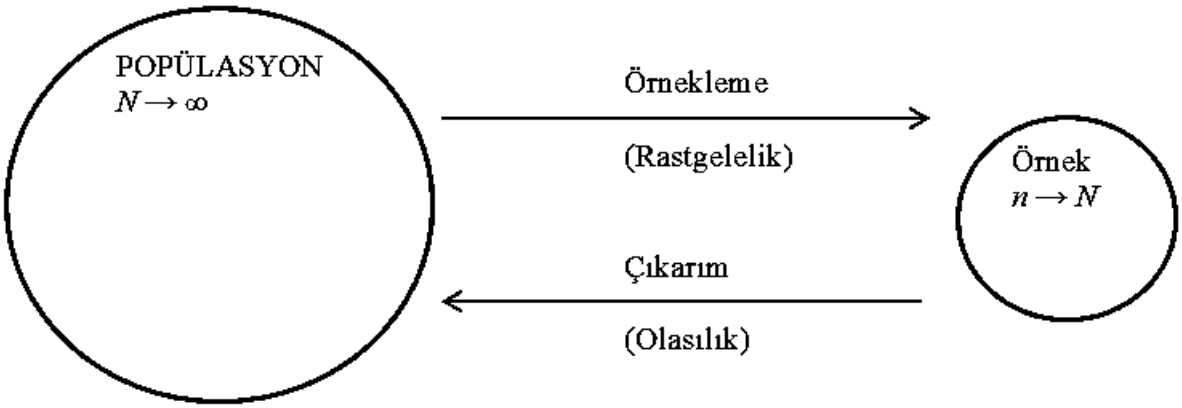
İstatistiksel Çıkarım

Şimdiye kadar istatistiksel çıkarım probleminin beş elemanından dördünü tanımış olduk: Bir popülasyon, ilgilenilen bir veya birden çok sayıda değişken, bir örnek ve bir çıkarım. Fakat çıkarımı yapmak sadece hikayenin bir kısmıdır; yaptığımız çıkarımın aynı zamanda güvenilirliğini de bilmek zorundayız. Yani, yaptığımız çıkarım ne kadar iyi? Bir popülasyon hakkında yapacağımız bir çıkarımın doğruluğundan emin olabilmemizin yegane yolu, popülasyonun tamamını örneğimize dahil etmektir. Bununla birlikte kaynak yetersizliklerinden (yani yetersiz para, zaman, emek vb.) dolayı çoğu zaman (neredeyse hiç bir zaman) popülasyonun tamamıyla çalışamayız. (Yani örnek hacmi hiç bir zaman popülasyonun hacmine eşit olmaz.) Bu sebeple, çıkarımlarımızı popülasyonun sadece bir kısmına (bir örneğine) dayandırırız. Böyle olunca, çıkarımlarımız bir bellisizlik unsuru taşımaktan kurtulamaz. Bu sebeple, her ne zaman olanaklıysa, yapılan her çıkarımın güvenilirliğini belirlemek ve rapor etmek önemlidir. Güvenilirlik, öyle ise, çıkarımsal istatistik probleminin beşinci elemanıdır.

Güvenilirlik ölçüsü, bir istatistiksel çıkarımın barındırdığı belllsizliğin derecesine dair nicel bir ifadedir. Bir çıkarıma eşlik eden güvenilirlik ölçüsü, istatistik bilimini falcılıktan ayırır. Bir falcı da, tıpkı bir istatistikçi gibi, bir örneği inceleyebilir (ki elimiz) ve popülasyon (ki yaşantımız) hakkında çıkarım yapabilir. Bununla birlikte, istatistiksel çıkarımda olduğunun aksine, falcının çıkarımına herhangi bir güvenilirlik ölçüsü dahil değildir.

İstatistiksel çıkarım, bir örneğin taşıdığı enformasyona dayanarak bu örneğin çekildiği popülasyon hakkında bir kestirim, tahmin veya diğer bir genellemedir. Bu genellemenin spesifikliği, olasılığı belirli bir genelleme olmasıdır. Yani daha dar hacimdeki örneğin taşıdığı enformasyonu kullanarak daha geniş hacimli popülasyonu öğreniriz. Ve bu öğrenmenin doğruluğu veya yanlışlığı hakkında matematiksel bir kesinliğe sahip oluruz. Çünkü bu süreçte istatistik bilimi matematik biliminin olasılık teorisinden yararlanır. Bir istatistiksel çıkarıma dair bellisizliğin derecesi hakkında her zaman nicel bir ifade olarak bir güvenilirlik ölçüsü mevcuttur.

İstatistiksel çıkarım sürecini aşağıdaki şekilde özetleyebiliriz. Bilimsel araştırmacı popülasyondan rastgele bir örnek çeker. Örnekten elde ettiği enformasyonu popülasyona, belirli bir olasılıkla geneller. İstatistik bilimi notasyonuna göre popülasyonun hacmini N ile ve örneğin hacmini de n ile gösteririz.



Dikkat edersen, bu sürecin bir tümdengelim ve tümevarım döngüsü yarattığını görebilirsin. Akabinde, bilimsel araştırmalar bir kez yapılmakla da kalmaz. Farklı bilimciler tarafından yinelenerek yapılır aynı bilimsel araştırma ve bu sayede bulguların doğruluğu ve güncelliği sağlanmaya çalışılır.

Misal 5 ABC kablo TV kanalı, abonelerinin yaşını merak etmekteydi. Bir istatistikçinin yardımıyla,

abonelerinden 600'ünü rastgele örneklediler ve herbirinin yaşını belirlediler.

a. Örnekleme ünitesini betimle
b. Gözlem ünitesini betimle.
c. Hedef populasyonu betimle.
d. Örnekleme çerçevesini betimle.
e. İlgilenilen değişkeni betimle.
f. Örneği betimle.
g. Yapılmak istenilen çıkarımı betimle.

Cevap:

a. ABC kablo TV kanalının abonesi olan her kişi bir örnekleme ünitesidir.
b. ABC kablo TV kanalına abonesi olan her kişi bir gözlem ünitesidir.
c. Hedef popülasyon, yaşı merak edilen, ABC kablo TV kanalının tüm aboneleridir.
d. Örnekleme çerçevesi ABC kablo TV kanalının tüm abonelerinin listesidir.
e. Abonelerin yaşı ilgilenilen değişkendir.
f. Örnek, hedef popülasyonun bir alt kümesi olmalıdır. Bu durumda ABC TV kanalı aboneleri arasından seçilen 600 kişidir.
g. Yapılmak istenilen çıkarım 600 aboneden oluşan örneğin taşıdığı enformasyonu ABC kablo TV kanalının tüm abonelerinden oluşan popülasyona genellemeyi ihtiva eder. (Örnekteki abonelerin yaşları ABC kablo TV kanalı abonelerinin hangi yaşlar arasında olduğunu veya yaşlarının aritmetik ortalamasını kestirmede kullanılabilir.)

Misal 6 Trakya Bölgesindeki ineklerin ortalama ağırlığını bulmak için bir örnekleme yapılıyor. Bölgedeki tüm çiftliklerin bir listesi mevcut ve bu listeden rastgele 30 çiftlik seçiliyor. Sonra, seçilen 30 çiftlikteki her bir ineğin ağırlığı kaydediliyor. Ağırlığı ölçülen tüm ineklerin sayısı 532.

a. Örnekleme ünitesini betimle
b. Gözlem ünitesini betimle.
c. Hedef populasyonu betimle.
d. Örnekleme çerçevesini betimle.
e. İlgilenilen değişkeni betimle.
f. Örneği betimle.
g. Yapılmak istenilen çıkarımı betimle.

Cevap:

a. Çiftlik.
b. İnek (yani, dişi sığır ki en az bir kez buzağılamış olsun).
c. Trakya bölgesindeki tüm inekler.
d. Trakya bölgesindeki tüm çiftliklerin listesi.
e. İneklerin ağırlığı.
f. Ağırlığı ölçülen 532 inek.
g. Beşyüz otuz iki inekten oluşan örneğin ortalamasını hesaplayarak Trakya Bölgesinde mevcut inek popülasyonun ortalamasını kestirmek.

Parametre ve İstatistik

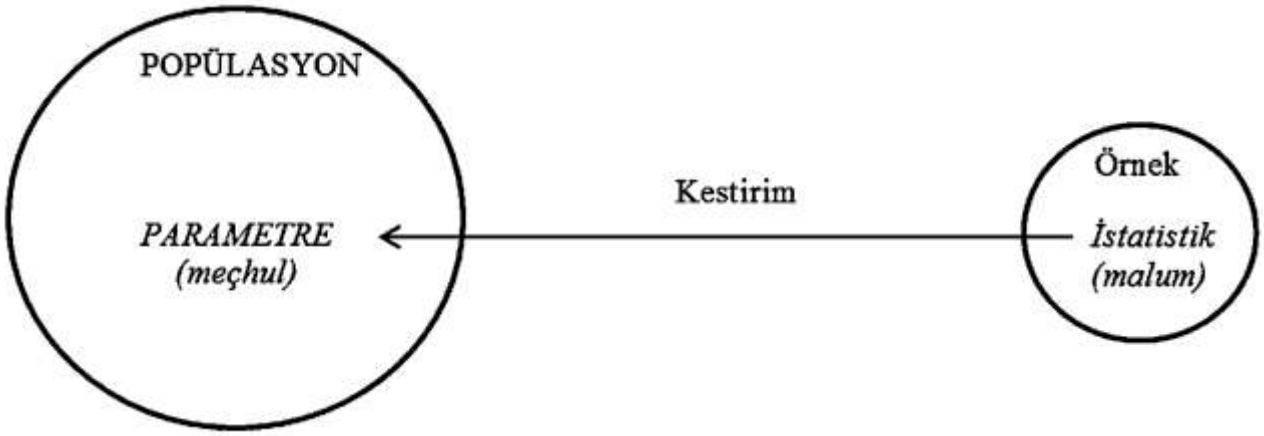
İlgilendiğimiz özelliğe sahip tüm gözlem ünitelerini ölçebilseydik, elde edeceğimiz veri ki buna popülasyon diyoruz, kesinliği olan bir enformasyonu taşıyor olacaktı. Popülasyonun nüfusu sayısınca veri tarafından taşınan enformasyonu elde edebilmek için bu verinin *özetlenmesi* veya başka bir deyişle *indirgenmesi* yani *betimlenmesi* gerekir. İşte popülasyonu oluşturan olanaklı tüm sonuçların birbirinden farklılık gösteren değerlerini daha anlaşılır/yararlanılabilir bir tarzda özetleyen veya indirgeyen yani betimleyen sayısal ölçüler **popülasyon parametreleri** diye adlandırılır. Bilimsel araştırmacının değişkenleri ve aralarındaki ilişkileri öğrenmeye çalışırken esasen peşinde olduğu da bu popülasyon parametreleridir. Bunlara misal olarak *aritmetik ortalama, ortanca, orantı, varyans, çeyreklikler, korelasyon katsayısı, regresyon doğrusunun kesimi ve eğimi* verilebilir.

Popülasyon parametreleri, farklı değerlere dağılan veriyi, genel manada bir merkezi değerle ve bu

merkezi değerden sapmaları da başka bir merkezi değerle betimler. Parametrik ölçülere kolayca ulaşabilseydik, hayat nispeten daha kolay olabilirdi belki; ancak, bir değişkenin olanaklı tüm sonuçlarına ulaşmak, yani popülasyonu oluşturan tüm gözlem ünitelerini ölçmek, pek çok zaman olanak dışıdır. Çünkü bilimsel araştırma için harcanabilecek kaynaklar sınırlıdır. Bu sebeple, pek çok zaman *popülasyon parametreleri meçhuldür*.

Örnekleme fikri, bütün hakkında enformasyon sağlamak üzere bütünün bir parçasını öğrenmeye-çalışmaktır. Popülasyona ulaşamadığımız durumda, örneğe giren sonuçların biribirinden farklılık gösteren değerlerini betimleyen sayısal ölçüleri hesaplarız. Örnekteki veriyi betimlemek için hesaplanan bu sayısal ölçüler, **örnek istatistikleri** diye adlandırılır. Örnek istatistikleri popülasyonun meçhul parametreleri için birer *kestirimdir*. Mesela Türkiye İstatistik Kurumu (TİK) aylık ortalama enflasyon oranını kestirmek için Türkiye pazarında alınıp satılan bütün mal ve hizmetleri temsil etmek üzere "enflasyon sepeti" adı verilen bir örnek çeker. Bu örnekten hesaplanan ortalama aylık enflasyon oranı bir istatistiktir ve mal ve hizmetlerin tümünden oluşan popülasyonun meçhul parametresi için bir kestirimdir.

Bir örnekteki ölçümlere dayanarak bir popülasyon hakkında bir takım iddilarda bulunduğumuz zaman, esasında parametrelere ve istatistiklere atıf yaparız. Özetlersek, bir parametre, bir popülasyonun bir özelliğini betimleyen bir sayısal ölçüdür. Bir istatistik, bir popülasyondan çekilen bir örneğin aynı özelliğini betimleyen bir sayısal ölçüdür. (Açık ki eğer örnek hacmi popülasyon hacmine eşit olsaydı, bunlar biribirine eşit olacaktı.) Bu manada, istatistikler enformasyon temin etmenin en genel yoludur. Aşağıdaki figür istatistiksel kestirim kavramını resmediyor.



Misal 7 Erzurum ilinde hane başına yıllık enerji tüketimi nedir acaba? Sağlamaya çalıştığımız enformasyon Erzurum ilindeki tüm hanelerin yıllık enerji tüketimi verisinde barınmaktadır. Erzurum ilindeki 135087 hane bizim popülasyonumuzdur. Bu popülasyonun (minimum, maksimum, ortalama, varyans gibi) parametreleri de sağlamaya çalıştığımız enformasyondur. Popülasyonun tamamı için sayım yapmaya gücümüz yetmediği zaman, daha az sayıda haneyi rastgele örnekleyeceğiz. Popülasyon parametrelerinin birer kestirimi olarak kullanmak üzere örnek istatistiklerini hesaplarız.

Misal 8 Daha yukarıda, bir ilkokuldaki 20 öğrencinin ağırlıklarına ait toplanmış veriyi ve o verinin taşıdığı enformasyonun bir betimlemesini bir tabloda görmüştük. Şimdi biliyoruz ki o enformasyonu parametreler veya istatistikler cinsinden de ifade edebiliriz. Ama önce, bu verinin bir popülasyon verisi mi yoksa bir örnek verisi mi olduğuna karar vermemiz lazım.

Şöyle ki, hedef popülasyonun sınırlarını bilimsel araştırmacının niyetinin çizdiğini öğrenmiştin. Yani, ağırlıkları ölçülen öğrenciler, bu veriyi toplayanın niyetine göre, bir popülasyon olarak telakki edilebilecekleri gibi, bir örnek olarak ta telakki edilebilirler. Eğer veri toplayıcısının niyeti yalnızca 20 öğrenciden oluşan, belirli bir sınıftaki öğrencilerin ağırlığını öğrenmeye-çalışmak idiyse, o sınıftaki öğrencilerin tümü, yani 20 öğrenci, elbette bu öğrenme-çalışması için tanımlanmış bir popülasyondur. Veri toplayıcısı hedef popülasyonun tüm bireylerini gözlemiştir. Öyleyse, bu veri bir

popülasyon verisidir. Eğer veri toplayıcısının niyeti belirli bir ilkokuldaki öğrencilerin ağırlığını öğrenmeye-çalışmak idiyse, o ilkokuldaki öğrencilerin tümü, bu halde, elbette bu öğrenme-çalışması için tanımlanmış bir popülasyondur. Veri toplayıcısı bu popülasyonun tüm bireylerini gözlemek yerine, o ilkokuldaki öğrenciler içerisinden 20 sini örneklemiştir ve gözlemiştir. Öyleyse, bu veri bir örnek verisidir.

Bilimsel araştırmacı, peşinde olduğu enformasyona ulaşmak için, bu veriyi bazı sayısal ölçüler ile de özetlemiştir. Bu sayısal ölçülerden birisi aritmetik ortalamadır ve değerini 26.55 kg olarak hesaplamıştır. Eğer bu bir popülasyon verisiyse, hesaplanan bu değer popülasyon parametresinin değeridir; yani belirli bir sınıftaki öğrencilerin ağırlığıdır. Eğer bu bir örnek verisiyse, hesaplanan bu değer örnek istatistiğinin değeridir ve popülasyon parametresinin bir kestirimidir; yani örneklenen öğrencilerin ağırlığıdır ve ait oldukları ilkokuldaki tüm öğrencilerin ağırlığının bir kestirimidir.

Veri	Enformasyon	Parametre/İstatistik
21 kg, 23 kg, 24 kg, 24 kg, 25 kg, 26 kg, 26 kg, 26 kg, 26 kg, 27 kg, 27 kg, 27 kg, 27 kg, 28 kg, 28 kg, 28 kg, 29 kg, 29 kg, 30 kg, 30 kg	5 bireyin ağırlığı 21 kg'dan 25 kg'a gidiyor, 15 bireyin ağırlığı 26 kg'dan 30 kg'a gidiyor	Aritmetik ortalama ağırlık = 26.55 kg

İstatistiklerin parametreleri kestirmedeki isabetliliği örnekleme veya veri toplama metodu ile doğrudan ilgilidir. Eğer örnekleme doğru metotla yapılmaz ise, yani örnek popülasyonu iyi temsil edemezse, örnek istatistikleri popülasyon parametrelerinin isabetli kestirimleri olamazlar. Mamafih, örnekleme doğru metotla yapılmış olsa dahi, istatistiğin parametreye isabet edip etmediğini hiç bir zaman bilemeyiz, ancak bunun olasılığını kestirebiliriz, ki bu da ancak örneklemenin rastgele yapılmış olmasıyla mümkündür. Bazı popülasyon parametrelerine ve mütekabil örnek istatistiklerine ait notasyonu aşağıdaki tabloda görebilirsin.

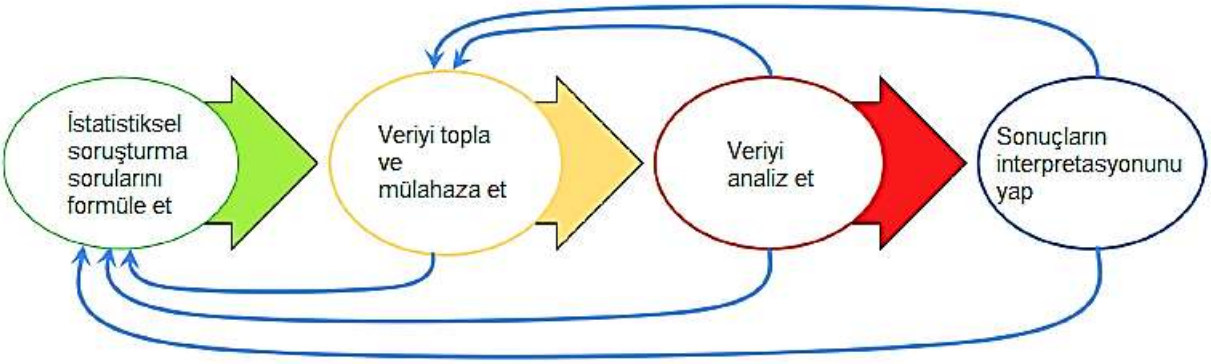
Sayısal betimleyici	Popülasyon parametresi	Örnek istatistiği
Aritmetik ortalama	μ	$\bar{y}$
Varyans	σ^2	s^2
Standart sapma	σ	s
Binomiyal orantı	π	p
Korelasyon	ρ	r
Regresyon doğrusunun eğimi	β	b
Regresyon doğrusunun kesimi	α	a

Hülasa, bize, ilgilendiğimiz popülasyon hakkında, elimizdeki veri kümesine dayanan olasılık ifadeleri temin eden istatistiksel çıkarım adını verdiğimiz sonuçların iki tipi vardır. (I) **Kestirim**: ilgilendiğimiz parametreyi kestirmek için çektiğimiz örneğin barındırdığı enformasyonu kullanmak. Mesela, iktidar partisinin ülke çapındaki güncel oy orantısını kestirmek için yapılacak bir örnek yoklamasının sonucunu kullanmak. (II) **Anlamlılık testi**: İlgilendiğimiz parametre hakkındaki belli bir ifadenin doğru olup olmadığını belirlemek için örneğin barındırdığı enformasyonu kullanmak. Mesela, iktidar partisi ülke çapındaki güncel oy orantısınının %75 ten fazla olduğunu iddia etmektedir. Bu iddianın yapacağımız örnek yoklamasına ait veri tarafından desteklenip desteklenmediğini belirlemek istiyoruz. Anlamlılık testlerine başka kaynaklarda, aynı zamanda, *hipotez testleri* de denildiğini görürsün. Kabaca bu iki terim biribirinin yerine kullanılıyor olsa dahi, teknik olarak farklıdır. Kısaca, burada, hipotez testinin, anlamlılık testine göre, daha geniş bir kavramı tanımladığını söylemekle yetineceğim.

İstatistiksel Problem-Çözme Prosesi

İstatistiksel problemleri nasıl çözeceğini öğrenirken, şunun farkında olmalısın ki istatistik dört bileşenli bir prosestir: (1) İstatistiksel soruşturma sorularının formülasyonu. (2) Verinin toplanması

ve mülahazası. (3) Verinin analizi (4) Sonuçların interpretasyonu.



Spesifik formüllere ayrı ayrı yoğunlaşmaktan daha ziyade, bu prosesi bir bütün olarak kavrayabilen her öğrenci, yaşama dair problemleri çözmek için istatistiksel düşünebilmeyi ve büyük resme bakabilmeyi başaracaktır.

SAS istatistiksel yazılım şirketinin eş-kurucusu John Sall'ın JMP bloğunda "İstatistik Niçin Esansiyeldir?" başlığıyla yayımlamış olduğu bir yazıdan bir bölümün çevirisi:

"*İnsanların çoğu, istatistik biliminin ne kadar esansiyel olduğunun farkında değildir. Günlük yaşantımız, oysa, istatistik biliminin ürünleri ile çevrilidir. Dişini fırçalıyorsun, mesela. Diş macunundaki florid belirli öğrenme-çalışmasının eseridir. Bu öğrenme-çalışmaları, güvenilir ve etkili bir içeriği ve münasip konsantrasyonu hassas bir şekilde sağlamak için istatistiksel metotları kullanan bilimcilerce yürütülmüştür. Diş macununun formülasyonu, istatistiksel modelleme vasıtasıyla optimal formülasyonu belirleyebilmeyi sağlayan bir seri dizayn-edilmiş deneyin bir sonucudur. Diş macunu üretimi, belirli bir kaliteyi ve bu kalitenin tutarlılığını garanti etmek ve değişkenliği azaltmak için istatistiksel kontrol süreçleri kullanılarak takip edilmiştir. Ürünün vasıfları, istatistiksel metotlar kullanılarak tüketici deneyleri vasıtasıyla öğrenilmeye-çalışılmıştır. Fiyatlandırmada, paketlemede ve pazarlamada, en iyi kararları verilebilmek için istatistiksel metotlardan yararlanılmıştır. Hatta süpermarket rafında diş macununun sergilendiği yer bile istatistiksel öğrenme-çalışmalarının bir sonucudur. Reklam kampanyaları istatistiksel metotlar kullanılarak yönetilmiştir. Senin satınalma işlemin bir veri olarak istatistiksel analize tabii olmuştur. Satınalmada kullandığın kredi kartının, güvenliğini sağlayan prosedürler bir istatistiksel modele dayalı olarak çalışmıştır. Yani, istatistik bilimi bütünsel süreci içerisinde yalnızca diş macunu için değil, tükettiğimiz her ürün, aldığımız her hizmet ve katıldığımız her aktivite için önemlidir. Her ne kadar farkında olmasan bile, istatistik bilimi yaşamında yer alan her sürece dahildir. İstatistik bilimi baktığın her yerdedir.*"

Sall J. (2013), "Why Statistics is Essential?" JMP Blog, SAS Blogs Home, link: http://blogs.sas.com/content/jmp/2013/01/14/why-statistics-is-essential/ .

Mini Proje 01

Belirlediğin bir popülasyondan rastgele seçtiğin bireyler üzerinde farklı tiplerde 10 ayrı değişkene ait veri topla. Şunu unutma ki istatistiksel manada "birey" illaki canlı olmak zorunda değil.

- **a.** Hedef popülasyonu belirt.
- **b.** Örnekleme çerçevesini belirt.
- **c.** Rastgele seçim prosedürünü belirt.
- **d.** Örnekleme ünitesini belirt.
- **e.** Gözlem ünitesini belirt.
- **f.** Örneği belirt.
- **g.** Değişkenlerin adlarını, tiplerini, ölçüm birimlerini, ölçüm düzeylerini ve veri tiplerini belirt.
- **h.** Bir veri tablosu hazırla.

a-g şıklarını bir MS Word belgesine ve **h** şıkkını bir MS Excel çalışma-sayfasına kaydet.

BÖLÜM 02 | GÖRSEL BETİMLEME

Giriş

Görsel betimleme ile kastettiğimiz şey, verinin taşıdığı enformasyonu tablo ve grafikler yardımıyla görülebilir hale getirmektir. Bu manada, veriyi organize etmenin ve düzenlemenin çeşitli yolları mevcuttur. Tabloları kullanmak suretiyle oldukça fazla enformasyon göz önüne serilebilir fakat çıplak rakamlardan müteşekkil böylesi düzenlemeler çoğu zaman ruhsuz ve bazen de ürkütücü gelebilmektedir insanlara. Grafiksel gösterimler, hangi formu kullanılırsa kullanılsın, elbette verinin biçimini göstermenin ve veri hakkındaki her şeyin anlık bir algısını yaratmanın en iyi yoludur.

Verinin görsel betimlemesinde en yaygın kullanılan teknikler çubuk grafikler, pasta grafikler, nokta grafikler, dal-yaprak grafikler ve histogramlardır. Bu tekniklerin seçimi verinin tipi ve örnek hacmi ile alakalıdır. Şunu aklımızda her zaman bulundurmalıyız ki tüm grafiksel gösterimler açıkça etiketlenmelidir ki resmin ne anlattığına dair bir şüpheye mahal bırakılmış olmasın.

Hangi görsel betimleme tekniğini kullanırsak kullanalım, yapacağımız iş veriyi bir frekans dağılımına dönüştürmek suretiye, taşıdığı enformasyonu açığa çıkarmak olacaktır. **Frekans dağılımı** gözlem değerlerini ve her bir değerin hangi frekansla ortaya çıktığını kaydeder. Belirli bir gözlemin **frekansı** o gözlem değerinin veri kümesinde tekrarlanma sayısıdır. *Bir değişkenin dağılımı* dediğimiz zaman anlayacağımız şey o değişkene ait gözlemlerin frekanslarının dağılım biçimidir. Yani bir değişkeni bu değişkene ait bir veri kümesi ile öğrenmeye-çalışırken, önce onun dağılımını göstereceğiz.

Frekans Dağılımı Tablosu

Bir değişkene ait gözlem değerlerini ve bu değerlerin frekanslarını bir tabloda gösterebiliriz. Bir nitel değişken doğası gereği sayısal olarak ölçülemeyeceğinden, kategorik değerler alacaktır. Mesela göz rengi nitel bir değişkendir ve kahverengi, siyah, ela, mavi, yeşil ve benzeri olanaklı kategorik değerler bir veri kümesinde belirli frekanskarla tekrarlanacaktır. *Bir nitel değişkenin dağılımı*, gözlenen kategorileri ve bu kategorilere düşen bireylerin sayılarını (yani, *frekanslarını*) veya yüzdelerini (yani, *nispetlerini*) listeler.

Misal 1 Bir sınıfta 35 öğrenci mevcut. Bu öğrencilere en çok ilgi duydukları spor dalı soruluyor. Alınan yanıtlar şunlar: {futbol, tenis, boks, futbol, boks, futbol, güreş, basketbol, basketbol, güreş, güreş, basketbol, boks, güreş, güreş, futbol, güreş, futbol, tenis, futbol, futbol, güreş, basketbol, basketbol, boks, boks, tenis, basketbol, tenis, tenis, atletizm, futbol, futbol, futbol, atletizm}. Bu veri bir frekans dağılımı tablosunda şöyle düzenlenebilir:

Kategori	Frekans	%
Futbol	10	28.6
Güreş	7	20.0
Basketbol	6	17.1
Boks	5	14.3
Tenis	5	14.3
Atletizm	2	5.7
Toplam	**35**	**100**

Bu frekans dağılımı tablosunu incelediğimizde, öğrencilerin 10 unun futbol, 7 sinin güreş, 6 sının basketbol, 5 inin boks, 5 inin tenis ve 2 sinin atletizm yanıtını vermiş olduklarını görüyoruz. Bu rakamlar – 10, 7, 6, 5, 5 ve 2 – altı kategorinin frekanslarıdır ve ham verinin organize edilmiş bir halidir. Frekans dağılımı tablosu ayrıca her altı kategori için yüzdeleri de gösteriyor. Bu yüzdeler, bu öğrenme-çalışmasında yer alan 35 bireyin en çok ilgi duydukları spor dallarının nispetleridir. Öğrencilerin sadece %5.7 sinin atletizme ilgi duyması şaşıtıcı mı?

Bir nicel değişkenin gözlem değerleri az sayıdaki tam sayının tekrarı olarak tezahür ediyorsa, frekans

dağılımı tablosunda biribirinden farklı her gözlem değerini bir **sınıf** olarak telakki ederek, her bir sınıfın karşısına frekansını yazmak yeterlidir. Fakat, eğer veri çok geniş gidimli tam sayılardan veya kesirli sayılardan müteşekkil ise, o zaman eşit aralıklı ardışık sınıflar oluşturup, bu sınıfların karşılarına frekanslarını yazmak daha yararlıdır. **Sınıf aralıklarını** tespit ederken göz önüne alınacak hususlardan birincisi aralıkların çakışmaması ve ikincisi de olanaklı tüm gözlem değerlerinin bu aralıklardan herhangi birine düşmesi gerektiğidir. Nicel değişkenler için frekans dağılımı tabloları yapılırken, sınıf sayısı genellikle 5 ve 15 arasında seçilir, öyle ki değişkenliğin biçimi açıkça ortaya koyulabilsin. *Bir nicel değişkenin dağılımı*, gözlenen sınıfları ve bu sınıflara düşen bireylerin sayılarını (yani, *frekanslarını*) veya orantılarını (yani, *nispi frekanslarını*) listeler.

Misal 2 On beş ev kadınının haftalık TV izleme saatleri şöyle gözlenmiştir: {54, 59, 35, 41, 46, 25, 47, 60, 54, 46, 49, 46, 41, 34, 22}. Şimdi, bu veriye bakarak, ev kadınlarının TV seyretme saatleri hakkında enformasyon elde etmeye çalışalım. Daha sonra, aynı enformasyon arayışına veriyi bir frekans dağılımı tablosuna dönüştürerek devam edelim. Bu değişkene ait frekans tablosunun her bir sınıfı veri kümesini oluşturan tam sayılardan birine tekabül edecektir.

Sınıf	Frekans
22	1
25	1
34	1
35	1
41	2
46	3
47	1
49	1
54	2
59	1
60	1

Frekans dağılımı tablosu, ham verinin sahip olduğu fakat açıkça gösteremediği enformasyonu bize daha açıkça gösterebiliyor. Değişkenin frekans dağılımına bakarak müşterek veya tipik değeri ve gözlem değerlerinin hangi aralıkta değiştiğini şimdi daha kolayca görebiliyoruz. Fakat sınıf sayısını manipüle ederek, değişkenliğin biçimini çok daha açıkça göz önüne serebiliriz. Şöyle ki, bu dafa sınıfları 10 ar birimlik aralıklar ile yeniden dizayn edeceğiz.

Sınıf	Frekans
20 – 29	2
30 – 39	2
40 – 49	7
50 – 59	3
60 – 69	1

Yeni dizaynıyla, frekans dağılımı tablomuz daha fazla enforme edici görünüyor mu? Değişkenin tipik değerini ve değişkenliğini daha açıkça görebiliyor musun?

Misal 3 Bir mahalle bakkalına belirli bir günde gelen 50 müşterinin harcamaları (TL) şöyle kaydedilmiştir: {2.32, 6.61, 6.90, 8.04, 0.45, 10.26, 11.34, 11.63, 12.66, 12.95, 13.67, 13.72, 14.35, 14.52, 14.55, 15.01, 15.33, 16.55, 17.15, 18.22, 18.30, 18.71, 19.54, 19.55, 20.58, 20.89, 20.91, 21.13, 23.85, 26.04, 27.07, 28.76, 29.15, 30.54, 31.99, 32.82, 33.26, 33.80, 34.76, 36.22, 37.52, 39.28, 40.80, 43.97, 45.58, 47.36, 52.57, 53.85, 54.30, 69.49}. Bu değişkenin frekans dağılımı tablosunu yaparken, sınıfları ardışık eşit aralıklar olarak teşkil edeceğiz ve her bir aralığa düşen gözlem değerlerini sayacağız.

Sınıf	Frekans
0.01 – 10.00	5
10.01 – 20.00	19

20.01 – 30.00	9
30.01 – 40.00	9
40.01 – 50.00	4
50.01 – 60.00	3
60.01 – 70.00	1

Değişkenin frekans dağılımı, şimdi bize aradığımız enformasyona daha kolayca ulaşabilme olanağı veriyor. Değişkenin tipik değerini ve verinin nasıl değiştiğini açıkça görebiliyoruz.

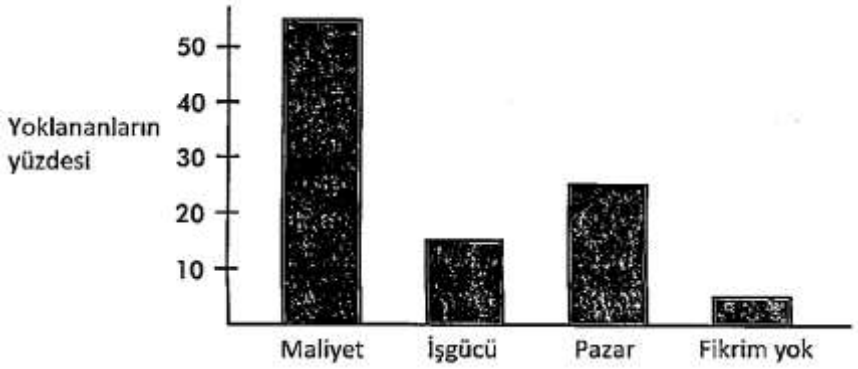
Her ne kadar, bir frekans dağılımı tablosu bize verinin taşıdığı enformasyonu daha net olarak gösterse dahi, mütenasip grafikleri kullanarak verideki değişkenliğin biçimini çok daha hızlıca görebiliriz ve elde ettiğimiz enformasyonu başkalarıyla daha kolayca paylaşabiliriz. Birinci bölümde, veri tipinin kullanacağımız istatistiksel analiz metodunu belirleyeceğini söylemiştim. Bu bölümden itibaren bu seçimi sürekli yapmak durumunda olacağız. Kategorik verinin görsel betimlemesinde çubuk grafikleri ve pasta grafikleri tercih ederken, sayısal verinin görsel betimlemesinde nokta grafikleri, dal-yaprak grafikleri ve histogramları tercih edeceğiz.

Veri ile yapılacak ilk iş daima bir resmini çizmek olmalıdır.

Çubuk Grafik

Nitel bir değişkeni, ürettiği kategorik veri ile öğrenmeye-çalışırken çubuk grafik ve pasta grafik bilhassa yararlı tekniklerdir. Kategorilere (veya gruplara) düşen gözlem sayılarını çabucak karşılaştırmak için yaygınca kullanılan bir görsel betimleme tekniği *çubuk grafiktir*. Çubukların uzunlukları frekanslara veya yüzdelere göre ölçeklendirilebilir. Bildiğin üzere, bu tip değişkenler için her bireyin ait olduğu kategoriyi kaydediyorduk. (Bunun aksine, nicel değişkenler sayısal değerler alıyordu.)

Misal 4 Ülke genelinde tarımsal işletmelerle yapılan bir yoklamada tarım işletmelerinin en önemli sorunun hangisi olduğu soruldu. Yoklamaya katılanların %55 i maliyet, %15 i işgücü, %25 i pazar ve ve %5 i fikrim yok yanıtlarını verdi. Bu veri aşağıdaki çubuk grafikle gösterilebilir:



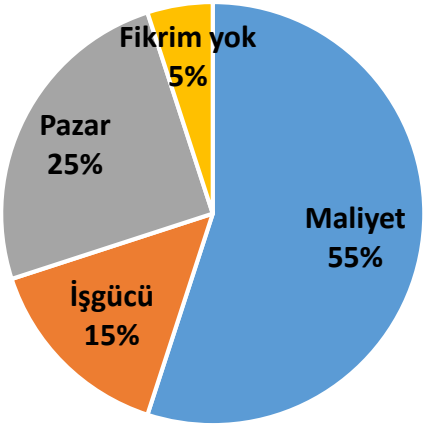
Farklı kategorilerin yüzdeleri, bu kategorileri temsil eden mütenasip uzunluktaki çubuklarla belirtilmiştir.

Bir değişkene ait gözlemlerin sayısal değerler ile etiketleniyor olması, onun bir nicel değişken olduğu manasına gelmez. Mesela futbolcuların forma numaraları gibi.

Pasta Grafik

Çubuk grafiğin kullanıldığı durumlarda alternatif olarak kullanılabilen bir grafik türü de pasta grafiktir. Bir pasta grafik her bir kategorinin bütünün ne kadarını teşkil ettiğini görmemize yardımcı olur. Pasta grafik bir dairenin sahip olduğu 360 derecelik yayın olanaklı kategorilerin frekanslarına nispeten parçalanmasıyla oluşturulur. Pasta grafiğin görünürde bir ölçeği olmadığı için, çoğu zaman kategorilerin frekansları veya yüzdeleri etiketlerinin yanına yazılır.

Misal 5 Misal 4 teki kategorik verinin dağılımını bu defa bir pasta grafik ile gösterelim. Microsoft Excel yazılımı ile çizilmiş pasta grafiği aşağıda görebiliyorsun. Farklı kategorilerin yüzdeleri, bu kategorileri temsil eden mütenasip alanlara sahip dilimlerle belirtilmiştir.



Nokta Grafik

Bilhassa tam sayı değerler alan nicel değişkenlerin frekans dağılımını göstermek için nokta grafik kullanılabilir. Örnek hacmi geniş olduğu zaman pek kullanışlı değildir fakat dar hacimli örnekler için hızlı bir grafiksel araçtır. Bir nokta grafikte, değişkenin her çıkımının frekansı o çıkımın sayısına eşit sayıdaki noktalarla belirtilir. Yatay veya dikey bir eksen üzerinde ölçeklendirilen değerlerin karşısına noktaların yerleştirilmesi ile çizilir. (Aslında, frekans dağılımının hemen hemen her türde grafiksel gösteriminde, resmin hangi şekilde daha iyi okunabileceğine veya sayfaya nasıl daha iyi uyacağına göre dikey veya yatay seçimi yapılabilir.)

Misal 6 Bir bölgede rastgele seçilen 30 ailenin çoçuk sayıları şöyledir: {7, 7, 6, 9, 4, 0, 4, 5, 4, 7, 8, 6, 6, 8, 5, 11, 8, 7, 4, 6, 6, 7, 3, 2, 3, 4, 10, 8, 6, 7}. Bu veriyi bir nokta grafik ile betimleyelim.

Basit bir gösterim olan bu nokta grafiğe bakarak, dağılıma ait temel olguları görebiliriz. İki uçtaki noktalar bize minimum (veya en az) ve maksimum (veya en çok) çocuk sayılarını veriyor. Yine, tipik çocuk sayısının ne olabileceği ve verinin değişkenliği hakkında fikir edinebiliyoruz.

Dal-yaprak Grafik

Sayısal değer alan nicel değişkenlerin dağılımı, eğer veri kümesi yine çok geniş hacimli değilse, bir dal-yaprak grafik ile gösterilebilir. Dal-yaprak grafik veri kümesinde yer alan gözlem değerlerini yeniden organize ederek, verinin taşıdığı enformasyonu açığa vurmanın zekice bir yoludur.

Dal-yaprak grafik, en çok, hepsi de sıfırdan büyük, az sayıda gözlem için uygundur. Gözlem sayısı fazla olduğunda, sayısal değerleri fiilen grafikte barındırması sebebiyle özetleyici avantajını kaybeder. Dal-yaprak grafik yaparken şu adımlardan geçeriz:

1. Her bir gözlemi en sağdaki son basamak hariç bir *dala* ve son basamaktan ibaret bir *yaprağa* ayırırız. Dal gerektiği kadar basamaktan oluşabilir fakat yaprak sadece tek basamak olmak zorundadır.
2. Dalları yukarıdan aşağıya doğru artan sırada yazarız ve sağ tarafına dikey bir çizgi çekeriz. Her

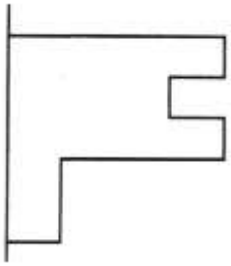
bir yaprağı ait olduğu dalın sağına artan sırada yazarız.

Misal 7 Bir gitar kursuna katılan kursiyerlerin yaşlarına ait gözlem sonuçları şöyledir: $\{17, 17, 18, 13, 28, 38, 31, 27, 35, 50, 43, 37, 24\}$. Bu veriyi bir dal-yaprak grafik ile betimleyelim. Önce, 10 lu, 20 li, 30 lu, 40 lı ve 50 li değerler için 1, 2, 3, 4 ve 5 rakamlarını birer *dal* olarak yukarıdan aşağıya doğru sıralıyoruz. Sonra, orijinal kümedeki her bir değerin son basamağındaki rakamı mütenasip *dalın* sağına bir yaprak olarak yazıyoruz.

Elimizdeki verinin dal-yaprak grafiği, nihayeti itibariyla şöyledir:

Dallar	Yapraklar
1	7 7 8 3
2	8 7 4
3	8 1 5 7
4	3
5	0

Yaprakların etrafına sürekli bir çizgi çekmek suretiyle, dağılımın biçimini daha net görebiliriz.



Misal 8 İngiliz bilimci Henry Cavendish, 1798 yılında, bir burkulma terazisi kullanarak yer kürenin yoğunluğunu ölçmeye girişti. Yaptığı 29 ölçümün sonucunu şöyle kaydetti: 5.50, 5.57, 5.42, 5.61, 5.53, 5.47, 4.88, 5.62, 5.63, 4.07, 5.29, 5.34, 5.26, 5.44, 5.46, 5.55, 5.34, 5.30, 5.36, 5.79, 5.75, 5.29, 5.10, 5.86, 5.58, 5.27, 5.85, 5.65 ve 5.39 gr/cm^3.

Bu veriye ait dal-yaprak grafiği aşağıda. Yana yatırılmış halde. Dikkat edersen, grafikte veri öyle bir ölçeklendirilmiş ki her bir değeri 0.01 ile çarpmak suretiyle orjinal değerine ulaşılabilir. Mesela, $407 \times 0.01 = 4.07$. (Yani, grafikteki 40 | 7 nin manası 4.07 gr/cm^3 tür.)

40	7
41	
42	
43	
44	
45	
46	
47	
48	8
49	
50	
51	0
52	6 7 9
53	0 4 4 9
54	2 4 6 9
55	0 3 5 7
56	1 2 3 7
57	5 9 5
58	5 6

Bu dal-yaprak grafiğin yapısı yine dikkat ettiysen eğer, Misal 7 dekinden daha farklı. Her satırdaki rakamlar küçükten büyüğe doğru yeniden düzenlenmiştir. Bu şekilde, verinin taşıdığı enformasyonun grafiğe yansıtılması daha üst bir seviyede gerçekleşmiş olur.

Dalları bölmek

Bazı veriyi dal-yaprak grafiğine dökerken frekans dağılımının biçimini daha iyi gösterebilmek için dalları bölerek grafiği genişletmek grekebilir. Mesela, Erzurum'da 15 kişiye günde kaç bardak çay içtikleri sorulmuş ve şu yanıtlar alınmış: 5, 7, 4, 9, 6, 3, 5, 0, 4, 7, 2, 8, 4, 6, 4. Bu verinin dal yaprak-grafiği ilk elde şöyledir:

0 | 0 2 3 4 4 4 4 5 5 6 6 7 7 8 9

Bu haliyle dal-yaprak grafiği veri hakkında fazla bir enformasyon sağlamıyor. Yalnızca bir dal üzerinde yapraklar aşırı yığılmış. Eğer burada olduğu gibi yapraklar fazla kalabalıklaşırsa, o halde her bir dalı iki veya daha çok dallara bölmek yararlı olabilir. Buna göre, 0 – 9 aralığını 0 – 4 ve 5 – 9 aralıklarına bölebiliriz. Bu halde dal-yaprak grafiğimiz şöyle görünecektir.

$0^{(0)}$	0 2 3 4 4 4 4
$0^{(5)}$	5 5 6 6 7 7 8 9

Dikkatini çekmek istiyorum ki $0^{(0)}$ dalı 0 – 4 aralığındaki tüm veriyi ve $0^{(5)}$ dalı 5 – 9 aralığındaki tüm veriyi ihtiva ediyor. Ya da, benzerce, 0 – 9 aralığını beş aralığa bölebiliriz: 0 – 1, 2 – 3, 4 – 5, 6 – 7, 8 – 9. Bu halde dal-yaprak grafiğimiz şöyle görünecektir.

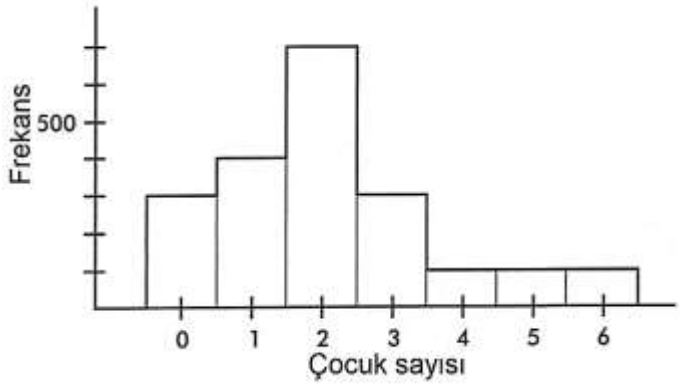
$0^{(0)}$	0
$0^{(2)}$	2 3
$0^{(4)}$	4 4 4 4 5 5
$0^{(6)}$	6 6 7 7
$0^{(8)}$	8 9

Histogram

Nicel değişkenleri havi geniş veri kümeleri söz konusu olduğu zaman, histogramları kullanmak daha yararlıdır. Bir histogram ya belli değerlerin ya da belli değerler arasına (yani, *sınıflara*) düşen değerlerin sayılarını (yani, *frekanslarını*) gösterir.

Bir histogramda, yatay eksen üzerinde işaretlenen eşit genişliklerden her biri bir sınıf aralığıdır. Sınıf aralıkları biribirlerine bitişik olmalı ve biribirleriyle çakışmamalıdır. Öyle ki nicel değişkenin olanaklı tüm gözlem değerleri bu sınıf aralıklarınnın ya birine ya diğerine düşmelidir. Her bir sınıf aralığına ait frekans dikey eksene göre ölçeklendirilerek bir kutu ile temsil edilir. Böylece bir histogram yanyana bitişik kutulardan müteşekkil bir biçim arzeder. Sınıf aralıklarının yatay eksende ve frekansların dikey eksende işaretlenmesi adet olmakla beraber, tersi de yapılabilir.

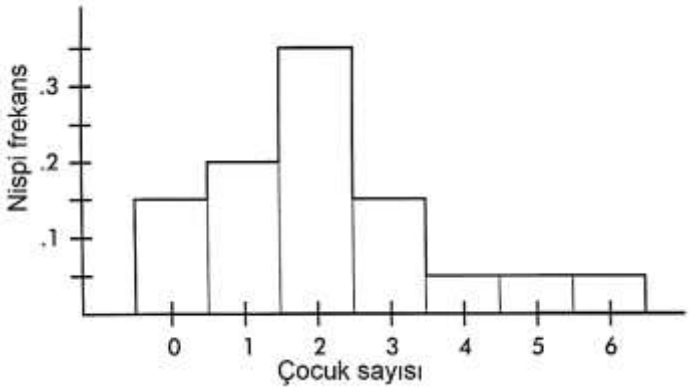
Misal 9 Bir ilçede 2000 aile yerleşik. Bu ailelerin sahip oldukları çocuk sayılarının dağılımı şöyledir: 300 aile çocuksuz, 400 aile bir çocuklu, 700 aile 2 çocuklu, 300 aile üç çocuklu, 100 aile dört çocuklu, 100 aile beş çocuklu ve 100 aile altı çocuklu. Bu veri müteakip histogramla gösterilebilir:



Bazen dikey ekseni, frekanslarla etiketlemek yerine *nispi frekansları* kullanmak daha elverişli veya daha manalı olabilir. Nispi frekansları hesaplamak için frekansları veri kümesinin hacmine, yani tüm sınıfların frekanslarının toplamına böleriz.

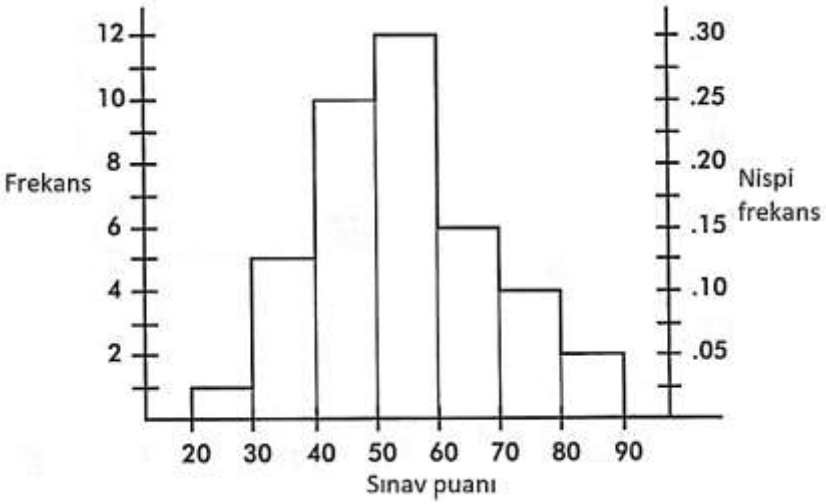
Çocuk sayısı	Frekans	Nispi frekans
0	300	$300/2000 = 0.15$
1	400	$400/2000 = 0.20$
2	700	$700/2000 = 0.35$
3	300	$300/2000 = 0.15$
4	100	$100/2000 = 0.05$
5	100	$100/2000 = 0.05$

6	100	$100/2000 = 0.05$
Toplam:	2000	1



Dikkatini çekmek isterim ki dikey eksen ister frekanslarla ister nispi frekanslarla etiketlensin, histogramın şekli aynıdır. Bazen aynı grafikte hem frekanslar hem de nispi frekanslar gösterilebilir.

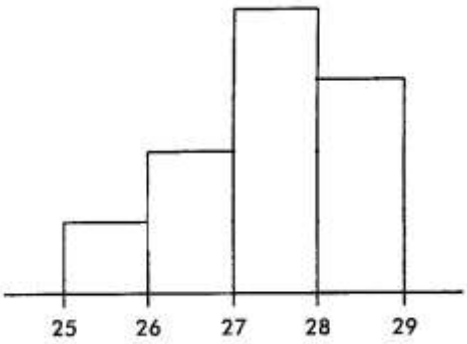
Misal 10 Aşağıdaki histogram 40 öğrencinin istatistik dersinin final sınavından aldıkları puanları göstermektedir.



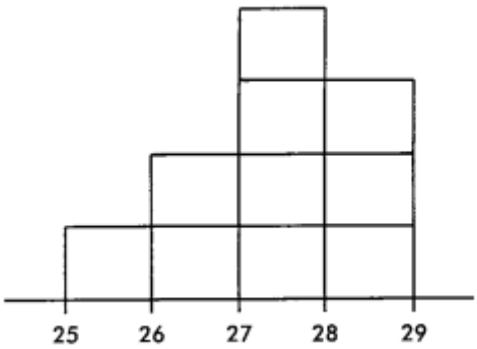
Bu histogramdan ne öğrenebiliriz? Mesela, öğrencilerin hiç biri 90 dan yüksek ve 20 den düşük puan almamıştır. On iki kişi 50 ve 60 arasında bir puan almıştır. Öğrencilerin %25 inin puanı 40 ve 50 arasındadır. Gördüğün gibi bu histogram belli değerler arasına düşen öğrencilerin (yani puanların) sayısını gösteriyor. Oysa, bir önceki histogram her bir çocuk sayısına sahip ailelerin (yani çocuk sayılarının) sayısını gösteriyordu. Mesela, Misal 10 da görüyoruz ki puanların on tanesi 40 ve 50 arasına düşmekteyken, Misal 9 da görüyoruz ki 700 ailenin tam iki çocuğu var.

Biribirinden farklı çok sayıda değerler alan bir değişkene ait kalabalık bir veri kümesini belirli değer aralıklarının frekansları ile betimlemek iyi bir fikir elbette fakat diğer taraftan bireysel puanların tam değerlerini de kaybetmiş oluyoruz. Yani, bir dal-yaprak grafiğin çizildiği veri kümesini, bu grafikten tekrar tam olarak elde edebiliriz. Ancak bir histogram için bu her zaman olanaklı değildir.

Misal 11 Aşağıdaki histogramda dikey eksen etiketlenmemiştir. Bu histogramdan ne öğrenebiliriz?



Cevap: Gerçek frekansları belirlemek olanaksız. Mamafih, her bir aralık üzerindeki alanı tüm aralıkların üzerindeki toplam alana orantılayarak nispi frekansları belirleyebiliriz:



Toplam alanı on eşit parçaya böldükten sonra, görüyoruz ki alanın $\frac{1}{10}$ u veya %10 u 25-26 aralığının üzerindedir, %20 si 26-27 aralığının üzerinde, %40 ı 27-28 aralığının üzerinde ve %30 u 28-29 aralığının üzerindedir.

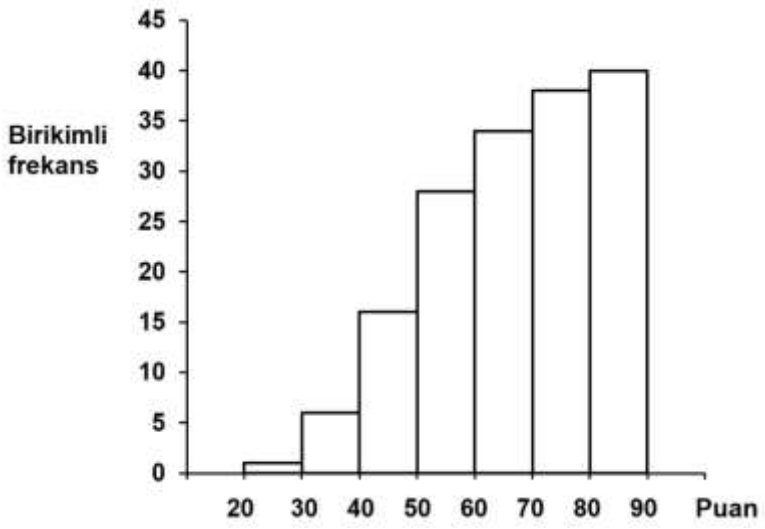
Elbette her histogramı buradaki gibi kolayca on eşit parçaya bölmek mümkün olmayabilir fakat nispi frekansların nispi alanlara tekabül etmesi fikri daima geçerlidir.

Eğer farklı hacimdeki veri kümelerinin dağılımını karşılaştırmak durumu hasıl olmuş ise, nispi frekansları kullanmak daha yararlıdır.

İstatistiksel manada, herhangi bir popülasyon ve herhangi bir örnek esasen birer frekans dağılımıdır.

Birikimli Frekans Grafikleri

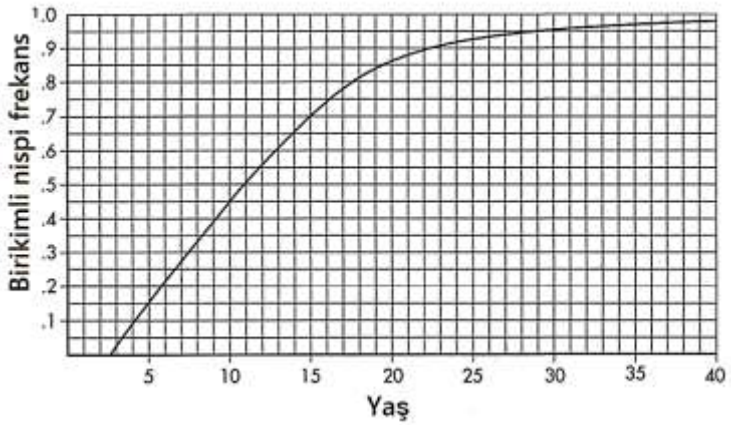
Yatay eksen üzerinde biribirini takip eden aralıkların (yani, sınıfların) frekanslarını üst üste ekleyerek bir grafik çizilebilir. Buna birikimli frekans grafiği denir. Misal 10 daki 40 öğrencinin istatistik dersinin final sınavından aldıkları puanların dağılımını gösteren histogramın sınıflarında yer alan frekansları üst üste topladığımız zaman, şu birikimli frekans grafiğini elde ederiz.



Dikkatini çekmek isterim ki birikimli frekans grafiğinde son sınıfın frekansı her zaman toplam frekansa eşittir çünkü kendisinden önceki tüm sınıfların frekanslarının ve kendi frekansının toplamını havidir.

Birikimli frekans grafikleri çoğu zaman birikimli nispi frekanslar ile, sınıf aralıklarının orta noktaları üzerinden bir poligon formunda çizilir ve *ogiv* adı da verilir.

Misal 12 Aşağıdaki grafik ABD'de 1996 yılında okula kayıt yaşlarının birikimli frekans poligonunu (ogiv) gösteriyor.



Bu birikimli frekans grafiğinden ne öğrenebiliriz? Mesela, grafikte 5 yaş hizasında yukarıya gidersek, görüyoruz ki okula kayıtların 0.15 i veya %15 i 5 yaşın altındadır. Dikey eksende 0.5 in hizasında sola gidersek, görüyoruz ki okula kayıtların %50 si 11 yaşın altında ve %50 si 11 yaşın üstündedir, yani 11 yaş ortadaki değerdir. Otuz yaş hizasında yukarıya gidersek, görüyoruz ki okula kayıtların 0.95 i veya %95 i 30 yaşın altında ve böylece %5 i 30 yaşın üstündedir. Dikey eksende 0.25 in ve 0.75 in hizalarında sola gidersek, görüyoruz ki okula kayıtların ortadaki %50 si 6 küsür yaş ve 16 yaş arasındadır.

Merkez ve Yayılım

Bir grafiksel gösterime baktığımız zaman, frekans dağılımının *genel biçimine* dair, şeklinin yanısıra, iki önemli hususiyet karşımıza çıkar:

1. *Merkez*, ki gözlem değerlerini (veya eğer bir histograma bakıyorsak eğrinin altında kalan alanı) kabaca ortasından ikiye ayırır.
2. *Yayılım*, ki en küçükten en büyüğe kadar değerlerin kapsamıdır.

Misal 9 daki histogramın merkezi 2 çocuk iken, yayılımı 0 dan 6 çocuğadır. Misal 10 daki histogramın merkezi 50 ve 60 puan arasında ve yayılımı 20 den 90 puana kadardır. Misal 12 deki birikimli frekans

grafiğinde dağılımın merkezi 10 ve 11 yaşları arasında ve yayılımı ise 3 yaştan 40 yaşa kadardır. Misal 7 deki dal-yaprak grafiğinin merkezi 28 yaş iken, yayılımı 13 ten 50 yaşa kadardır. Misal 8 deki dal-yaprak grafikte ise dağılımın merkezi 5.46 gr/cm^3 ve yayılımı 4.07 den 5.86 gr/cm^3 e kadardır.

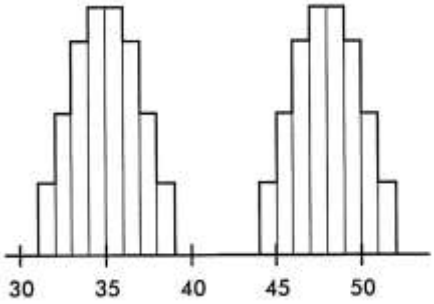
Merkez ve yayılım daima beraber betimlenmelidir.

Öbekler ve Boşluklar

Dağılımın genel biçimine dair diğer önemli hususiyetler şunlardır:

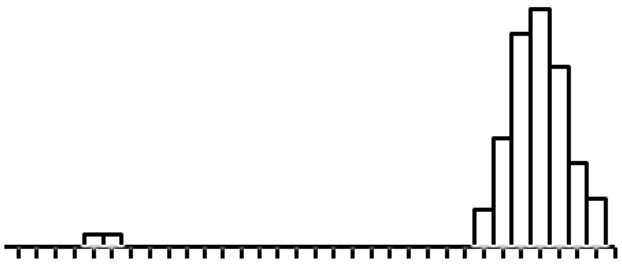
1. *Öbekler*, ki gözlem değerlerince oluşturulan doğal alt-gruplaşmalardır.
2. *Boşluklar*, ki hiç bir değerin düşmediği bölgelerdir.

Misal 13 Aşağıdaki histograma bakalım.



Basitçe, dağılımın merkezinin 42 civarında ve yayılımının 31 den 52 ye kadar olduğunu söylersek, açıkça bir şeyi ıskalamış olmaz mıyız? Gözlem değerleri, aralarında bir boşluk barındıran iki ayrı öbek oluşturmaktadır.

Dışlaklar



Dışlaklar (/*outliers*) diye adlandırılan aşırı değerler pek çok dağılımda kendini gösterebilir. Bazen bu dışlaklar, ölçüm hatalarının sonucu ortaya çıkarlar ve incelenmeleri gerekir. Diğer taraftan, dışlaklar doğal şansa bağlı değişimin bir sonucu da olabilirler. Dağılımın her iki yanında veya yalnızca bir yanında görülebilirler. Misal 8 deki 4.07 değeri açıkça bir dışlaktır.

Dışlaklara dikkat et!

Misal 14 Selcan bir yüzücüdür ve Türkiye Şampiyonası için antremanlarını sürdürüyor. Otuz gün boyunca her gün yüzdüğü 50 metrelik tur sayısı aşağıda iki farklı dal-yaprak grafinde betimlenmiştir.

1	0 8 9
2	0 1 2 2 4 4 4 5 5 6 6 6 7 7 7 7 8 8 8 8 8 9 9
3	1 1 2 9

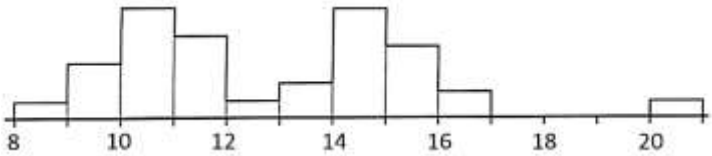
$1^{(0)}$	0
$1^{(5)}$	8 9
$2^{(0)}$	0 1 2 2 4 4 4
$2^{(5)}$	5 5 6 6 6 7 7 7 7 8 8 8 8 8 9 9
$3^{(0)}$	1 1 2
$3^{(5)}$	9

Frekans dağılımının sağdaki genişletilmiş grafiği bize gösteriyor ki Selcan günlük antremanlarında mutadı vechile 25 ve 29 tur arasında yüzmüş. Dağılımın iki ucunda yer alan 10 ve 39 turluk gözlemleri dışlak olarak telakki edebiliriz. Dışlakları görebilmenin frekans dağılımının organizasyonuyla alakalı olduğuna dikkat etmiş olmalısın.

Mod

Bazı dağılımlar bir veya birden çok sayıda başlıca tepeye sahiptir ki bu tepelere *mod* adı verilir. Dağılımın sadece bir tepesi varsa *tek-modlu,* iki tepesi varsa *iki-modlu,* ikiden çok tepesi varsa, *çok-modlu* denir. Fakat verideki her küçük yükselti bir mod değildir. Her zaman büyük resme bakmaya çalışmalı ve hangi fenomenlerin dağılımın biçimini etkilediğine karar vermelisin.

Misal 15 Aşağıdaki histogram bir şirketin elemanlarının internete erişim zamanlarını göstermektedir.



Bunun iki-modlu bir dağılım olduğunu görebiliyoruz. Bu şirketin internete erişimi, görünen o ki kuşlukta (sabah ortası) ve ikindide (öğleden sonra ortası) yoğunlaşmakta ve öğle yemeği saatinde seyrekleşmektedir. Bir akşam saati dışlağı mevcut ki akşam yemeğinden dönen elemanlara (veya bekçilerin veya temizlik işçilerinin bir internet molasına) olanaklıca işaret etmektedir.

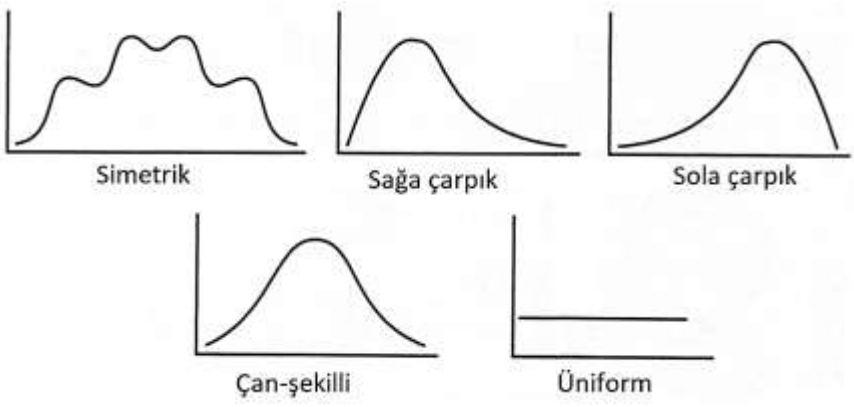
Yukarıda misallendirilerek anlatılmaya çalışıldığı üzere, dışlakların ve modların arkasındaki sebepleri aramak değişken hakkında çoğu zaman daha çok şey öğrenmemizi sağlar.

Tek-modlu bir dağılımın modunu o verinin ait olduğu değişkenin aldığı en tipik değer olarak düşünebiliriz. (Bu fikri, bir sonraki bölümde, dağılımın merkezileşme ölçüleri bahsinde daha geliştireceğiz.) Nitel değişkenlerin dağılımı için mod olanaklı kategoriler arasında frekansı en yüksek olanı belirtir.

Şekil

Dağılımların şekli sonsuz çeşitlilikte olabilir. Mamafih, yaygın rastlanan bazı şekiller özellikle ilgi alanımızdadır:

1. *Simetrik* bir dağılım, biribirinin yansıması olan iki yarıma sahiptir. Mesela, toplumları meydana getiren insanların ağırlıkları, boyları ve IQ puanları gibi değişkenlere ait verinin dağılımı şekil itibariyla simetriktir.
2. Bir dağılım *sağa çarpıktır*, eğer ki daha büyük değerler yönünde zayıflayarak uzağa yayılıyor ise. Mesela, emeklilerin yaşlarının dağılımı öyle bir dağılımdır ki büyük yaşlara doğru gittikçe dramatik olarak azalan sayılara sahiptir.
3. Bir dağılım *sola çarpıktır*, eğer ki daha küçük değerler yönünde zayıflayarak uzağa yayılıyor ise. Mesela, kolay bir sınavdan alınan puanlar daha yüksek uçta yığılırken, az sayıda düşük puan görülecektir.
4. Bir merkezi tümseğe ve eğimli iki kuyruğa sahip simetrik bir dağılım *çan-şeklindedir*. Mesela, genel nüfusun IQ puanlarının dağılımı kabaca simetriktir, 100 puan etrafında merkezi bir tümseğe sahiptir ve eğimli iki kuyruğu vardır.
5. Bir dağılım *üniformdur*, eğer ki bu dağılımın histogramı yatay bir çizgi ise. Mesela, hilesiz bir zarı çok çok çok sayıda kere atar ve üste gelen yüzdeki noktaların sayısını kaydedersek, hepsi eşit olabilirlikli olmak üzere 1 den 6 ya kadar sayıların üniform bir dağılımını elde etmeyi bekleriz.



Elimizdeki verinin dağılımının şekli yukarıda özellikle üzerinde durduğumuz şekillerden herhangi biriyle temelde benzerlik gösteriyor olsa bile, eğer verinin bir kısmı bu şekilden sapıyor ise, bunu not etmek gerekir. Mesela, Misal 8 de 4.07 bir dışlak, 4.88 bir aşırı değerdir ve geriye kalan gözlem değerleri kabaca çan-şeklinde bir dağılım göstermektedir.

Mini Proje 02

MP 01 de topladığın verinin mütenasip tekniklerle görsel betimlemesini yap.

BÖLÜM 03 | SAYISAL BETİMLEME

Giriş

Elimizde bir ham veri kümesi mevcut olduğu zaman, verinin taşıdığı enformasyona ulaşabilmek maksadıyla, önce verinin görsel bir betimlemesini yaparak, dağılımın genel bir biçimini ortaya koyarız. Dağılımın grafiği hangi değerlerin hangi frekanslarla ortaya çıktığı, aşırı değerler ve değerlerin gidimi (/*range*) hakkında bizim fikir sahibi olmamızı sağlar. Betimleyici istatistiğin birinci aşaması olan görsel betimlemeden sonra, verinin sayısal olarak betimlenmesi aşamasına geçilir. Sayısal betimleme dağılımın merkezine, yayılımına, çeşitli değerlerin pozisyonlarına ve şekline ait bir takım ölçülerin hesaplanmasını ihtiva eder. Diğer bir deyişle, verinin sayısal betimlemesi dediğimiz zaman kastettiğimiz şey, verinin taşıdığı enformasyonu açığa çıkarmak ve daha kolay iletebilmek maksadıyla frekans dağılımının özetlenerek bir kaç sayısal betimleyiciye indirgenmesidir. Örnek veriden popülasyona genelleştirilecek neticelerin çıkarılması prosesine çıkarımsal istatistik adını verdiğimizi ve sayısal betimleyicilerin bu prosesin girdisi olduklarını daha önce öğrenmiştin.

Bir frekans dağılımının en basit sayısal betimlemesi bir merkez ölçüsünden ve bir yayılım ölçüsünden ibarettir.

Merkezin Ölçülmesi

Bir önceki bölümde bir frekans dağılımında en yüksek frekansa sahip olan gözlem değerini veya sınıf aralığını mod sınıfı olarak tanımlamış ve bunu değişkenin tipik bir değeri olarak addetmiştik. İlgilenilen değişken için tipik bir değer bulma fikrini bu bölümde ölçmenin derecelendirme ve aralık düzeylerinde ele alıp daha kesinlikli *merkezileşme* (veya *merkezi temayül*) ölçüleri üzerinde duracağız. Bir sayısal verinin frekans dağılımının merkezini ölçen ölçülere genel manada *ortalama* adı verilir.

Gündelik konuşmalarımıza ait cümlelerde, *ortalama* kelimesi sıkça geçer. İnsanlar ortalama fiyatlardan, not ortalamalarından, bir pilin veya bir insanın ortalama ömründen bahseder. Esasen, *ortalama* (/*average*) terimi, Fransızca "*averie*" kelimesinden türetilmiştir ki bu kelime deniz tacirlerinin, malını güvenle limana ulaştıramayan meslektaşlarının uğdadıkları zararı telafi etmek için katkıda bulundukları para manasında kullanılırdı (yani, mağdurun zararı, katkıda bulunanların her birince bir ortalama miktar ile paylaşılırdı). İstatistik bilimi jargonunda, *ortalama* terimi ile bir dağılımın temsilci bir değerini veya tipik bir değerini veya merkezini kastederiz. Matematiksel olarak, bir veri kümesinin ortalamasını tanımlamanın çeşitli yolları vardır. Pratikte, ele alınan durumun hususiyetine en uygun metot hangisi ise onu kullanırız.

Bu bölümde bir ortalamayı ifade etmek için başta gelen üç metodu öğreneceksin, ki bunlardan ikisi istatistiksel veri analizinde etkin olarak kullanılmaktadır:

1. *Orta-gidim*, ki veri kümesinin iki aşırı değeri (minimum ve maksimum) arasındaki mesafenin orta noktasıdır. Bu iki değeri toplayıp, ikiye bölerek hesaplarız.
2. *Ortanca*, ki bir veri kümesi küçükten büyüğe doğru tertip edildiği zaman, sıra olarak ortaya düşen değerdir.
3. *Aritmetik ortalama*, ki bir veri kümesindeki tüm değerleri topladıktan sonra, toplamı bu değerlerin sayısına bölmek suretiyle bulunur.

Misal 1 Müteakip veri kümesinde ateşli bir hastalık geçiren bir kişinin 14 günlük ateş ölçümleri (^{o}C) küçükten büyüğe doğru sıralanmıştır:

$$\{38.7, 39.4, 40.0, 40.0, 41.0, 41.0, 41.0, 41.4, 41.5, 42.0, 42.0, 42.1, 45.7, 46.1\}.$$

Ortalama nedir?

Cevap: Orta-gidimi bulmak için, verinin minimumunu (ki 38.7) ve maksimumunu (ki 46.1) toplayıp ikiye böleriz,

$$\frac{38.7 + 46.1}{2} = 42.4\ ℃.$$

Ortancayı bulmak için, küçükten büyüğe doğru sıralanmış 14 değerin ortasına düşen yedinci ve sekizinci sıradaki değerleri toplayıp ikiye böleriz. Çünkü, gözlem değelerinin sayısı çift, bu yüzden, orta sıraya bir değil iki değer düşüyor.

$$\frac{41.0 + 41.4}{2} = 41.2 \text{ derece Celsius.}$$

Aritmetik ortalamayı bulmak için, tüm değerleri toplayıp 14 e böleriz,

$$\frac{38.7 + 40.0 + 40.0 + 41.0 + 41 + 41.0 + 41.4 + 41.5 + 42.0 + 42.0 + 42.1 + 45.7 + 46.1}{13}$$

$$= \frac{542.5}{13} = 41.73 \text{ derece Celsius.}$$

Ortanca

Ortanca (*/median*) terimi Latince "*medius"* kelimesinden türetilmiştir ki "ortadaki" manasına gelmektedir. Ele alınan gözlem değerleri küçükten büyüğe (veya büyükten küçüğe) doğru sıraya dizilir. Sıranın ortasına düşen gözlem değeri ortancadır. Eğer değerlerin sayısı tek ise, ortanca basitçe sıranın tam ortasına düşen değerdir. Eğer değerlerin sayısı çift ise, sıranın tam ortasına düşen tek bir değer yoktur; bu durumda, sıranın ortasına düşen iki değer toplandıktan sonra ikiye bölmek suretiyle bir ortanca bulunur. Böylece, bir veri kümesinin ortancasından daha küçük gözlemlerin sayısı kaç ise, daha büyük gözlemlerin sayısı da aynıdır. Popülasyonun ve örneğin ortancasını simgelemek için sırasıyla *M* ve *m* harfleri genellikle kullanılır.

Büyük değerler istedikleri kadar büyük olsun veya küçük değerler istedikleri kadar küçük olsun, ortanca bunlardan kati surette etkilenmez. Böylece, *dışlaklar* diye adlandırdığımız, aşırı değerler bir manada kuşku verici bulunduğu zaman, veya bunların ortalamaya etkisini azaltmak istediğimiz zaman, ortanca bilhassa yararlı bir ölçüdür. Mesela, eğer on tane fare bir labirenti çözmeye çabalıyor ve dokuz tanesi 15 dakikadan kısa bir zamanda başarırken, bir tanesi 24 saat sonra bile hala çabasını sürdürüyorsa, en temsilci değer ortancadır (aritmetik ortalama değil, çünkü 2 saaten daha büyüktür). Benzerce, eğer dört şirket müdürünün maaşları yıllık 240000TL ve 265000TL arasında ve bir beşincisinin maaşı yıllık 20000TL den daha düşük ise, yine en temsilci değer ortancadır (aritmetik ortalama 200000TL nin altındadır). Ortancanın aşırı değerlere karşı daha dirençli olduğunu söyleriz.

Ortancayı bulurken, veriyi küçükten büyüğe doğru bir sırada tertiplemeyi unutma!

Belli durumlarda ortanca, ortalamayı hesaplamanın en ekonomik ve en çabuk yoludur. Mesela, bir fabrikada belirli bir marka ampullerin 10000 tanesinin aynı anda yakıldığını farzedelim. Ampuller için ortalama bir ömür, ampullerin tam olarak yarısının değiştirilmesi için ne kadar zaman geçtiğini kaydetmek suretiyle kolayca bulunabilir. Ortanca, aynı zamanda, tıbbi bilimsel-araştırmaların belli çeşitleri için de kullanışlıdır. Mesela, muhtelif zehirlerin nispi kuvvetlerini karşılaştırmak için, bir bilimci her zehirin hangi dozunun test hayvanlarının tam olarak yarısını öldürdüğünü kaydeder. Eğer hayvanlardan biri zehirlerden belirli birine karşı münhasıran dayanıksız çıkarsa, ortanca öldürücü doz bundan etkilenmez. Ölçmenin adlandırma düzeyi hariç, diğer tüm düzeylerinde ölçülen veri için ortanca uygun bir ortalama ölçüsüdür.

Aritmetik ortalama

Ortanca, istatistiksel betimlemede çoğu zaman yararlıdır, bununla beraber, *aritmetik ortalama (/mean),* istatistiksel çıkarımda ve analizde çok önemlidir. Ve sokaktaki insan için ortalama genellikle aritmetik ortalama olarak anlaşılır.

Bir *popülasyonun* (yani ilgi duyulan bireylerin tümünün) aritmetik ortalaması çoğu zaman Yunan

harfi μ ("mü" diye okunur) ile gösterilirken, bir örneğin (yani popülasyonun bir kısmının) aritmetik ortalaması çoğu zaman $\bar{y}$ ile gösterilir. Mesela, Türkiye'de yaşayan tüm insanlar kümesinin aritmetik ortalama ağırlığı $\mu = 67.4$ kg olabilirken, rastgele seçilen 100 kişinin ağırlığı $\bar{y} = 61.6$ kg olabilir veya belki $\bar{y} = 78.5$ kg veya hatta $\bar{y} = 86.4$ kg bile olabilir.

İstatistik-bilimi dersinde bir popülasyon aritmetik ortalamasını bir örnek aritmetik ortalaması vasıtasıyla nasıl kestireceğini öğrenirsin. Bu ders boyunca, *örnek* kelimesi ile *basit rastgele örnek* kastedilecektir; yani, bu öyle bir örnek ki aynı hacimdeki başka her olanaklı örnek ile aynı çekilme şansına sahiptir. (Şu da aynı zamanda doğru ki popülasyonun her bir elemanının çekilen örneğe girme şansı eşittir.) Gerçek dünyada, bu rastgele seçim prosesini başarmak çoğu zaman hayli zordur ve bu yüzden, örneğimizin popülasyonun temsilcisi olduğuna dair iyi bir sebebimiz daima olmak zorunda.

Bir örneğin aritmetik ortalamasını matematiksel olarak şöyle yazarız,

$$\bar{y} = \frac{\sum y}{n}$$

ki burada $\sum y$ ile ele alınan örnekteki gözlem değerlerinin toplamı ve n ile örnek hacmi (yani gözlem sayısı) simgelenmektedir. (Σ, Yunan büyük harfi sigmadır ve toplama işlemini temsil eder.) Eğer tüm olanaklı gözlemleri yapabiliyorsak, bir popülasyonun da aritmetik ortalamasını hesaplayabiliriz elbette. Bu halde matematiksel ifademiz,

$$\mu = \frac{\sum y}{N}$$

olarak yazılmalıdır ki burada N popülasyonun hacmini (yani olanaklı tüm gözlemlerin sayısını) simgelemektedir.

Misal 2 Belli bir mahalledeki polis karakolunun nezarethanesine belirli bir günde atılan beş şüphelinin geçmişe ait sabıka kaydı sayıları $\{2, 2, 8, 20, 33\}$ kümesinde verilmiştir. Görüyoruz ki ortanca 8 ve aritmetik ortalama $\frac{2+2+8+20+33}{5} = 13$ tür. Daha sonra, beşinci şüphelinin sabıka kaydının 33 değil de aslında 58 olduğu farkedildiğine göre, ortanca ve aritmetik ortalama bundan nasıl etkilenir?

Cevap: Küme şimdi $\{2, 2, 8, 20, 58\}$ halini aldı. Ortanca hala 8; mamafih, aritmetik ortalama $\frac{2+2+8+20+58}{5} = 18$ e değişir.

Yukarıda misallendirerek göstermeye çalıştığım üzere, ortancanın aksine, veri kümesindeki herhangi bir gözlemin değerinin değiştirilmesine karşı, aritmetik ortalama çok hassastır.

Aritmetik ortalamanın hesaplanacağı veri en az aralık veya oran düzeylerinde ölçülmüş olmalıdır. Fakat uygulamada bu kabülün zorlandığı durumlar da çoktur. Mesela Likert ölçeği, 1930'ların başından beri kullanılan bir tutum ölçeğidir ve yanıtları "*tamamen katılıyorum*", "*katılıyorum*", "*kararsızım*", "*karşıyım*" ve "*tamamen karşıyım*" kategorilerinde kademelendirerek ölçer. Genelde, örnek yoklamalarında, araştırmacılar bu kategorileri 1'den 5'e kadar rakamlarla kodlarlar ve topladıkları verinin aritmetik ortalamasını hesaplarlar.

Ağırlıklandırılmış Aritmetik Ortalama

Bir toplaşık aritmetik ortalama elde etmek üzere, iki veya daha fazla sayıda aritmetik ortalama katıştırıldığı zaman, hesaplandıkları değerlerin sayısına göre, her bir aritmetik ortalamanın tesiri ağırlıklandırılır. Yani, n_1, n_2 ve n_m ile m sayıdaki aritmetik ortalamanın hesaplandıkları değerlerin sayılarını gösterirsek,

$$\bar{y}_{Ağırlıklandırılmış} = \frac{n_1\bar{y}_1 + n_2\bar{y}_2 + \cdots + n_m\bar{y}_m}{n_1 + n_2 + \cdots + n_m}$$

Mesela, bir müzik okulunda piyano öğrenen 10 kişinin yaş aritmetik ortalaması 12, keman öğrenen 15 kişinin yaş aritmetik ortalaması 14 ve gitar öğrenen 40 kişinin yaş aritmetik ortalaması 18 ise bu okulun tüm öğrencilerinin aritmetik ortalama yaşını bulmak için, verilen aritmetik ortalamaları

ağırlıklandırmamız gerekir:

$$\bar{y}_{Ağırlıklandırılmış} = \frac{10(12) + 15(14) + 40(18)}{10 + 15 + 40} = 16.2$$

Yayılımın Ölçülmesi

Bir sayılar kümesini betimlerken, yalnızca bir ortalama değer belirlemek değil fakat aynı zamanda ölçümlerin *değişkenliğini* veya *yayılımını* ifade edebilmek te önemlidir. Bir saatli bomba üreticisi az değişkenliği hedefler – kullandığı 30 dakikalık fünyelerin fiilen patlama öncesi sürelerinin 10 dakikayla 50 dakika arasında bir değişkenliğe sahip olması iyi olmayacaktır. Diğer taraftan, iyi öğrencileri zayıf öğrencilerden ayırmak isteyen bir öğretmen, sonuçlarının değişkenliği fazla olan sınavlar dizayn etmek ister – öğrencilerin hepsinin tam olarak aynı puanı alması iyi olmayacaktır. İki basketbol takımının oyuncularının boy ortalamaları biribirine eşit olabilir, fakat bu gözlem hikayenin tamamını bize anlatmaz. Eğer yayılımları biribirinden hayli farklıysa, bir takımın 2.10 metre boyunda bir oyuncusu varken, diğer takımın 1.90 metreden uzun herhangi bir oyuncusu olmayabilir. İki tatil oteli reklamlarında müşterilerinin yaş ortalamaları için aynı rakamı telaffuz etmiş olabilirler. Mamafih, birisinin müşterileri yalnızca 20 ve 25 yaşları arasındayken, diğerinin müşterileri yalnızca 40 larındaki orta-yaşlı ebeveynler ve onların 10 yaşın altındaki çocukları olabilir.

Değişkenliği veya yayılımı betimlemenin başlıca dört yolu vardır:

1. *Gidim* (/*range*), ki gözlem değerlerinin en büyüğü ve en küçüğü arasındaki mesafedir.
2. *Çeyreklikler*arası *gidim* (/*interquartile range*), *IQR*, ki veri kümesinin alt ve üst çeyrekleri atıldıktan sonra en büyük değer ve en küçük değer arasındaki farktır (yani, IQR ortadaki %50 nin gidimidir); şöyle ki $IQR = Q_3 - Q_1$ veya 75 inci yüzdelik ve 25 inci yüzdelik arasındaki fark.
3. *Varyans*, ki tüm gözlem değerlerinin kendi aritmetik ortalamalarından farklarının karelerinin aritmetik ortalamasıdır.
4. *Standart sapma*, ki varyansın kareköküdür.

Değişimi anlamak istatistik bilimini anlamanın anahtarıdır.

Misal 3 Liberya'nın Monrovia'sında Mayıs ve Ekim ayları arası yağışlı dönem ve Kasım ve Nisan ayları arası kurak dönemdir. Belirli bir yılın aylık yağış miktarları (inç cinsinden) şöyledir:

Ay:	Oc	Şub	Mar	Nis	May	Haz	Tem	Ağ	Ey	Ek	Kas	Ar
Yağış (inç):	1	2	4	6	18	37	31	16	28	24	9	4

Veriyi küçükten büyüğe doğru sıraladıktan sonra, $\{1, 2, 4, 4, 6, \mathbf{9}, \mathbf{16}, 18, 24, 28, 31, 37\}$, ortanca,

$$\frac{9 + 16}{2} = 12.5 \text{ inç}$$

Aritmetik ortalama,

$$\frac{1 + 2 + 4 + 6 + 18 + 37 + 31 + 16 + 28 + 24 + 9 + 4}{12} = 15 \text{ inç.}$$

Değişkenlik ölçüleri nedir?

Cevap:

<u>Gidim:</u> Maksimum 37 inç (Haz) ve minimum 1 inçtir (Oc). Böyle olunca, gidim $37 - 1 = 36$ inçlik yağış miktarıdır.

<u>Çeyreklikler</u>arası gidim: Verinin alt yarımının ortancası,

$$Q_1 = \frac{4+4}{2} = 4 \text{ inç}$$

üst yarımının ortancası,

$$Q_3 = \frac{24+28}{2} = 26 \text{ inç}$$

ve

$$IQR = Q_3 - Q_1 = 22 \text{ inç.}$$

Varyans:

$$\frac{(-14)^2 + (-13)^2 + (-11)^2 + (-9)^2 + 3^2 + 22^2 + 16^2 + 1^2 + 13^2 + 9^2 + 6^2 + 11^2}{12} = 143.7 \text{ inç}^2$$

Standart sapma:

$$\sqrt{143.7} = 11.98 \text{ inç.}$$

Gidim

En basit, en kolay hesaplanan değişkenlik ölçüsü *gidimdir (/range)*. Gözlem değerlerinin *maksimumu* (yani en büyüğü) ve *minimumu* (yani en küçüğü) arasındaki fark çabucak bulunabilir ve verinin gidimine bakılarak yayılımına dair bir fikir edinilebilir. Mamafih, bu yayılım ölçüsü tamamen iki yöndeki aşırı değerlere bağlıdır ve ortadaki değerlere karşı duyarsızdır.

Gidim, çok az sayıda gözlem barındıran örnekleri değerlendirmede yararlıdır. Mesela, bazı kalite kontrol teknikleri periyodik olarak dar örneklerin alınmasını ve bu örneklerin gidimlerine bakarak sonraki aşamalara geçilmesinden ibarettir.

Çeyreklikleraras Gidim

Aşırı değerlerin gidim üzerindeki tesirini elemine etmenin bir metodu verinin *çeyreklikleraras gidimini* (/*interquartile range*) bulmaktır. Çeyreklikleraras gidimi (IQR) hesaplamak için veri küçükten büyüğe doğru sıralanır, üst ve alt çeyreklerde yer alan gözlem değerleri atılır ve geriye kalan değerlerin gidimi bulunur. Yani, gözlem değerlerinin ortadaki %50 sinin gidimi çeyreklikleraras gidimdir.

Çeyreklikler

Ortanca bir veri kümesini ikiye böler; gözlemlerin yarısı ortancanın altında yarısı üstündedir. Ortancayı 50'nci yüzdelik olarak adlandırabiliriz, çünkü tüm değerlerin %50 si ortancadan daha küçüktür. *Alt çeyreklik* verinin alt yarısının ortancasıdır. Yani tüm değerlerin %25 i alt çeyreklikten daha küçüktür. Benzerce, *üst çeyreklik* verinin üst yarısının ortancasıdır. Yani tüm değerlerin %75 i üst çeyreklikten daha küçüktür. Ortancayla beraber çeyreklikler veriyi dört eşit parçaya böler; her bir parça verinin %25'ini barındırır. Alt çeyrekliğe ve üst çeyrekliğe, aynı zamanda, sırasıyla, *birinci çeyreklik* (Q_1) ve *üçüncü çeyreklik* (Q_3) adı verilir.

Misal 4 On beş erkeğin nabız sayıları şöyle ölçülmüştür: {70, 70, 73, 73, 73, 78, 78, 69, 81,71, 72, 76, 77, 77, 74}. Bu verinin birinci ve üçüncü çeyreklikleri nedir?

Veriyi önce küçükten büyüğe doğru giden bir sırayla tertipliyoruz,

69	70	70	**71**	72	73	73	**73**	74	76	77	**77**	78	78	81

Ortanca, listenin merkezinde yer alan ve değeri koyu ve altı çizili 8 inci gözlemdir, $M = 73$. Birinci çeyreklik, ortancanın pozisyonuna göre solda kalan 7 gözlemin ortancasıdır; $Q_1 = 71$. Benzerce, üçüncü çeyreklik, sağda kalan 7 gözlemin ortancasıdır; $Q_3 = 77$. Eğer veri kümesinin hacmi çift olsaydı, o durumda ortanca olarak tek bir gözlem değeri yerine orta sıraya düşen iki gözlem değerinin aritmetik ortalamasını alırdık.

Eğer yazılım ile çalışmıyorsak diğer yüzdelikleri elle daha kabaca hesaplayabiliriz. Verisayfası veya istatistik yazılımları çeyreklikleri biraz farklı değerlerde bize verebilir. Bunun nedeni çeyreklikleri hesaplamanın kesin kuralları olmamasıdır. Ama farklar her zaman çok küçüktür.

Çeyreklikleararası gidim, bir ampirik kural olarak, dışlak zanlılarının teşhisinde bir kıstas olarak kullanılır. Şöyle ki eğer bir gözlemin değeri en az $1.5\,IQR$ kadar birinci çeyrekliğin altında veya üçüncü çeyrekliğin üstünde ise bir *dışlak zanlısı* olarak kabul edilir. (Bu mesafe $3\,IQR$ veya daha fazla olduğunda *uzak dışlak zanlısı* alarmı verilir.)

Misal 5 Bir köydeki işletmelerin arazi genişlikleri şu karakteristikleri haizdir: en küçük gözlem değeri 16.6 dönüm, değerlerin %10 u 23.5 dönümden daha küçük, değerlerin %25 i 41.1 dönümden daha küçük, ortanca 57.6 dönüm, %60 ı 87.2 den daha küçük, %75 i 101.9 dönümden daha küçük, %90 ı 124.0 dönümden daha küçük ve en büyük değer 201.7 dönüm.

a. Gidim nedir?
Cevap: Gidim $201.7 - 16.6 = 185.1$ dönümdür.

b. Çeyreklikleararası gidim nedir?
Cevap: Çeyreklikleararası gidim, değerlerin en alttaki ve en üstteki çeyrekleri atıldıktan sonra, $101.9 - 41.1 = 60.8$ dönümdür. Böylece, en geniş arazi en dar araziden 185 dönüm daha genişken, arazi genişliklerinin ortaya düşen %50 si 61 dönümlük bir aralıkta gitmektedir.

c. Dışlaklar için ampirik kuralı uygularsak, en küçük değer veya en büyük değer dışlak zanlısı olarak teşhis edilebilir mi?
Cevap:

$$1.5IQR = 1.5(60.8) = 91.2.$$

Eğer bir arazi genişliği birinci çeyreklikten 91.2 dönüm daha aşağıda veya üçüncü çeyreklikten 91.2 dönüm daha yukarıda ise, bir dışlak zanlısı olarak teşhis edilebilir. En büyük değer olan 201.7, $101.9 + 91.2 = 193.1$ den daha büyük olduğu için, ampirik kurala göre, bir dışlak zanlısıdır.

Varyans

Yayılım çoğu zaman çeşitli şans eseri hallerin bir sonucudur. Mesela, bir sıvı içerisinde sürüklenen mikroskopik parçacıkların hareketini mülahaza edelim. Herhangi bir parçacığın tahmin edilemez hareketi, diğer parçacıklarla rastgele çarpışmaların neden olduğu muhtelif yönlerdeki pek çok küçük hareketin sonucudur. Eğer, biz, tüm parçacıkların başlangıç noktalarından itibaren toplam yer değişimlerinin ortalamasını alırsak, sonuç zamanla doğru orantılı artmayacaktır. Mamafih, eğer tüm parçacıkların toplam yerdeğişimlerinin *karelerinin* ortalamasını alırsak, bu sonuç zamanla doğru orantılı artacaktır.

Aynı şey terliksi hayvanın hareketi için de geçerlidir. Terliksi hayvanların, bir mikroskop altında rastgele görünen hareketleri de gözlendiğinde, hakeza, başlangıç noktalarından itibaren yerdeğişimlerinin karelerinin ortalamasının zamanla doğru orantılı olduğu görülebilir. Benzerce, yüksek bir binanın tepesinden aynı noktadan aşağıya doğru pinpon toplarını bıraktığımızı düşünelim, toplar havada şans eseri savrulurlar. Zeminde topların fiilen düştükleri noktaların merkezi bir noktadan sapmalarını ölçebiliriz. Binanın yüksekliği arttıkça, sapmaların karelerinin ortalaması orantısal olarak artacaktır.

Pek çok halde, filhakika, şans etkilerinin çokluğu hasebiyle, aritmetik ortalama etrafında gözlenen yayılımı ölçmeye çalışırız. Bu hallerde mütenasip metot aritmetik ortalamadan sapmaların karelerinin aritmetik ortalamasıdır; bunu **varyans** diye adlandırırız ve popülasyonun varyansını σ^2 ile simgeleriz (σ, Yunan alfabesinin küçük *sigma* harfidir):

$$\sigma^2 = \frac{\sum(y-\mu)^2}{N}.$$

Bir örneğin varyansı s^2 ile simgelenir ve popülasyonun varyansından biraz farklı olmak üzere, şöyle

hesaplanır,

$$s^2 = \frac{\sum(y-\bar{y})^2}{n-1}$$

ki paydada yer alan $n-1$ terimi dar örneklerin yanlı (yani, popülasyon varyansından daha küçük çıkma eğilimi gösteren) varyansı için matematiksel bir düzeltmedir ve örnek hacmi artarken ($n > 30$), katkısı ihmal edilebilecek kadar küçülür.

Misal 6 Aston Martin otomobil markasının Vanquish modelinin Avrupa'da 2017 yılı satış sayıları aylara göre müteakip tabloda veriliyor. Bu veri kümesini bir popülasyon olarak mülahaza etmek kaydıyla, varyansı nedir?

Ay:	Oc	Şub	Mar	Nis	May	Haz	Tem	Ağ	Ey	Ek	Kas	Ar
Satış sayısı:	27	15	71	11	37	33	33	17	37	26	22	30

Cevap: Bir hesap makinası, bir hesap tablosu yazılımı veya basit bir istatistiksel yazılım kullanılmak suretiyle, varyans çabucak bulunabilir; ya da şu şekilde basitçe hesaplayabilirsin:

$$\mu = \frac{27+15+71+11+37+33+33+17+37+26+22+30}{12} = 29.91 \text{ adet}$$

$$\sigma^2 = \frac{(27-29.91)^2+(15-29.91)^2+\cdots+(22-29.91)^2+(30-29.91)^2}{12} = 220.08 \text{ adet}^2$$

Misal 7 Organik ürünlerle beslenmeye meraklı bir aile, bahçesindeki kümeste yetiştirmiş olduğu tavukların yumurtalarından 4 tanesini rastgele örnekleyerek tartıyor (gr): $Y = \{48, 65, 80, 59\}$. Varyans nedir?

Cevap: Yine bir hesap makinası, bir hesap tablosu yazılımı veya basit bir istatistiksel yazılım kullanılmak suretiyle, ya da örneğin aritmetik ortalamasını,

$$\sum y = 48+65+80+59 = 252 \text{ gr},$$

$$\bar{y} = \frac{\sum y}{n} = \frac{252}{4} = 63 \text{ gr},$$

hesapladıktan sonra varyansını hesaplayabiliriz,

$$\sum(y-\bar{y})^2 = (48-63)^2+(65-63)^2+(80-63)^2+(59-63)^2 += 534 \text{ gr}^2,$$

$$s^2 = \frac{\sum(y-\bar{y})^2}{n-1} = \frac{534}{4-1} = 178 \text{ gr}^2.$$

Standart sapma

Belli bir popülasyonun değişkenliğini temsil edecek bir değer tespit etmek istiyoruz. Şimdiye kadar yaptığımız tartışma gösteriyor ki bu değer, doğal olarak, aritmetik ortalamadan sapmaların karelerinin aritmetik ortalamasının karekökü̇dür. Yani, varyansın karekökü, ki bu değere **standart sapma** adı verilir, popülasyon için σ ile simgelenir ve şöyle hesaplanır:

$$\sigma = \sqrt{\frac{\sum(y-\mu)^2}{N}}$$

Benzerce, bir örneğin standart sapması s ile simgelenir ve şöyle hesaplanır:

$$s = \sqrt{\frac{\sum(y-\bar{y})^2}{n-1}}$$

Tekrar dikkatini çekmek istiyorum ki varyans kareli birimlerle ölçülürken, standart sapma veriyle aynı birimlerde ölçülür. Muhtelif y değerlerinin $y - \bar{y}$ sapmaları için s tipik bir değerdir. Şunu, ancak, iyi anlamak gerekir ki s sapmaların doğrudan aritmetik ortalaması değildir (sapmaların aritmetik ortalaması sıfırdır), s bize y değerlerinin aritmetik ortalama etrafındaki yayılımının bir ölçüsünü verir.

Misal 8 Bir istatistik sınavında sizden müteakip verinin standart sapmasını hesaplamanız istenmiş olsaydı ne yapardınız?

$$Y = \{1, 6, 3, 8\}$$

Cevap: Eğer bu bir popülasyon verisi ise,

$$\sigma = 2.693;$$

eğer bir örnek verisi ise,

$$s = 3.109.$$

Nispi Değişkenlik

İki veri kümesinin değişkenliklerinin varyansları veya standart sapmaları cinsinden biribiriyle karşılaştırılması, eğer bu veri kümelerinin aritmetik ortalamaları da göz önüne alınırsa, daha manalıdır. Bir veri kümesinin standart sapmasını aritmetik ortalamasına bölmek suretiyle *nispi değişkenliğinin* bir tanımını yapabiliriz. Sonuç genellikle yüzde olarak ifade edilir. Bu ölçüye *varyasyon katsayısı* da denir.

Mesela, istatistik dersinin vize sınavında kızların puanlarının aritmetik ortalaması 45 ve standart sapması 9 ve oğlanların puanlarının aritmetik ortalaması 35 ve standart sapması 7.7 ise, bu veri kümelerinin nispi değişkenlikleri sırasıyla $\frac{9}{45} = 0.2$ ve $\frac{7.7}{35} = 0.22$ veya %20 ve %22 dir. Buna göre, kızların puanlarının değişkenliği mutlak olarak daha fazlayken, aritmetik ortalamalar arasındaki fark göz önüne alındığı zaman, oğlanların puanlarının değişkenliği nispi olarak daha fazladır.

Pozisyonun Ölçülmesi

Bir dağılımın merkezini temsil edecek bir değeri tayin etmenin çeşitli yollarını gördük. Bununla beraber, herhangi başka değerlerin *pozisyonları* hakkında da konuşabilmek isteriz. Bazı hallerde, mesela şarap tadımında olduğu gibi, basitçe sıralarla ilgilenilir. Başka hallerde, mesela okul başvurularının değerlendirilmesinde öğrencilerin yüzdelik sıralarına bakılabilir. Aynı zamanda, öyle haller vardır ki hem merkezileşme ve hem de değişkenlik ölçüleri kullanılarak pozisyon belirlenebilir.

Gözlem değerlerinin pozisyonlarını ifade etmenin üç önemli prosedürü şunlardır:

1. *Basit sıralama*, ki gözlem değerlerini bir sıraya sokarak, belirli bir değerin hangi sıraya düştüğünü kaydetmektir.
2. *Yüzdelik sıralaması*, ki ilgilenilen gözlem değerinden daha küçük değerlerin tüm değerler içerisindeki yüzdesini (başka bir deyişle, ilgilenilen gözlem değerinin tüm değerlerin yüzde kaçından daha büyük olduğunu) belirtir.
3. *Z-puanı*, ki bir gözlem değerinin aritmetik ortalamanın kaç standart sapma uzağında olduğunu çok spesifik olarak ifade eder.

Misal 9 Otuz yedi otomobil markasının Avrupa'da 2017 yılındaki pazar payları (% olarak) sırasıyla şöyledir: 10.9, 7.3, 6.7, 6, 5.9, 5.9, 5.3, 5.3, 5, 4.5, 4.4, 3.7, 3.6, 3.3, 3, 3, 2.6, 1.9, 1.6, 1.5, 1.4, 1.1, 0.9, 0.7, 0.7, 0.6, 0.6, 0.5, 0.4, 0.4, 0.3, 0.3, 0.2, 0.2, 0.1, 0.1 ve 0.1. Pazar payı %4.4 olan Toyota'nın pozisyonu nedir?

Cevap: Toyota basit sıralamada yukarıdan 11 incidir (37 marka içerisinde), çünkü listede kendisinden daha fazla pazar payına sahip 10 marka var. Yirmi altı markanın pazar paylarının daha düşük olması hasebiyle, Toyota'nın yüzedelik sıralaması $\frac{26}{37} \times 100 = \%70.3$ tür. Listelenen verinin

aritmetik ortalaması %2.7 ve standart sapması %2.6 olduğuna göre, Toyota $\frac{4.4-2.7}{2.6} = 0.65$ lik bir z-puanına sahiptir.

Basit Sıralama

Basit sıralama kolayca hesaplanır ve kolayca anlaşılır. Birisi için 324 kişilik bir sınıfta üçüncü olarak mezun olmanın veya bir sporcu için 64 kişinin katıldığı bir spor müsabakasında dokuzuncu gelmenin ne manaya geldiğini biliriz. Hatta verinin sayılardan oluşmadığı durumlarda bile, basit sıralama yararlıdır. Mesela, deterjanlar, herhangi bir sayısal ölçüm yapılmaksızın, nispi temizleme güçlerine göre sıralanabilir.

Yüzdelik Sıralaması

Yüzdelik sıralaması, diğer bir kolayca anlaşılır pozisyon ölçüsü olarak, farklı veri kümelerindeki pozisyonları karşılaştırmada yararlıdır. Mesela, 704 birey içerisindeki 176 ıncı sırayı 935 birey içerisindeki 187 inci sırayla, ilkinin %75 inci sıraya ve ikincinin %80 inci sıraya sahip olduğunu kaydetmek suretiyle, kolayca karşılaştırabiliriz. Bir veri kümesindeki bir gözlemin *yüzdeliği* (*/percentile*) o gözlemin tüm diğer gözlemlere göre pozisyonunu temsil eder ve kendisinden daha küçük gözlemlerin yüzdesini verir. Bir dağılımın *p-inci yüzdeliği* gözlemlerin %p'sinin daha küçük veya eşit olduğu değerdir. Bir yüzdeliği hesaplamak için gözlemleri artan sırada düzenler ve listenin altından başlayarak gerekli yüzdeye kadar sayarız. Her zaman verinin tam %p'sinin daha küçük veya eşit olduğu bir değer mevcut olamayabileceğinden bu tanım biraz eksik kabul edilebilir. Belki çoğu yüzdelik için en yakın gözlemi almak suretiyle bu eksikliği giderme yolunu seçebiliriz.

Çeyreklikler (*/quartiles*), Q_1 ve Q_3, küçükten büyüğe doğru sıralanmış bir listenin ilk çeyreğini ve ilk üç çeyreğini işaretlerler, sırasıyla. Bunların yüzdelik sıraları %25 ve %75 tir, sırasıyla. Daha önce çeyreklikleraası gidimi, aynı zamanda, $Q_3 - Q_1$ olarak tanımlamıştık.

Ondalıklar (*/deciles*) küçükten büyüğe doğru sıralanmış bir listenin birinci onda birden dokuzuncu onda bire kadar her onda birini işaretlerler, sırasıyla ve %10 dan %90 a kadar yüzdelik sıralara tekabül ederler.

Z-Puanı

Bir z-puanı dağılımın hem merkezini (yani aritmetik ortalamasını) ve hem de yayılımını (yani standart sapmasını) hesaba katarak veri değerlerini/noktalarını yatay eksende yeniden pozisyonlandıran bir ölçüdür. Daha spesifik olmak gerekirse, bir ham değerin z-puanı bu ham değerin ait olduğu veri kümesinin aritmetik ortalamasından kaç standart sapma uzakta olduğunu bize söyler. Matematiksel olarak, y nin μ ye ham uzaklığı $y - \mu$ dür; bunu σ ya bölerek kaç standart sapma ettiğini buluruz. Yani ham değerin aritmetik ortalamaya olan uzaklığı kaç standart sapma ise *z-puanı* da odur. Böylece,

$$z = \frac{y - \mu}{\sigma}$$

ki burada y ham değer, μ aritmetik ortalama ve σ standart sapmadır. Eğer y değeri aritmetik ortalamadan (μ) daha büyük ise, *z-puanı* pozitiftir; eğer y değeri aritmetik ortalamadan (μ) daha küçük ise, *z-puanı* negatiftir. Z-puanlarına aynı zamanda *standartlaştırılmış (veya standart) puanlar* da denir.

Bir z-puanı verildiği zaman, prosedürü ters-düz ederek mütekabil ham değeri bulabiliriz. Eğer yukarıdaki ifadeyi y ye göre çözersek, $y = \mu + z\sigma$ formunu alır.

Misal 10 Bir öğrenme-çalışmasının sonucunda 18−24 yaşları arasındaki kadınların boylarının ortalaması (aritmetik ortalaması) 166.5 cm ve standart sapması 6.5 cm olarak tespit edilmiştir. O halde, 179.5 cm'lik boy için z-puanı,

$$\frac{179.5 - 166.5}{6.5} = 2$$

ve 147 cm'lik boy için *z*-puanı,

$$\frac{147 - 166.5}{6.5} = -3.$$

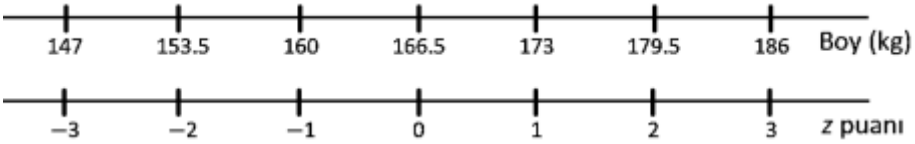
Diğer taraftan, 2.2 lik bir *z-puan*ının tekabül ettiği ham değer,

$$166.5 + 2.2(6.5) = 166.5 + 14.3 = 180.8 \text{ cm}$$

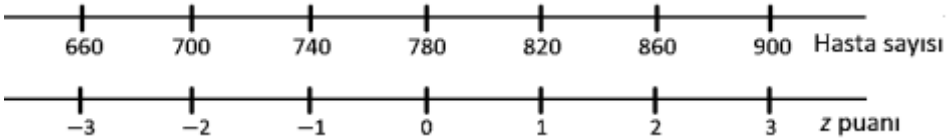
ve -1.6 lık bir *z*-puanının tekabül ettiği ham değer,

$$166.5 + (-1.6)(6.5) = 166.5 - 10.4 = 156.1 \text{ cm}.$$

Çoğu zaman tam sayı *z*-puanlarını ve tekabül ettikleri ham değerleri aşağıdaki gibi göstermek yararlıdır:



Misal 11 Bir sağlık ocağına müracaat eden günlük hasta sayısının aritmetik ortalaması 780 kişi ve standart sapması 40 kişidir. Yediyüz seksen kişilik aritmetik ortalamaya 40 ın katlarını ekler ve çıkartırsak şunu elde ederiz:

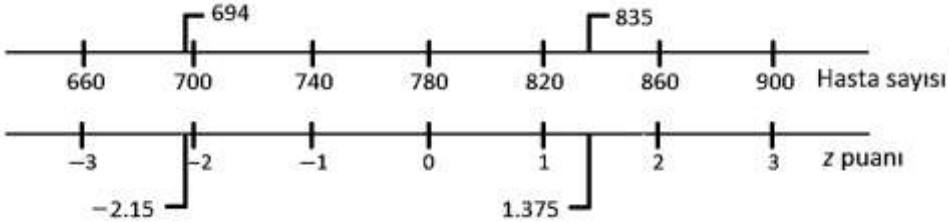


Bir günlük hasta sayısı 835 ise buna tekabül eden *z-puan*ı:

$$\frac{835 - 780}{40} = \frac{55}{40} = 1.375.$$

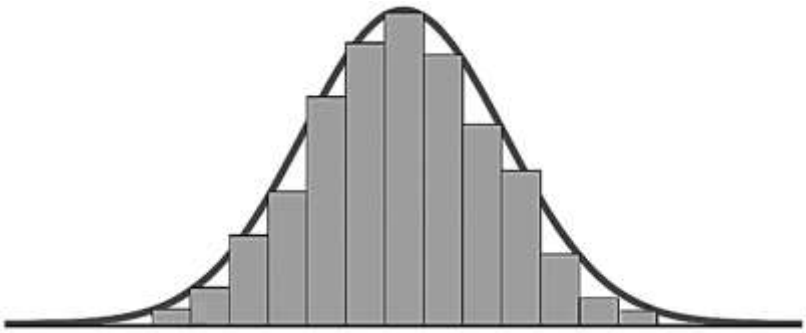
Diğer taraftan, -2.15 lik bir *z-puan*ına tekabül eden günlük hasta sayısı:

$$780 + (-2.15)(40) = 694 \text{ kişi}.$$



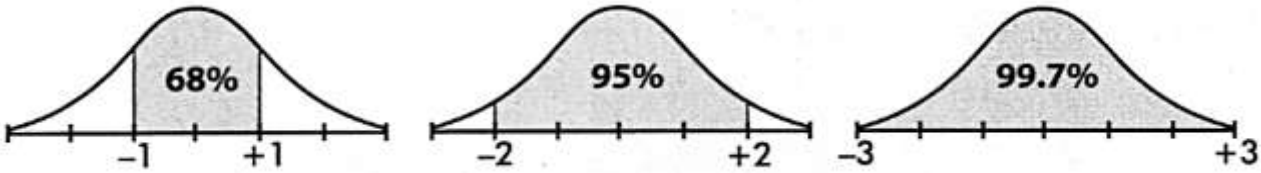
Ampirik Kural

Verinin dağılımı simetrik ve çan-şekilli olduğunda, *ampirik kural* adı verilen spesifik bir kural uygulanır ki aynı zamanda 68-95-99.7 kuralı olarak ta adlandırılır. Bu kurala göre, gözlem değerlerinin yaklaşık %68 i aritmetik ortalamanın sağlı-sollu 1 standart sapma civarına düşer, gözlem değerlerinin yaklaşık %95 i aritmetik ortalamanın sağlı-sollu 2 standart sapma civarına düşer ve gözlem değerlerinin %99.7 si aritmetik ortalamanın sağlı-sollu 3 standart sapma civarına düşer.



Misal 10 daki 18-24 yaşları arasındaki kadınların boylarının aritmetik ortalaması 166.5 cm ve standart sapması 6.5 cm olarak tespit edilmişti. Dağılımın çan-şekilli olduğunu varsayarak, 68-95-99.7 kuralına göre dağılımın ortada kalan %95'inin 153.5 cm ve 179.5 cm arasında boyları olduğunu söyleyebiliriz. Bunun gibi, bu kadınların en uzun %2.5 u 179.5 cm den daha uzundur. (Kadınların dışarıda kalan %5 i aritmetik ortalamanın 2σ ya da 13.0 cm daha uzağına düşer. Dağılımın simetrik olmasından dolayı, bu kadınların yarısı bir uca yarısı ise diğer uca düşer.) Yine 68-95-99.7 kuralına göre neredeyse tüm kadınların (%99.7) boylarının 147.0 ile 186.0 cm arasında olduğunu söyleyebiliriz.

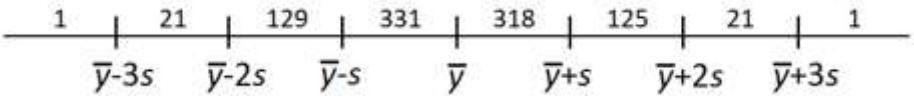
Müteakip figürde yer alan nispi frekans grafiklerinde yatay eksen *z*-puanlarını göstermektedir:



Misal 12 Biliniyor ki standart bir teste giren öğrencilerin testi bitirme sürelerinin aritmetik ortalaması 75 dakika ve standart sapması 12 dakikadır. Ampirik kural bize hangi enformasyonu temin eder?

Cevap: Dağılımın çan-şekilli olduğunu varsayarak, öğrencilerin yaklaşık %68 inin testi 63 dakika ve 87 dakika arasında bitirdikleri, yaklaşık %95 inin testi 51 dakika ve 99 dakika arasında bitirdikleri ve neredeyse taümünün (%99.7) testi 39 dakika ve 111 dakika arasında bitirdiği neticesine varabiliriz.

Misal 13 Dokuz yüz kırk yedi öğrenciye uygulanan bir dil testi sonucunda puanların aritmetik ortalaması 6.86 ve standart sapması 1.59 bulunmuştur. Daha sonra bu puanlar yatay eksen üzerinde aritmetik ortalamanın 1*s*, 2*s* ve 3*s* etrafında frekanslarına göre şöyle dağılmıştır:



Bu enformasyona göre, gözlem değerlerinin %68.5 i $[= 100(649/947)]$ aritmetik ortalamanın 1*s* etrafında, %95.4 ü $[= 100(903/947)]$ aritmetik ortalamanın 2*s* etrafında ve %99.8 i $[= 100(945/947)]$ aritmetik ortalamanın 3*s* etrafında yer almaktadır. Dağılımın simetrik ve çan-şekilli olduğu kolayca görülebiliyor ve 68-95-99.7 kuralına uygunluğu ise açık.

Ampirik kural, aynı zamanda bize standart sapmanın gidim cinsinden çabuk bir kestirimini de verir. Yukarıdaki figürde görebileceğin üzere, verinin %95 i 4 standart sapmalık bir aralığa (yani, *z*-puanları ekseninde -2 den +2 ye kadar olan aralığa) düşer ve verinin %99.7 si 6 satandart sapmalık bir aralığa (yani, *z*-puanları ekseninde -3 ten +3 e kadar olan aralığa) düşer. Bu sebeple, çan eğrisi şeklinde bir frekans dağılımına sahip olması koşulu ile, bir veri kümesinin standart sapmasının, kabaca, gidiminin dörtte biri ve altıda biri arasında olduğu neticesine varabiliriz. Bir veri kümesinin gidimini neredeyse anında bulabileceğimiz için, standart sapmanın kestirimi için ampirik kural çoğu zaman absürt aritmetik hataların görülmesinde yararlıdır.

Misal 14 Eğer bir veri kümesinin gidimi 12 birim ise, standart sapmasının bir kestirimi nedir?

Cevap: Ampirik kurala göre, standart sapmanın $\left(\frac{1}{6}\right)12 = 2$ birim ve $\left(\frac{1}{4}\right)12 = 3$ birim arasında

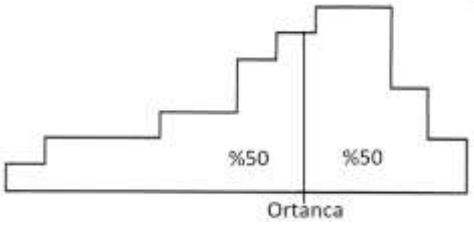
olmasını bekleyebiliriz. Eğer standart sapmayı 0.04 birim veya 11 birim olarak hesaplamışsan, muhtemelen aritmetiksel bir hata yapmışsındır; bununla beraber, 3.2 birimlik bir hesaplama akla yatkındır.

Yine de üzerine bastırarak söylemek gerekir ki gidimi yukarıdaki gibi kullanmak, standart sapma için isabetli bir değer elde etmeye matuf değildir. Ancak ve sadece, mesela kompüter çıktılarını körü körüne kabul etmek gibi, mantıksız sonuçları işaret etmekte yararlı bir araçtır.

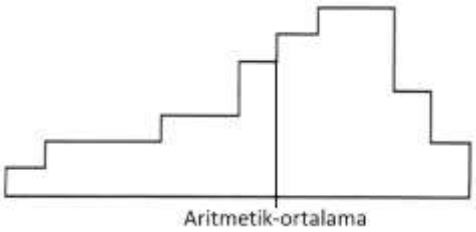
Bir frekans dağılımının merkezine ait sayısal betimlemeyi o frekans dağılımını meydana getiren gözlem değerleri için seçilen bir tipik değer olarak ve yayılımına ait sayısal betimlemeyi de gözlem değerlerinin seçilen bu tipik değere ne kadar benzediklerinin bir ölçüsü olarak düşünebilirsin.

Histogramlar ve Merkezileşme Ölçüleri

Bir verinin histogramı çizildiği zaman, o verinin merkezileşmesi, dağılımın şekli itibarıyla, kendini gösterir. Farzedelim ki elimize aşağıdaki üzere bir histogram var, merkezileşme ölçülerimiz doğal olarak şöyle bir durumda olacaktır:



Ortanca bir dağılımı eşit iki parçaya böler ve histogramın alanını yarıya bölen bir çizgi ile temsil edilir.



Aritmetik ortalama değerlerin eksen üzerindeki yerinden (yani büyüklüğünden) etkilenir. Böylece, eğer histogram katı bir alan olarak mülahaza edilirse, aritmetik ortalama ağırlık merkezinden veya denge noktasından geçen bir çizgiye tekabül eder.

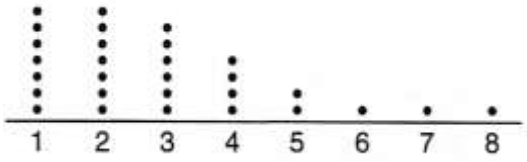
Yukarıdaki dağılım, küçük değerler yönünde incelerek ve uzayarak yayılmaktadır ki böylesi dağılımlara *sola çarpık* denir. Dikkatini çekmek isterim ki aritmetik ortalama, bu durumda, ortancadan çoğu zaman daha küçüktür. Benzerce, büyük değerler yönünde incelerek ve uzayarak yayılan bir dağılıma *sağa çarpık* denir ve aritmetik ortalaması çoğu zaman ortancasından daha büyüktür.

Misal 15 Bir büyük şirketin çalışanlarına ödediği yıllık ücretlerin ortancası 32500 TL ve aritmetik ortalaması 38700 TL dir. Yıllık ücretlerin dağılımının şekli hakkında, bu ne ifade eder?

Cevap: Ortanca aritmetik ortalamadan daha az ve böylece yıllık ücretlerin dağılımı muhtemelen sağa çarpık. Ücret skalasının aşağısında yer alan pek çok çalışanla beraber, çok az sayıda yüksek ücretli yöneticiler mevcut.

Şunu not etmek durumundayım ki yukarıdaki prensip çok kullanışlı olmakla beraber istisna kabul etmez bir kural değildir.

Misal 16 Aşağıdaki nokta grafikte veri kümesi sağa çarpıktır; mamafih, ortancası (ki 3) aritmetik ortalamasından (ki 2.97) daha büyüktür.



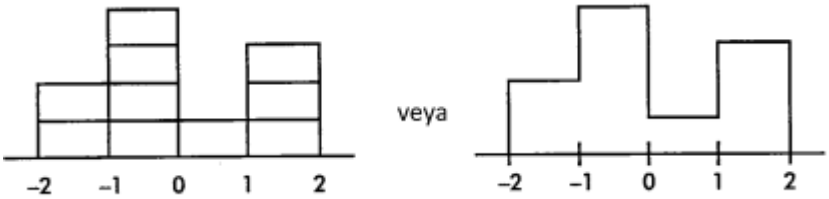
Histogramlar, *Z*-puanları ve Yüzdelik Sıralamaları

Görmüştük ki nispi frekanslar histogramlarda nispi alanlarla temsil edilmektedirler ve böylece dikey eksenin nasıl etiketlendiğinin fazlaca önemi yoktur. Eğer standart sapmayı biliyorsak, yatay eksen *z*-puanları cinsinden etiketlenebilir. Aslında, eğer bize muhtelif *z*-puanlarının yüzdelik sıraları verilirse, bir histogram inşa edebiliriz.

Misal 17 Farzedelim ki müteakip veriden bir histogram inşa etmemiz isteniyor:

z-puanı:	-2	-1	0	1	2
Yüzdelik sıralaması:	0	20	60	70	100

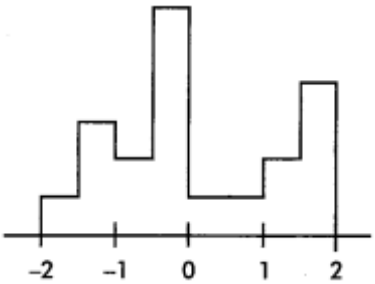
Görüyoruz ki alanın tamamı +2 *z*-puanı ve -2 *z*-puanı arasındadır. Aynı zamanda, alanın %20 si -2 ve -1 *z*-puanları arasında, %40 ı -1 ve 0 *z*-puanları arasında, %10 u 0 ve 1 *z*-puanları arasında ve %30 u 1 ve 2 *z*-puanları arasındadır. Buna göre, histogram aşağıdaki üzeredir:



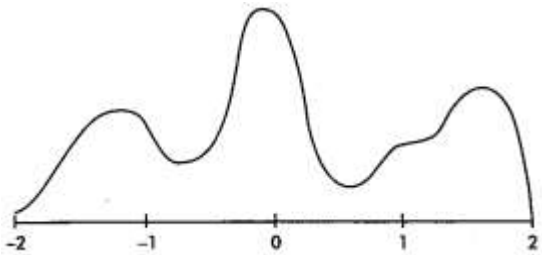
Şimdi farzedelim ki yukarıda verilmiş olan *z*-puanlarının arasına düşen dört yeni *z*-puanı ve yüzdelik sıralamaları daha verilmiş olsun:

z-puanı:	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2
Yüzdelik sıralaması:	0	5	20	30	60	65	70	80	100

Buna göre, histogramımızın yeni hali şudur:



Şimdi de düşünelim ki 1000 tane *z*-puanımız ve yüzdelik sıralamaları var, histogram belki de şöyle görünecek:



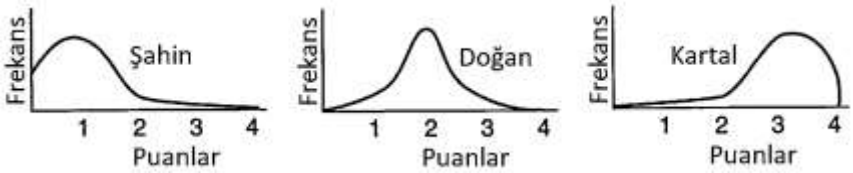
Grafiğin, yatay eksenin herhangi bir noktasındaki yüksekliği manasızdır; önemli olan nispi alanlardır. Mesela, yukarıdaki sonuncu diyagramda +1 ve +2 *z*-puanları arasındaki alanın yüzdesi nedir? Cevap: Hala %30. Sıfırın solundaki alanın yüzdesi kaçtır? Cevap: Hala %60.

Birikimli Frekans ve Çarpıklık

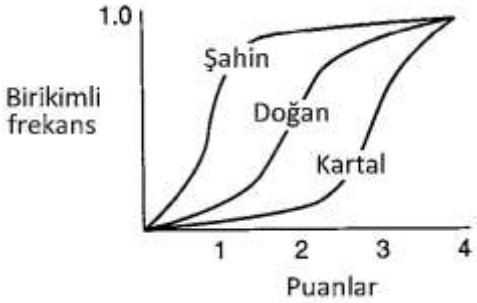
Sola çarpık bir dağılımın birikimli frekans grafiği önce hafif ve sonra şiddetli bir eğimle yükselirken, sağa çarpık bir dağılımın birikimli frekans grafiği önce şiddetli sonra hafif bir eğimle yükselir.

Misal 18 Üç ayrı istatistik profesörünün verdikleri sözlü sınavı puanlarının dağılımlarına ait

histogramlar şöyledir:



Histogramlara göre, Bay Şahin daha çok sayıda düşük puanlar veriyor, Bay Doğan eşit sayılarda düşük ve yüksek puanlar veriyor ve Bayan Kartal daha çok sayıda yüksek puanlar veriyor. Bu histogramların birikimli frekans grafikleri şöyledir:

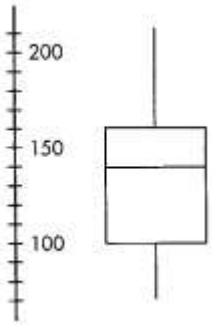


Beş-sayıyla Özet ve Kutu Grafikler

Bir veri kümesinin beş-sayıyla özeti o veri kümesinin en küçük değerini (minimum), en büyük değerini (maksimum), ortadaki değerini (ortanca), alttaki yarısının ortasını (Q_1) ve üstteki yarısının ortasını (Q_3) kapsar. Beş-sayıyla özet bilhassa geniş veri kümelerinin betimsel analizinde yararlıdır.

Verinin beş-sayıyla özeti *kutu grafik* (aynı zamanda kutunun her iki yanında uzayan bıyıkları da vardır) adı verilen bir grafik ile gösterilebilir. Kutu grafik verinin yayılımının görsel bir bir temsilidir. Bir kutu grafiğe bakarak verinin biribirini takibeden %25 lik kısımlarının gidimlerini karşılaştırabilirsin. Dağılımın şekli ve dışlaklar hakkında fikir sahibi olabilirsin.

Misal 19 Toplam tarımsal ürün endeksinin 27 yıl boyunca aldığı değerler yıllar itbarıyla şöyledir: {215, 210, 130, 140, 150, 150, 160, 150, 140, 150, 150,125, 85, 70, 75, 90, 115, 120, 125, 100, 95, 100, 130, 160, 200, 200, 210}. En büyük değer (maksimum) 215, en küçük değer (minimum) 70, ortadaki değer (ortanca) 140, üstteki yarımın ortasındaki değer (üçüncü çeyreklik) 160 ve alttaki yarımın ortasındaki değer (birinci çeyreklik) 100. Bir istikrarsızlık, çiftçiler için, ürün fiyatları bakımından, söz konusu. Bu beş-sayının bir kutu grafiği şöyledir:



Dikkat etmeni isterim ki bu görsel biribirine yapışık iki "kutu" ve iki yandaki "bıyıklardan" müteşekkil. (Bu grafiğin, zaten, İngilizce'deki adı "*box and whiskers*".) Kutular, ortadaki iki çeyreğin yayılımını gösteriyor; bıyıklar, dıştaki iki çeyreğin yayılımını gösteriyor. Bu nispeten basit gösterim, histogramların veya dal-yaprak grafiklerin detaylıca ortaya koyacağı enformasyonu bize bir bakışta aksettirir.

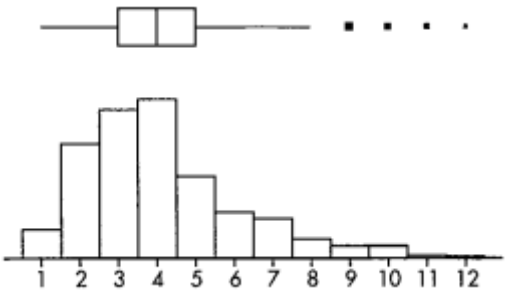
Bir dağılım ağırca çarpık olduğu zaman, ya da açıkça dışlaklara sahip olduğu zaman, bu dağılımın ortancasını, çeyrekliklerini ve aşırı değerlerini havi bir kutu grafik çizmek suretiyle, aritmetik

ortalamasını ve standart sapmasını hesaplamaktan daha yararlı bir betimlemesini yapmış olursun.

Misal 20 Shakespare'in oyunlarında geçen kelimelerin uzunluklarının beş-sayıyla özeti şöyledir:

Minimum = 1, Q_1 = 3, Ortanca = 4, Q_3 = 5, Maksimum = 12.

$Q_1 - 1.5IQR = 3 - 1.5(5-3) = 0$ ve $Q_3 + 1.5IQR = 3 + 1.5(5-3) = 8$ olduğuna göre, bunlardan daha küçük ve daha büyük kelime uzunlukları dışlak zanlılarıdır. İstatistiksel yazılım paketlerince çizdirilen kutu grafikler çoğu zaman dışlakları işaret eder ve ayrıca verinin bir histogramını da verir.

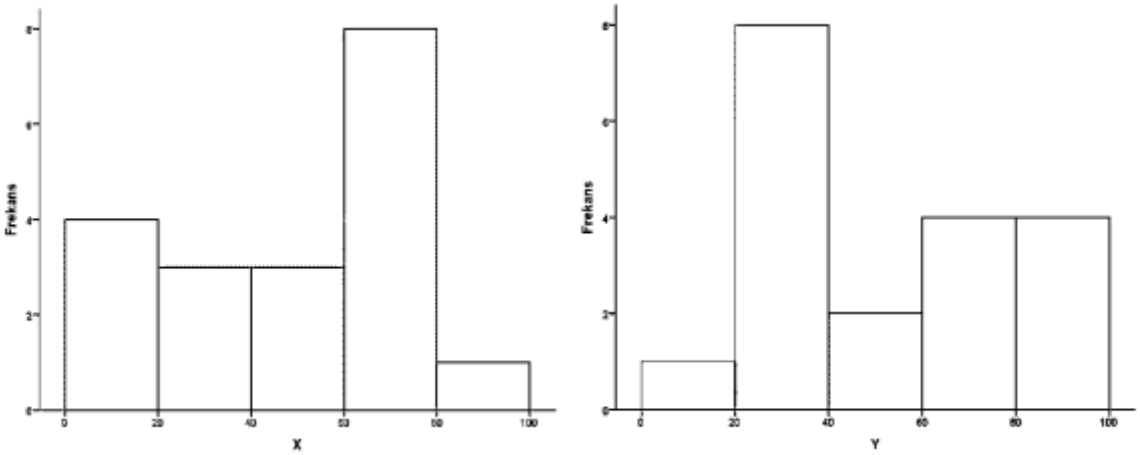


Nihayetinde, şunu da not etmek gerekir ki farklı iki veri kümesinin beş-sayıyla özetleri aynı olabilir ve böylece aynı kutu grafiklere sahip olabilirler, fakat bu veri kümelerinin dağılımları aynı zamanda biribirinden hayli farklı olabilir de.

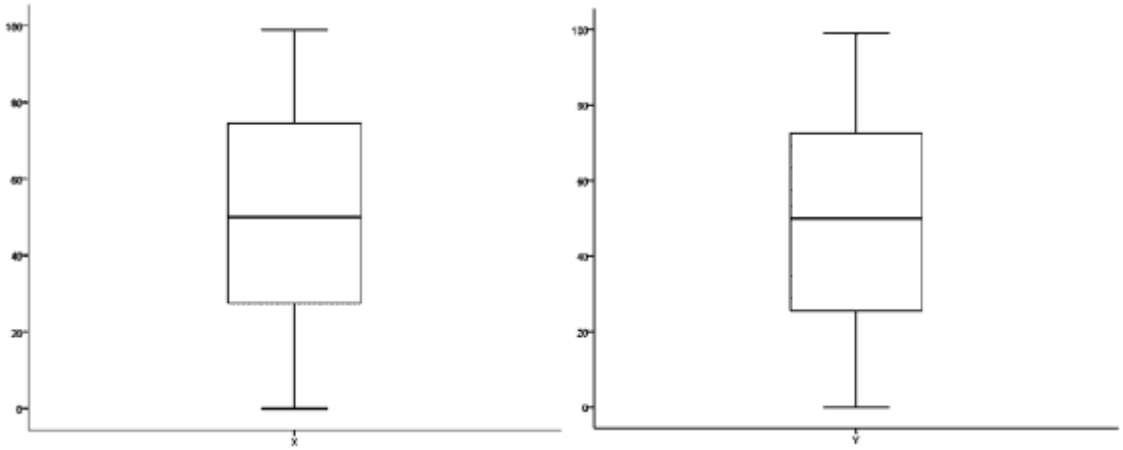
Misal 21 Şu iki veri kümesine sahibiz:

$X = \{0, 5, 10, 15, 25, 30, 35, 40, 45, 50, 71, 72, 73, 74, 75, 76, 77, 78, 100\}$ ve

$Y = \{0, 22, 23, 24, 25, 26, 27, 28, 29, 50, 55, 60, 65, 70, 75, 85, 90, 95, 100\}$.



Histogramlarına baktığımız zaman, biribirinden oldukça farklı iki dağılım görüyoruz. Oysa, her iki değişken için, Minimum = 0, Q_1 = 25, Ortanca = 50, Q_3 = 75, Maksimum = 100 ve dolayısıyla X ve Y değişkenleri biribirinin aynı iki kutu grafikle betimlenecektir:



Birim Değiştirme

Birimlerin, mesela Türk Lirası'ndan Amerikan Doları'na ya da milden kilometreye ya da kilogramdan libreye dönüştürülmesi hayli sık uygulanan işlemlerdendir. Peki ölçü birimini değiştirmek, bir veri kümesinin merkez ve yayılım ölçülerini nasıl etkiler?

Veri kümesindeki her değere aynı sabiti eklemek (veya çıkarmak), aritmetik ortalamayı ve ortancayı aynı sabit miktarınca büyütür veya küçültür; bununla beraber, büyütülmüş veya küçültülmüş değerler arasındaki mesafeler yine aynı kaldığı için gidim ve standart sapma değişmez.

Misal 22 Bir bilimci belirli bir sıvının donma noktasını öğrenmeye-çalışırken yaptığı deneysel ölçümlerin aritmetik ortalamasını 25.32 derece Celsius ve standart sapmasını 1.47 derece Celsius olarak bulur. Eğer tüm ölçümler Kelvin ölçeğine dönüştürülürse, yeni aritmetik ortalama ve standart sapma nedir?

Cevap: Biliyoruz ki $\text{Kelvin} = \text{Celsius} + 273.16$. Buna göre, yeni aritmetik ortalama $25.32 + 273.16 = 298.48$ derece Kelvin. Mamafih, standart sapma sayısal olarak aynı kalır, yani 1.47 derece Kelvin. Grafiksel olarak, dağılım yatay eksen üzerinde 273.16 sabitince kayar; dağılımın merkezi kayar (yani aritmetik ortalama değişir), fakat yayılımı aynı kalır (yani standart sapma değişmez).

Veri kümesindeki her bir değeri aynı sabitle çarpmak (veya bölmek), aritmetik ortalamayı, ortancayı, gidimi ve standart sapmayı aynı sabit çarpımınca büyütür veya küçültür.

Misal 23 Tarım işletmelerinin arazi genişliklerine ait ölçümlerin aritmetik ortalaması 9.36 hektar ve standart sapması 2.24 hektardır. Eğer tüm ölçümler hektardan dekara dönüştürülürse, yeni aritmetik ortalama ve standart sapma ne olur?

Cevap: Bir hektar, 10 dekara eşittir. Buna göre, yeni aritmetik ortalama $10(9.36) = 93.6$ dekar ve yeni standart sapma $10(2.24) = 22.4$ dekar olur. Grafiksel olarak, veri kümesindeki her bir değeri 10 sabitiyle çarpmak, dağılımı hem yerinden kaydırmış hem de yayılımını artırmıştır.

Misal 24 Bir şirkette çalışan altı kişinin maaşları (TL) sırasıyla şöyledir: 2000, 5000, 5000, 7000, 8000 ve 9000.

a. Ortanca ve aritmetik ortalama maaş nedir?
Cevap: Ortanca,

$$\frac{5000 + 7000}{2} = 6000$$

Aritmetik ortalama,

$$\frac{2000 + 5000 + 5000 + 7000 + 8000 + 9000}{6} = 6000\text{TL}$$

b. Eğer çalışanların her biri 1000TL zam alırsa, yeni aritmetik ortalama ne olur?
Cevap: Ortanca,

$$\frac{6000 + 8000}{2} = 7000$$

Aritmetik ortalama,

$$\frac{3000 + 6000 + 6000 + 8000 + 9000 + 10000}{6} = 7000\text{TL}$$

Dikkat edersen, hem ortanca hem de aritmetik ortalama için, $6000\text{TL} + 1000\text{TL} = 7000\text{TL}$.

c. Eğer her biri %10 zam alırsa?
Cevap: Ortanca,

$$\frac{5500 + 7700}{2} = 6600$$

Aritmetik ortalama,

$$\frac{2200 + 5500 + 5500 + 7700 + 8800 + 9900}{6} = 6600\text{TL}$$

Dikkat edersen, hem ortanca hem de aritmetik ortalama için, 6000TL nin %110 u 6600TL dir.

Yukarıda misallendirerek göstermeye çalıştığım üzere, veri kümesindeki her bir değere aynı sabiti eklemekle, ortancaya ve aritmetik ortalamaya aynı sabiti eklemiş oluruz. Benzerce, her bir değeri aynı sabitle çarpmakla, ortancayı ve aritmetik ortalamayı aynı sabitle çarpmış oluruz.

Mini Proje 03

MP 01 de topladığın verinin mütenasip tekniklerle sayısal betimlemesini yap.

BÖLÜM 04 | İLİŞKİLERİ BETİMLEME: BİR NİTEL VE BİR NİCEL DEĞİŞKEN

Giriş

Şimdiye kadar, tek bir değişkene ait ölçümler ile ilgilendik. Bir değişkeni ele alarak, gösterdiği dağılımı inceledik. Mamafih, istatistik biliminin pek çok önemli uygulamaları, iki veya daha çok sayıdaki değişkenin biribiriyle ilişkili olup olmadığını incelemeyi icap ettirir.

İlişkileri inceleme prosesi değişkenlerin tipiyle yakından alakalıdır. Yani iki değişken arasındaki bir ilişkiyi öğrenmeye-çalışırken, hangi metotları kullanacağımızı bu değişkenlerin tipleri belirleyecektir. Değişken tiplerinin üç kombinasyonunu ele alacağız: (1) Nitel-nicel, (2) nitel-nitel ve (3) nicel-nicel. Her üç halde, örnekleme ünitelerinin iki ayrı değişkene ait bir çift gözlem değeri veri toplayıcısı tarafından kaydedilmiş olur. Bu değer çiftlerine *iki-değişkeli veri* adı verilir.

İstatistik-biliminin pek çok gerçek yaşam uygulamaları iki popülasyonun karşılaştırılmasını ihtiva eder. Bu, genellikle, bir nicel değişkene ait verinin aynı bireyler üzerinde ölçülen başka bir nitel değişkenin kategorileri bakımından gösterdiği dağılımları karşılaştırmak suretiyle yapılabilir. Bu karşılaştırmalar, *nokta grafikleri*, *kutu grafikleri* ve *birikimli frekans grafiklerini* kullanarak, her iki veri kümesini eş zamanlı olarak görsel betimlemek suretiyle gerçekleştirilir. İki değişken arasındaki ilişki, nitel değişkenin farklı kategorilerine göre nicel değişkenin gösterdiği dağılımların şekillerine, merkezlerine ve yayılımlarına bakılarak değerlendirilir.

Akabinde, bu dağılımların merkezlerine ve yayılımlarına ait sayısal betimleyiciler hesaplanarak, yine biribirleriyle karşılaştırılabilir. Popülasyonları karşılaştırmak için kullanılan sayısal betimleyiciler dağılımların tipine bağlıdır. Bir frekans dağılımının *simetrik* veya *çarpık* olabileceğini daha önce öğrenmiştin. Eğer her iki popülasyon da simetrik dağılımlara sahipse, bunları karşılaştırmak için, aritmetik ortalama ve standart sapma kullanılabilir. Eğer popülasyonlardan en az birinin dağılımı çarpıksa, bunları karşılaştırmak için, *ortanca* ve *çeyreklikler arası gidim* kullanılabilir.

Bu, öte taraftan, şu soruya da cevap aramaktır ki karşılaştırılan iki örnek acaba aynı popülasyondan mı (veya bir başka deyişle, özdeş iki popülasyon mı) çekilmişlerdir? Her halükarda, iki popülasyonun ortalamaları arasındaki farkla ilgileneceğiz. Mesela, uygun örneklere dayanarak, belli bir işi adamların kadınlar kadar hızlı yapıp yapmadıklarına karar vermeyi arzu edebiliriz. Ya da, bir yoklamaya dayanarak, bir bölgedeki ailelerin aylık gıda harcamalarının bir başka bölgedekilerden fazla olup olmadığına karar vermek istiyebiliriz. İki değişken arasındaki ilişki bağlamında ifade edersek; birinci misalimizde cinsiyet ve iş yapma hızı arasındaki ilişkiyi, ikinci misalimizde ise bölge ve aylık gıda harcaması arasındaki ilişkiyi inceliyoruz.

Nihayetinde, bir nitel ve bir nicel değişken arasındaki ilişkiyi incelerken, bilimsel tahkikat nitel değişkenin kategorileri bakımından, nicel değişkenin gözlendiği istatistiksel popülasyonlar arasındaki doğal veya deneysel farkları değerlendirmeyi gerektirir.

Şimdi bir nitel ve bir nicel değişken arasındaki ilişkiyi incelemek için kullanabileceğimiz görsel ve sayısal betimleyicileri somut misaller üzerenden daha detaylıca ele alacağız.

Nokta Grafikler

Misal 1 İki farklı zayıflama programını (A programı ve B programı) uygulayan 25 er kişinin ağırlık kayıpları (kilogram) şöyle kaydedilmiştir:

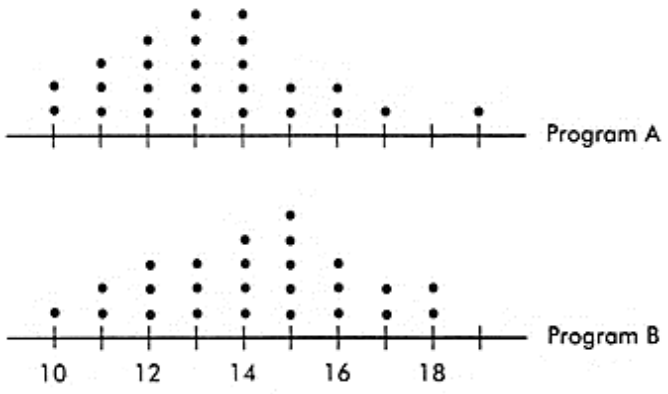
Denek	Zayıflama programı	Ağırlık kaybı (kg)	Denek	Zayıflama programı	Ağırlık kaybı (kg)
1	B	11	26	B	14
2	B	13	27	B	15
3	A	10	28	B	14
4	B	15	29	A	12
5	B	15	30	B	15
6	A	12	31	B	14
7	B	16	32	A	16
8	A	13	33	B	17
9	B	13	34	A	12
10	B	12	35	B	16
11	A	11	36	A	14
12	A	19	37	A	11
13	B	12	38	B	13
14	A	13	39	A	11
15	A	17	40	A	13
16	A	10	41	A	13
17	A	14	42	B	14
18	A	14	43	A	15
19	B	12	44	A	14
20	B	15	45	A	16
21	A	15	46	B	18
22	B	17	47	B	16
23	A	14	48	B	10
24	A	13	49	B	11
25	B	18	50	A	12

Uygulanan zayıflama programı ve ağırlık kaybı arasında bir ilişki var mı? Elimizdeki ham veriye bakarak, bu ilişkiyi görebilmek pek mümkün değil. Önce, bu ham veriyi taşıdığı enformasyonu daha kolayca görebileceğimiz bir formda yeniden organize edelim:

$$\text{A Programı: } 10, 10, 11, 11, 11, 12, 12, 12, 12, 13, 13, 13, 13, 13, 14, 14, 14, 14, 14, 15, 15, 16, 16, 17, 19$$

$$\text{B programı: } 10, 11, 11, 12, 12, 12, 13, 13, 13, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 17, 17, 18, 18$$

Veri bu formda organize edildiği zaman, iki değişken arasındaki ilişkiyi deneyimli gözler farkedebilir. Fakat, daha iyisi, bu ilişkiyi görsel olarak betimlemek üzere, her iki veri kümesinin dağılımlarını gösteren nokta grafikleri, aynı yatay ölçek üzerinde, üst üste çizerek karşılaştırmak olabilir.



Görülüyor ki A programı B programına göre daha düşük ortalama ağırlık kaybını çağrıştırıyor. Dağılımları şekil, merkez ve yayılım bakımlarından karşılaştırırsak:

Şekil: Görüyoruz ki her iki veri ikümesi de kabaca çan-şekillidir (ampirik kural geçerli) ve A programı 19 kilogramlık bir dışlağa sahip (B programında hiç bir dışlak yok).

Merkez: Görsel olarak ya da noktaları sayarak, dağılımların merkezlerini, A ve B programları için sırasıyla 13 kg ve 14 kg olarak bulabiliriz. (Bir hesap makinesi aritmetik ortalamaları $\bar{y}_{\mathrm{A}} = 13.36$ kg ve $\bar{y}_{\mathrm{B}} = 14.24$ kg olarak verecektir.) Hangi metotla bulursan bul, B programının merkezi daha yukarıdadır.

Yayılım: Her iki dağılımın yayılımları neredeyse aynı, A programı için 10 kg dan 19 kg a ve B programı için 10 kg dan 18 kg a. (Bir hesap makinesi standart sapmaları $s_{\mathrm{A}} = 2.18$ kg ve $s_{\mathrm{B}} = 2.18$ kg olarak verecektir.)

Sırt-sırta Dal-yaprak Grafikler

Misal 2 Kemoterapi uygulamasından sonra iki farklı ilacı kullanan kanser hastaları 40 yıl boyunca takip edilir ve sağkalım süreleri (yıl) kaydedilir. A ilacını kullanan 25 ve B ilacını kullanan 30 kişiye ait veri şöyledir:

$$\text{A ilacı: } 5, 10, 17, 39, 29, 25\ 20, 4, 8, 31, 21, 3, 12, 11, 19, 10, 4, 22, 17, 18, 13, 28, 11, 14, 21$$

$$\text{B ilacı: } 19, 12, 20, 28, 22, 35, 1, 21, 21, 26, 18, 28, 29, 20, 15, 32, 31, 24, 22, 26, 18, 20, 22, 35, 30, 18, 25, 24, 19, 21$$

Kemoterapiden sonra, kanser hastalarına uygulanan ilaç ve sağkalım süresi arasında bir ilişki var mı? Yine, bu ilişkiyi görsel olarak betimlemek üzere, her iki veri kümesinin dağılımlarını gösteren dal-yaprak grafikleri, aynı dikey ölçek üzerinde, sırt sırta çizerek karşılaştırabiliriz.

Elimizdeki verinin sırt-sırta bir dal-yaprak grafiğini çizmek üzere, dalların her iki yanına birer dikey çizgi çekiyoruz ve yaprakların bir kümesini sola doğru yerleştirirken, diğer kümesini sağa doğru yerleştiriyoruz.

İlaç B		İlaç A
1	0	3 4 4 5 8
9 9 8 8 8 5 2	1	0 0 1 1 2 3 4 7 7 8 9
9 8 8 6 6 5 4 4 2 2 2 1 1 1 0 0 0	2	0 1 1 2 5 8 9
5 5 2 1 0	3	1 9

(Dal birimi = 10)

Dikkat etmelisin ki A ilacı grubu en uzun sağkalımlı (39 yıl) hastaya ve B ilacı grubu en kısa sağkalımlı hastaya (1 yıl) sahipken, sırt-sırta dal-yaprak grafik gösteriyor ki B ilacını kullanan hastaların çoğunluğu A ilacını kullanan hastaların çoğunluğuna göre daha uzun süre sağkalmışlardır.

Dağılımın şeklini, merkezini ve yayılımını karşılaştırmak suretiyle, şunları söyleyebiliriz:

Şekil: Her iki dağılım da çan-şekillidir (ampirik kural geçerli). A ilacına ait dağılımın sağında 39 yıllık bir dışlağa sahip olduğu görülüyorken, B ilacına ait dağılımın solunda 1 yıllık bir dışlağa sahip olduğu görülüyor.

Merkez: Görsel olarak, ya da değerleri sayarak, dağılımların merkezlerini sırasıyla 17 yıl ve 22 yıl olarak bulabiliriz. (Bir hesap makinesi aritmetik ortalamaları $\bar{y}_{\mathrm{A}} = 16.48$ yıl ve $\bar{y}_{\mathrm{B}} = 22.73$ yıl olarak verecektir.) Hangi metotla bulursan bul, B ilacını kullanan hastaların sağkalım sürelerinin dağılımının merkezi, A ilacını kullanan hastaların sağkalım sürelerinin dağılımının merkezinden daha büyüktür.

Yayılım: Hastaların sağkalım sürelerinin dağılımlarının yayılımları, A ilacı için 3 yıldan 39 yıla ve B programı için 1yıldan 35 yıla. (Bir hesap makinesi standart sapmaları $s_{\mathrm{A}} = 9.25$ yıl ve $s_{\mathrm{B}} = 6.97$ yıl olarak verecektir.) Hangi metotla bulursan bul, A ilacını kullanan hastaların sağkalım sürelerinin dağılımının yayılımı, B ilacını kullanan hastaların sağkalım sürelerinin dağılımının yayılımından daha fazladır.

Paralel Kutu Grafikler

Misal 3 Organik tarım ürünleri son zamanlarda oldukça revaçta. Bir bilimsel-araştırmacı markette satılan endüstriyel tarım ürünü meyvelerin ve organik tarım ürünü meyvelerin fiyatlarını (TL/kg) karşılaştırıyor. Verinin bir betimleyici analizinin sonuçlarını ihtiva eden kompüter çıktısı şudur:

Endüstriyel tarım ürünü meyveler için:

```
Mean = 5.37   Standard deviation = 1.92   Min = 3.51
Max = 8.00   N = 18   Median = 4.77
Quartiles = 3.92, 6.45
```

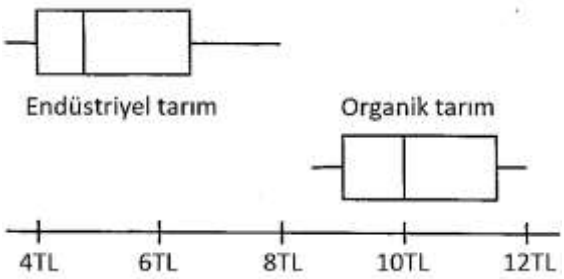
Organik tarım ürünü meyveler için:

```
Mean = 10.11   Standard deviation = 1.32   Min = 8.58
Max = 11.95   N = 15   Median = 10.08
Quartiles = 8.97, 11.51
```

Markette satılan meyvelerin üretim tipi ve fiyatı arasında bir ilişki var mı? Bu ilişkiyi hesaplanmış olan istatistiklere bakarak kolyca değerlendirebiliriz aslında(*Mean*: aritmetik ortalama, *Standard deviation*: standart sapma, *Min*: minimum, *Max*: maksimum, *N*: gözlem sayısı, *Median*: ortanca, *Quartiles*: çeyreklikler). Fakat dağılımların şeklini görmek, ilişkiye dair enformasyonu elbette artıracaktır.

Bu ilişkiyi görsel olarak betimlemek üzere, her iki veri kümesinin dağılımlarını gösteren kutu grafikleri, aynı ölçek üzerinde, üst üste veya yan yana çizerek karşılaştırabiliriz.

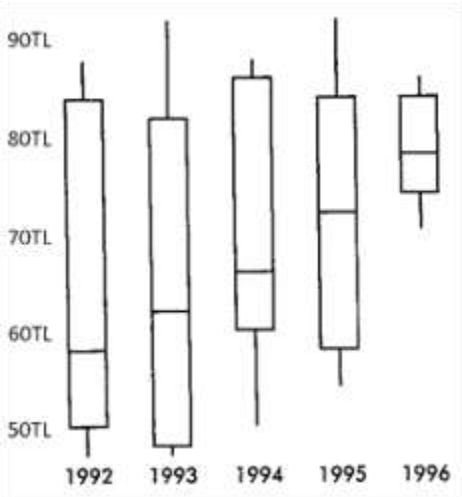
Elimizdeki verinin paralel kutu grafiklerini çizerken, her iki kutu grafiği, aynı ekseni kullanmak suretiyle bir diyagrama yerleştiriyoruz:



Kutu grafikler minimum, maksimum, ortanca ve çeyreklik değerlerini gösteriyor. Endüstriyel tarım ürünü meyvelerin fiyatlarının dağılımı daha aşağıda ve yayılımı organik tarım ürünü meyvelerin fiyatlarının dağılımından daha fazla. Her iki dağılımda yukarı uca (sağa) doğru hafif çarpık (çarpıklık, aynı zamanda, kompüter çıktısınında her iki dağılımın aritmetik ortalamalarının ortancalarından büyük olmasından da anlaşılabilir).

Paralel kutu grafikler, bilhassa, çok sayıda dağılımın karşılaştırılmasında yararlıdır.

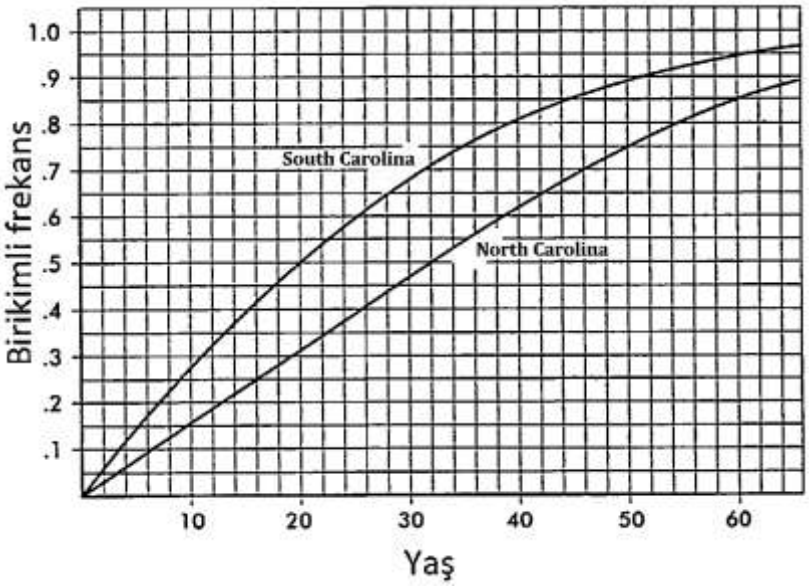
Misal 4 Müteakip paralel kutu grafikler, belirli bir şirketin günlük borsa değerinin yıl içerisindeki dalgalanmalarını biribirini takibeden beş farklı yıl için göstermektedir. Kutu grafikler bize hangi trendleri gösteriyor?



Paralel kutu grafikler gösteriyor ki şirketin ortanca günlük borsa değeri 58 TL den başlayarak 78 TL ye kadar yıldan yıla yükselerek 20 puan artmıştır. Üçüncü çeyrekliğin değeri 84 TL civarında neredeyse stabil kalmıştır. Yıllık minimum hiç bir yılda bir önceki yılın minimumundan düşük olmamıştır. Ve IQR hiç bir yılda bir önceki yıla göre artmamıştır ve hatta küçülen bir trendi olduğunu söyleyebiliriz.

Birikimli Frekans Grafikleri

Misal 5 Aşağıdaki birikimli frekans grafiği Amerika Birleşik Devletlerinin iki eyaletinin, South Carolina ve North Carolina, 1980 yılındaki nüfuslarının birikimli yaş dağılımlarını karşılaştırmaktadır.



Eyalet ve nüfusun yaşı arasında bir ilişki var mı? Bu soruya cevap vermek için, söz konusu dağılımların ortancalarını ve çeyreklikler arası gidimlerini nasıl karşılaştırırsın?

Cevap: Dikey eksende 0.5 noktasının karşısı boyunca baktığımız zaman, görüyoruz ki South Carolina'da nüfusun yarısı 20 yaşın altındayken, North Carolina'da bu değer 32 yaşa yükselmiştir. Dikey eksende 0.25 ve 0.75 noktalarının karşıları boyunca baktığımız zaman, görüyoruz ki South Carolina için $Q_1 = 9$ yıl ve $Q_3 = 35$ yıl ve dolayısıyla çeyreklikler arası gidim $35 - 9 = 26$ yılken, North Carolina için $Q_1 = 16$ yıl ve $Q_3 = 50$ yıl ve dolayısıyla çeyreklikler arası gidim $50 - 16 = 34$ yıldır.

İstatistiksel anlamlılık

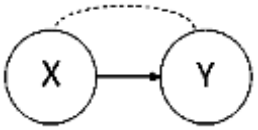
Peki, bir örnek veride gözlenen bir ilişki, bu örneğin içinden çekilmiş olduğu popülasyona genelleştirilebilir mi? Bu bir çıkarımsal istatistik problemidir ve bu dersin son bölümlerinde ele alınacaktır. Bu problemin çeşitli unsurları olmakla beraber, burada *istatistiksel anlamlılıktan* kısaca bahsetmek yerinde olacaktır. **İstatistiksel anlamlılık** iki değişken arasında gözlenen bir ilişkinin şans eseri olmamasının *olabilirliğine* matuftur. Anılan bölümlerde, bu olabilirliği görsel veya sayısal betimleyicilere yalnızca bakarak değerlendirmenin ötesine geçerek, olasılık modellerini kullanmak suretiyle matematiksel olarak değerlendirebileceğiz.

İlişki, Çağrışım ve Nedensellik

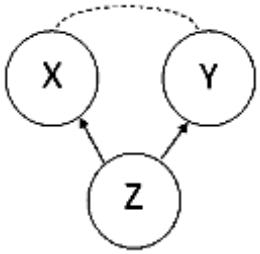
Her ne vakit iki değişken arasında bir ilişkiden (/*relationship*) bahsetsek, bir çağrışımı (/*association*) mı yoksa bir nedenselliği (/*causality*) mi kastettiğimiz açıklanmaya muhtaçtır.

İki değişken arasındaki ilişkiyi öğrenmeye-çalıştığımızda, çoğu zaman, belirli bir değişkendeki değişikliklerin başka bir belirli değişkende neden olduğu değişiklikleri göstermeyi ümit ederiz. İstatistiksel terminolojiye göre, birincisine *açıklayıcı değişken* ve ikincisine *yanıt değişkeni* adı verilir. Mamafih, iki değişken arasındaki güçlü bir çağrışım *neden-ve-etki* hakkında netice çıkarmak için yeterli değildir. Peki, iki değişken arasında hangi bağlantılar gözlenen bir çağrışımı açıklayabilir? Acaba, arkaplanda gizlenen başka değişkenler olabilir mi? Nedensellik için iyi bir kanıtı ne oluşturur? Çağrışımın 3 tipini mülahaza edeceğiz. Müteakip diyagramlarda, kesikli çizgi *X* ve *Y* değişkenleri arasında güçlü bir çağrışımı gösteriyor. Bu çağrışım oklarla gösterilen çeşitli nedensel ilişkilerin bir sonucu olabilir.

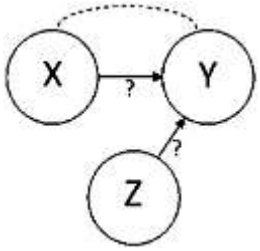
Nedensellik (Direk etki). İki değişken arasında doğrudan bir neden-ve-etki ilişkisi var. *X* teki değişiklikler *Y* de değişikliklere neden olur. Mesela, tüketilen alkol miktarı ve kandaki alkol düzeyi arasındaki ilişki.



Müşterek yanıt. Bir *gizli değişken* hem açıklayıcı değişkeni hem de yanıt değişkenini etkiler. Hem *X* teki hem de *Y* deki değişikliklere arka planda gizlenen bir *Z* değişkenindeki değişiklikler neden olur. Mesela, öğrencilerin üniversiteye giriş sınavında aldıkları puanları ve mezuniyet ortalamaları arasındaki ilişkiyi düşünelim. Zeki ve çalışkan öğrenciler hem üniversite giriş sınavında yüksek puan alma hem de yüksek mezuniyet ortalamasına sahip olma eğilimi gösterirler.



Karışma. İki değişkenin (ki çoğu zaman, bir gizli değişken ve bir açıklayıcı değişken) yanıt değişkenine etkisi biribirinden ayrılamaz. *X* in *Y* ye etkisi *Z* gizli değişkeninin etkisi ile karışır. bir öğrenme-çalışmasında, öğrenim düzeyi ve yıllık gelir arasındaki ilişkinin incelendiğini farzedelim. Meslek ve deneyim süresi karışan değişkenler olabilirler. Çünkü, bazı meslekler daha fazla kazandırmaktadır ve her meslekte deneyimli personele daha fazla ücret ödenmektedir. Meslek ve deneyim hesaba katılmadan, bu öğrenme-çalışması yanlış bir neticeye ulaşabilir.



- İki değişken arasındaki güçlü bir çağrışım her zaman bir değişkendeki değişikliklerin diğerinde değişikliklere neden olacağı manasına gelmez.
- İki değişken arasındaki çağrışıma çoğu zaman arka planda gizlenen başka değişkenler tesir eder.
- Nedensellik için en iyi kanıt *rastgeleleştirilmiş karşılaştırmalı deneylerden* elde edilir.
- İki değişken arasında gözlenen çağrışım; **direk ekiden**, **müşterek yanıttan** veya **karışmadan** dolayı olabilir.

Misal 6 Nedensellik iki değişken arasında bir çağrışımı gerektirir fakat çağrışımın nedensellik manasına gelmesi elzem değildir. Eğer bir öğrenme-çalışması yürütür ve gözlersek ki spor malzemelerine çok para harcayan kişilerde kalp krizi geçirme riski düşüktür, şu neticeye varamayız ki spor malzemesi satın almak kardiyovasküler hastalıklara karşı korur (nedensel ilişki). Belki spor malzemesi satın alanlar fizikselce aktif kişilerdir, öyle ki bu bilinen bir koruyucu faktördür. Burada fiziksel aktivite bir karışan değişkendir ve hem çok spor malzemesi satın almanın hem de düşük kalp krizi geçirme riskinin müşterek nedenidir.

Mini Proje 04

MP 01 de topladığın veri kümesinde yer alan bir nitel ve bir nicel değişken arasındaki ilişkiyi betimle. Sonuçların interpretasyonunu yap.

BÖLÜM 05 | İLİŞKİLERİ BETİMLEME: İKİ NİTEL DEĞİŞKEN

Giriş

Önceki bölümde, bir nicel değişkenin, başka bir nitel değişkenin kategorileri bakımından gösterdiği dağılımları biribiriyle karşılaştırmak suretiyle, aralarındaki ilişkiyi öğrenmeye-çalıştık. Ancak, pek çok istatistiksel öğrenme-çalışmasında olduğu üzere, aralarındaki ilişkiyi öğrenmeye-çalıştığımız değişkenlerin her ikisi de nitel birer değişken olabilmektedir.

Mesela yaş, gelir, öğrenim süresi, boy ve ağırlık gibi pek çok değişken doğası itibarıyla nicelken ya da sayısal değerler alırken; mesela cinsiyet, ırk, marka tercihi, meslek ve anksiyete durumu gibi başka değişkenler de niteldir ya da kategorik değerler alırlar. Bazen, doğası itibarıyla nicel olan değişkenler de, kullanılan ölçeğin tanımladığı yapay kategorik sınıflarda gruplanabilirler. Mesela insanların yaşını genç ve yaşlı olmak üzere iki kategoride ve ağırlığını zayıf, orta ve şişman olmak üzere üç kategoride ölçebilirsin.

Çapraz Tablo

Öğrenme-çalışmalarında elde edilen kategorik veri çoğu halde iki nitel değişken ihtiva eder ki bu değişkenler arasında bir ilişki mevcut olabilir veya olmayabilir. Böylesi veri bir *çapraz tabloda* gösterilebilir. Bir çapraz tablo iki nitel değişkenin ortak frekans dağılımı tablosudur. Bu tablolara *kontenjans tabloları* da denir.

Misal 1 Yetmiş yaşını geçmiş erkekler ile yapılan 4-yıllık bir öğrenme-çalışmasında, bireylerin kolesterol düzeyleri ölçüldü ve hangi tip kalp krizine maruz kaldıkları kaydedildi. Müteakip çapraz tablo bireylerin bu iki nitel değişkenin kategorilerine *ortak dağılımını* vermektedir. İki değişkenin ortak dağılımı, bu değişkenlerin kategorilerinin kombinasyonlarına ait ortak frekansların dağılımıdır. Mesela ölümcül-olmayan kalp krizi geçiren ve düşük kolesterol düzeyine sahip kişilerin ortak frekansı 29 dur.

		Kolesterol düzeyi		
		Düşük kolesterol	Vasat kolesterol	Yüksek kolesterol
Kalp krizinin tipi	Ölümcül-olmayan kalp krizi	29	17	18
	Ölümcül kalp krizi	19	20	9

Kalp krizinin şiddeti *satır değişkeniyken*, kolesterol düzeyi *sütun değişkenidir*. Bu veriyi analiz etmenin bir yolu, önce her satır ve her sütun için ortak frekansları toplamayı ihtiva eder:

		Kolesterol düzeyi			
		Düşük kolesterol	Vasat kolesterol	Yüksek kolesterol	Toplam
Kalp krizinin tipi	Ölümcül-olmayan kalp krizi	29	17	18	64
	Ölümcül kalp krizi	19	20	9	48
	Toplam	48	37	27	112

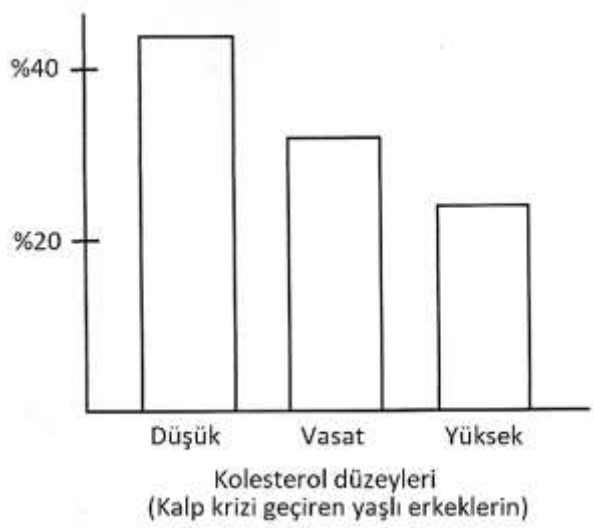
Satır ve sütun toplamları tablonun sağ ve alt marjinlerine (yani kenarlarına) yerleştirildiklerinden dolayı, bunlara *marjinal frekanslar* adı verilir. Bu marjinal frekanslar çoğu zaman orantılar veya yüzdeler formunda yazılır. Buna göre, her iki değişkenin kendi dağılımları bu tablodan hesaplanabilir. Kolesterol düzeyinin *marjinal frekans dağılımı* şöyledir:

Düşük: $$\frac{48}{112} = 0.429 = \%42.9$$

Vasat: $$\frac{37}{112} = 0.330 = \%33.0$$

Yüksek: $\frac{27}{112} = 0.241 = \%24.1$

Bu marjinal frekans dağılımını, aynı zamanda, bir çubuk grafik ile gösterebiliriz:

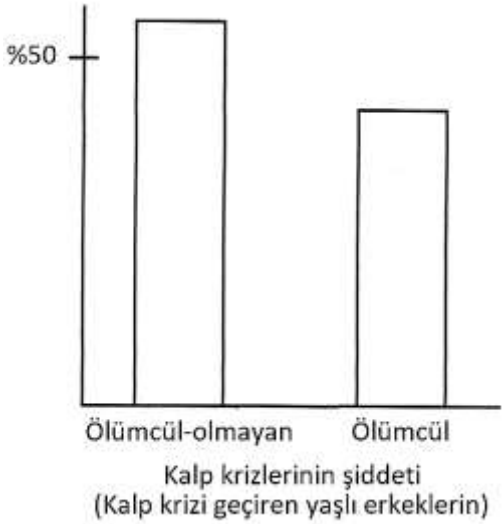


Benzerce, kalp krizinin şiddeti için de marjinal frekans dağılımını belirleyebiliriz:

Ölümcül-olmayan: $\frac{64}{112} = 0.571 = \%57.1$

Ölümcül: $\frac{48}{112} = 0.429 = \%42.9$

Bu marjinal frekans dağılımını temsil eden çubuk grafik:



Bir çapraz tablo her iki değişkenin ayrı ayrı marjinal frekans dağılımlarından daha fazla enformasyon taşır ve bu enformasyon değişkenler arasındaki ilişkiyi haizdir. Ancak, iki nitel değişken arasındaki ilişkiyi betimleyebilmek için ne değişkenlerin ortak frekans dağılımı ne de hesaplanan marjinal frekans dağılımları yeterli değildir.

Koşullu Nispi Frekanslar ve Çağrışım

Yukarıda gördük ki (Misal 1) bir çapraz tabloda betimlenen ortak dağılım ve hesaplanan marjinal dağılımlar iki nitel değişken arasındaki ilişkiyi betimleyebilmemize veya ölçebilmemize tek başına yetmedi. Bunu yapabilmek için, çapraz tablonun yalnızca marjinlerindeki (yani kenarlarındaki) toplam frekansları değil, gövdesindeki enformasyonu, başka bir perspektiten bakarak, mülahaza etmemiz gereklidir. İki nitel değişken arasındaki ilişki, çapraz tabloda yer alan ortak frekanslardan

mütenasip *koşullu nispi frekansların* hesaplanmasıyla betimlenir. Koşullu nispi frekanslar değişkenlerden herhangi birinin ayrı ayrı her bir kategorisi içerisinde diğer değişkenin kategorilerinin nasıl dağıldığını bize söyleyecek.

Şimdi, öyleyse, Misal 1 de yer alan çapraz tablo için koşullu nispi frekansları hesaplayalım. Zaten yukarıda, satır ve sütun toplamlarını bulmuştuk:

		Kolesterol düzeyi			
		Düşük kolesterol	Vasat kolesterol	Yüksek kolesterol	Toplam
Kalp krizinin tipi	Ölümcül-olmayan kalp krizi	29	17	18	64
	Ölümcül kalp krizi	19	20	9	48
	Toplam	48	37	27	112

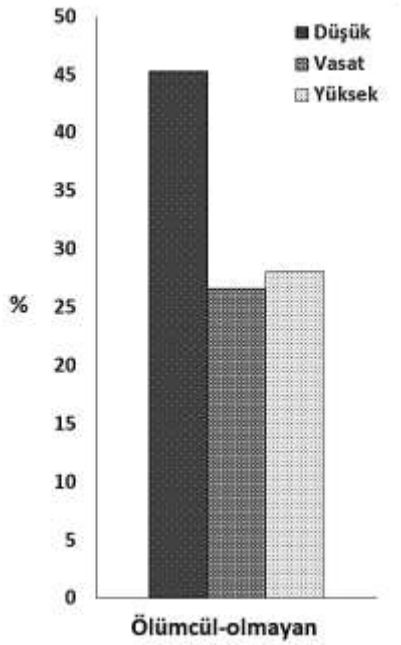
Kalp krizinin tipi ve kolesterol düzeyi arasında bir ilişkinin mevcut olup olmadığı ile ilgileniyoruz. Bunun için kalp krizi tipinin her iki kategorisi içerisinde kolesterol düzeyinin kategorilerinin nasıl dağıldığına, yani çapraz-tablonun her satırına ayrı ayrı bakacağız. Yani, ölümcül-olmayan kalp krizi geçirenler içerisinde her bir kolesterol düzeyinin orantısını veya yüzdesini hesaplayacağız.

Düşük: $\frac{29}{64} = 0.453 = \%45.3$

Vasat: $\frac{17}{64} = 0.266 = \%26.6$

Yüksek: $\frac{18}{64} = 0.281 = \%28.1$

Hesaplamış olduğumuz bu *koşullu nispi frekansları* bir bir çubuk grafikle gösterebiliriz:



Koşullu nispi frekansların bu dağılımına bakınca, ölümcül-olmayan kalp krizi geçiren hastalar içerisinde düşük kolesterol düzeyine sahip olanların nispi frekansının vasat ve yüksek kolesterol düzeyine sahip olanlara nazaran çok daha yüksek olduğunu görebiliyoruz.

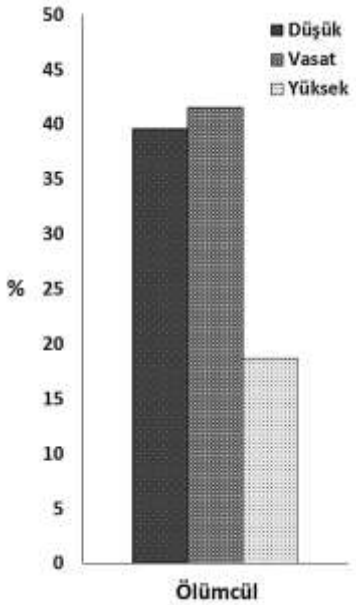
Benzerce, ölümcül kalp krizi geçiren 48 kişinin kolesterol düzeyine ait koşullu nispi frekansları şöyledir:

Düşük: $\frac{19}{48} = 0.396 = \%39.6$

Vasat: $\frac{20}{48} = 0.416 = \%41.7$

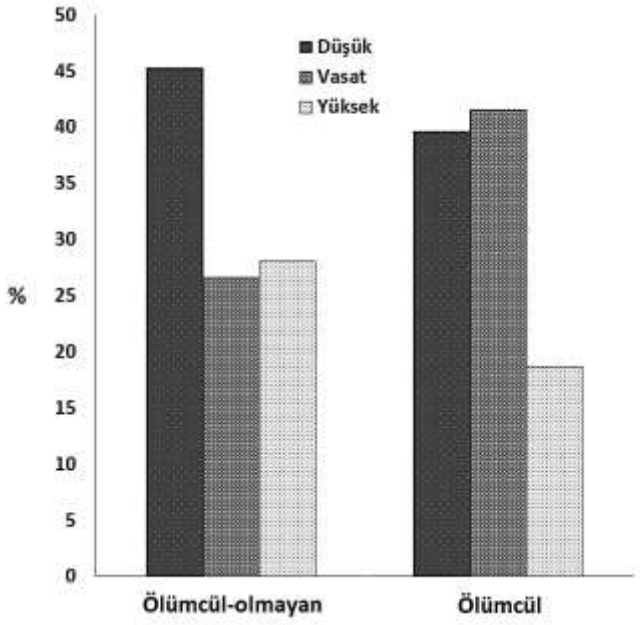
Yüksek: $\frac{9}{48} = 0.187 = \%18.8$

Bu koşullu nispi frekansları da yine çubuk grafikle gösterelim:



Koşullu nispi frekansların bu dağılımına bakınca, ölümcül-olmayan kalp krizi geçiren hastalarınkine pek benzemediğini görebiliyoruz. Ölümcül kalp krizi geçiren hastalar içerisinde koşullu nispi frekansı en yüksek olan kategori vasat kolesterol düzeyine sahip olanlarınki. Onu düşük kolesterol düzeyine sahip olanlarınki takip ediyor. Peki, iki değişken arasındaki ilişki hakkında bu bize ne söylemektedir?

Kalp krizi tipinin iki ayrı kategorisi için ayrı ayrı çizdiğimiz kolesterol düzeyinin koşullu nispi frekanslarının dağılımlarını daha kolay karşılaştırabilmek maksadıyla, şimdi, aynı grafikte gösterelim. Müteakip çubuk grafik iki değişken arasındaki ilişkinin güzel bir görsel betimlemesini vermektedir:



Yan-yana gruplandırılmış çubuk grafikler kolesterol düzeyi ve kalp krizinin tipi arasında bir ***çağrışıma*** işaret etmektedir. Çünkü ölümcül-olmayan kalp krizi geçiren kişiler içerisinde düşük kolesterol düzeyi daha yüksek bir nispi frekansa sahipken, ölümcül kalp krizi geçiren kişiler içerisinde vasat kolesterol düzeyi daha yüksek bir nispi frekansa sahiptir. Yani, kalp krizi tipi kategorilerine göre kolesterol düzeyinin dağılımları biribirilerine eşit görünmemektedir.

Misal 2 Saç dökülmesi tipi vücut kütle endeksiyle ilişkili midir? Yedi yüz altmış dokuz adamla yapılan bir öğrenme çalışması, aşağıdaki rakamları ortaya çıkardı:

		Saç dökülme tipi		
		Hiç	Ön	Tepe
Vücut kütle endeksi	< 25	137	22	40
	25-28	218	34	67
	>28	153	30	68

İcra edilecek analiz ilk evvela yukarıda yapmış olduğumuz üzere satır ve sütun toplamlarının bulunmasını ihtiva eder:

		Saç dökülme tipi			
		Hiç	Ön	Tepe	Toplam
Vücut kütle endeksi	< 25	137	22	40	199
	25-28	218	34	67	319
	>28	153	30	68	251
	Toplam	508	86	175	769

Saç dökülmesi tipini vücut kütle endeksinden tahmin etmekle ilgileniyoruz, ve böylece her satıra ayrı ayrı bakacağız. Mesela, vücut kütle endeksi 25 in altında olan 199 adamın her bir saç dökülmesi tipi için orantısı veya yüzdesi nedir?

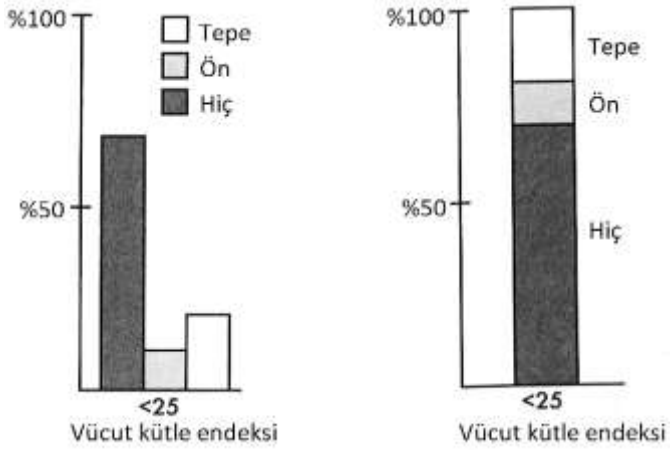
$$\text{Hiç:} \quad \frac{137}{199} = 0.688 = \%68.8$$

$$\text{Ön:} \quad \frac{22}{199} = 0.111 = \%11.1$$

$$\text{Tepe:} \quad \frac{40}{199} = 0.201 = \%20.1$$

Hesaplamış olduğumuz bu *koşullu nispi frekansları* grafikle gösterebilmek için ya çubukları yan-

yana ya da üst-üste gruplandırmamız gerekir. Her iki halde de, çubukların boyu nispi frekanslara tekabül edecektir:



Benzerce, vücut kütle endeksi 25 ve 28 arasında olan 319 adamın saç dökülmesi tipine ait koşullu nispi frekansları şöyledir:

$$\text{Hiç:} \quad \frac{218}{319} = 0.683 = \%68.3$$

$$\text{Ön:} \quad \frac{34}{319} = 0.107 = \%10.7$$

$$\text{Tepe:} \quad \frac{67}{319} = 0.210 = \%21.0$$

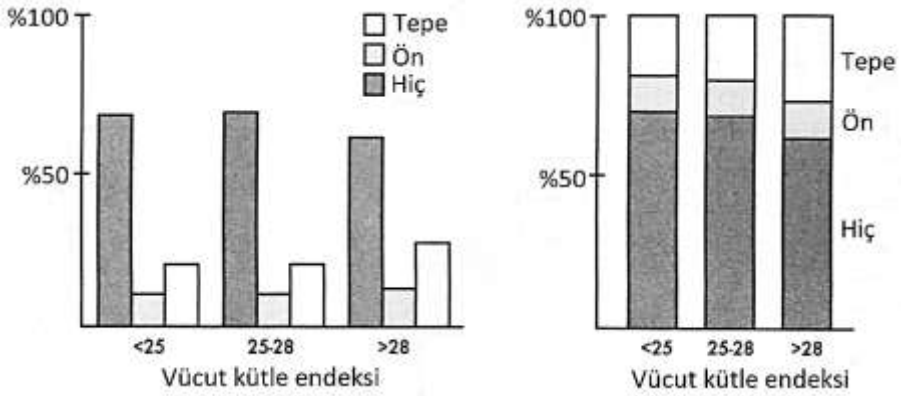
Vücut kütle endeksi 28 in üstünde olan 251 adamın koşullu nispi frekansları:

$$\text{Hiç:} \quad \frac{153}{251} = 0.610 = \%61.0$$

$$\text{Ön:} \quad \frac{30}{251} = 0.120 = \%12.0$$

$$\text{Tepe:} \quad \frac{68}{251} = 0.271 = \%27.1$$

Müteakip iki çubuk grafik iki değişken arasındaki ilişkinin güzel bir görsel betimlemesini vermektedir:



Hem yan-yana hem de üst-üste gruplandırılmış çubuklar saç dökülmesi tipi ve vücut kütle endeksi arasında zayıf bir ilişkiye işaret etmektedir. Çünkü vücut kütle endeksi 28 in üstünde olan adamlar için tepeden dökülme tipi biraz daha yüksek bir koşullu nispi frekansa sahiptir. Yani vücut kütle endeksi kategorilerine göre saç dökülmesi tipinin dağılımları biribirilerine eşit görünmemektedir.

Misal 3 Bir üniversite kampüsünde yapılan bir örnek yoklamasında öğrencilere hangi fakültede okudukları ve dondurmayı vanilyalı mı yoksa çikolatalı mı tercih ettikleri soruldu. Sonuçlar şöyle:

	Vanilyalı	Çikolatalı
Ziraat	20	10
Fen	24	12
Tıp	18	9
Edebiyat	22	11

Her bir fakülte için koşullu nispi frekanslar nedir?

Ziraat: $\frac{20}{30} = 0.667 = \%66.7$ vanilyalı ve $\frac{10}{30} = 0.333 = \%33.3$ çikolatalı tercih ediyor

Fen: $\frac{24}{36} = 0.667 = \%66.7$ vanilyalı ve $\frac{12}{36} = 0.333 = \%33.3$ çikolatalı tercih ediyor

Tıp: $\frac{18}{27} = 0.667 = \%66.7$ vanilyalı ve $\frac{9}{27} = 0.333 = \%33.3$ çikolatalı tercih ediyor

Edebiyat: $\frac{22}{33} = 0.667 = \%66.7$ vanilyalı ve $\frac{11}{33} = 0.333 = \%33.3$ çikolatalı tercih ediyor

Böylesi bir halde, ki tüm koşullu nispi frekans dağılımları özdeştir, iki değişken arasında *mükemmel bir çağrışımsızlık* olduğunu söyleriz. (Mamafih, şunu da belirtmek gerekiyor ki herhangi iki değişken tamamen biribiriyle çağrışımsız olsa dahi, gözlem sonuçlarını özetleyen bir çapraz tablonun bu mükemmel çağrışımsızlığı göstermesi şansı pratikte çok çok azdır.)

Gözlenen ve Beklenen Frekanslar

İki nitel değişken arasındaki ilişkiyi betimlemenin bir yolunun koşullu nispi frekansları (veya koşullu orantıları) karşılaştırmak olduğunu öğrendik. Başka bir yol ise gözlenen ve beklenen frekansları karşılaştırmaktır. Eğer iki değişken arasında herhangi bir ilişki yoksa, *beklenen frekanslar* örnekleme prosedürünün pek çok kere yinelenmesi sonucunda çapraz tablonun hücrelerinin ortalama frekanslarıdır. Bu durumda, örneğin çekildiği popülasyonda iki değişken biribirinden tam olarak bağımsızdır.

İlişkinin mevcut olmaması demek, bir başka deyişle, satır ve sütun kategorileri içerisinde frekansların her iki yönde benzer dağılması demektir. Bir çapraz tablonun beklenen frekanslarını hesaplamak için, elimizdeki örneğe ait veriyi kullanarak bir kestirimde bulunacağız. Her hücrenin satır ve sütun toplamlarını çarpıp, örnek hacmine böleceğiz. Yani,

$$\text{Beklenen frekans} = \frac{(\text{Satır toplamı})(\text{sütun toplamı})}{\text{Genel toplam}}$$

Gözlenen ve beklenen frekanslar arasındaki farklara bakarak iki nitel değişkenin çağrışımlı olup olmadığı hakkında bir gözlemde bulunabiliriz.

Misal 4 Bir öğrenme-çalışmasında, Amerika birleşik Devletleri'nde ve Kanada'da 2000 kişiye dini inancın bireysel ahlaka dahil olup olmadığı sorulmuştur. Toplanan veri müteakip çapraz tabloda özetlenmiştir.

	Evet	Hayır	Toplam
ABD	670	630	1500
Kanada	150	350	500
Toplam	1020	980	2000

Tabloya baktığımızda, "Evet" yanıtı verenlerin genel yüzdesinin $\frac{1020}{2000} = \%51$ olduğunu görüyoruz, fakat ABD ve Kanada için bu yüzdeler hayli farklı, sırasıyla $\frac{670}{1500} = \%58$ ve $\frac{150}{500} = \%30$. Bu çapraz tablo için beklenen frekansları şöyle hesaplarız:

	Evet	Hayır	Toplam
ABD	$\frac{(1500)(1020)}{2000} = 765$	$\frac{(1500)(980)}{2000} = 735$	1500
Kanada	$\frac{(500)(1020)}{2000} = 255$	$\frac{(500)(980)}{2000} = 245$	500
Toplam	1020	980	2000

Bu tablonun marjinal frekansları (yani, kenar toplamları) üstteki tabloyla aynı, fakat tablonun içindekiler, eğer 1500 Amerikalının ve 500 Kanadalının "Evet" yanıtlarının yüzdesi (%51) aynı olsaydı, bekleyeceğimiz frekansları yansıtmaktadır. Gözlenen frekanslar ve beklenen frekanslar arasında nasıl bir fark var? Eğer verilen yanıt milliyete bağlı olmasaydı, Amerakalıların 765 inin "Evet" demesi beklenirdi fakat 100 kişiden daha fazla fark var. Buna karşın, eğer bir çağrışım olmasaydı 255 Kanadalının "Evet" demesi beklenirdi fakat fiilen 150 Kanadalı "Evet" demiş. Tablonun diğer hücreleri de aynı minvalde farklara sahip. Bu görünen dört farka bakarak, formal olarak herhangi bir çıkarımı gerekçelendirmeksizin, söyleyebiliriz ki bu çapraz tabloda bir çağrışım gözlenmektedir.

İster koşullu nispi frekanslar vasıtasıyla olsun ve isterse beklenen değerler vasıtasıyla olsun, bir çapraz tabloda gözlediğin bir çağrışımın ancak ve ancak o tobloyu oluşturan fiili veri için geçerli olduğunu söyleyebilirsin. Eğer bir çapraz tablo örneklenmiş bir veriyi ihtiva ediyorsa, o çapraz tabloda gözlediğin bir çağrışımı, aynı değişkenlerin herhangi başka bir örneğinde de gözleyip gözleyemeyeceğini bilemezsin. Çıkarımda bulunabilmek için, bir anlamlılık testi yapman gerekir.

Karışan Değişkenler ve Simpson Paradoksu

Her ne zaman iki değişken arasında bir ilişkinin keşfine çıkılsa, neredeyse her zaman, değişkenlerden birinin diğerinde fiilen değişikliğe *neden olup olmadığı* sorusu gündeme gelir. Kadın olmak kontak lensi gözlüğe tercih etmeye neden olur mu? Botoks daha az terlemeye neden olur mu? Uygulanan tedavi iyileşmeye neden olur mu? Güneş ışınlarına maruz kalmak cilt kanserine neden olur mu? Fosfat vermek bitkinin yaprak sayısının artmasına neden olur mu? Eğer verinin nasıl toplandığını bilmiyorsak, hiç bir çapraz tablo iki değişken arasındaki çağrışımın bir *neden-ve-etki ilişkisi* olup olmadığını bize söyleyemez. İki değişken arasında gözlenen çağrışımın bir neden-ve-etki ilişkisi olup oladığı tamamen veri toplama metoduna bağlıdır. Eğer veri toplamak için *kontrol-altına-alınmış bir deney* dizayn etmiş ise, bir araştırmacı iki değişken arasındaki neden-ve-etki ilişkisini ancak o zaman açığa çıkarabilir. Aksi halde, sadece, bir *çağrışımsal ilişkiden* bahsedebilir. Veri toplama metotlarını tartışacağımız 7. Bölümde, gözlemsel öğrenme-çalışmalarında, karışan değişkenlerin olanaklı tesirinden dolayı, nedenselliği tesis etmenin zorluğunu vurgulayacağız.

Müteakip misalimiz *karışan değişkenlerin*, eğer onları hesaba katmaksızın (yani veri toplama prosesini kontrol-altına-almaksızın), arkaplanda gizlenmelerine izin verilirse, çapraz tabloların analizinde nasıl yanıltıcı nedensellik neticelerine yol açacaklarını tanıtlamaktadır.

Misal 5 Farzedelim ki –Tanrı korusun– bir kalp ameliyatı geçirmen gerekiyor ve mevcut iki cerrah arasında bir seçim yapmaya çalışıyorsun: Dr. Söğüt ve Dr. Kavak. Geçen yıl bu doktorların 250 şer hastayı ameliyat etmiş olduklarını ve şu sonuçları aldıklarını öğrendin:

	Dr. Söğüt	Dr. Kavak
Ölü	60	50
Sağ	190	200

Hangisine amaliyat olmak isterdin? Dr. Söğüt'ün 250 hastasından 190 ı sağ kalmış ki sağkalım oranı $\frac{190}{250} = 0.76$ veya %76. Dr. Kavak'ın 250 hastasından 200 ü sağ kalmış ki sağkalım oranı $\frac{200}{250} = 0.80$ veya %80.

Mamafih, her şey göründüğü kadar açık ve sarih olmayabilir. Farzedelim ki daha derin bir soruşturmayla öğreniyorsun ki doktorların ameliyat etmiş oldukları hastaların bazıları iyi bazıları ağır durumdaymış. Buna ait veri şöyle:

	İyi durumda	
	Dr. Söğüt	Dr. Kavak
Ölen	8	17
Sağkalan	60	120

	Ağır durumda	
	Dr. Söğüt	Dr. Kavak
Ölen	52	33
Sağkalan	130	80

Dikkat edersen, sonki iki tablodaki mütekabil hücreleri topladığımız zaman, daha yukarıdaki ilk tabloyu veriyor.

İyi durumdaki hastaları ameliyat ettikleri zaman cerrahları karşılaştıralım. Dr. Söğüt'ün iyi durumdaki 68 hastası $\frac{60}{68} = 0.882$ veya %88.2 lik sağkalım oranına sahipken, Dr. Kavak'ın iyi durumdaki 137 hastası $\frac{120}{137} = 0.876$ veya %87.6 lik sağkalım oranına sahip. Benzerce, görülüyor ki Dr. Söğüt'ün ağır durumdaki 182 hastası $\frac{130}{182} = 0.714$ veya %71.4 lük sağkalım oranına sahipken, Dr. Kavak'ın ağır durumdaki 113 hastası $\frac{80}{113} = 0.708$ veya %70.8 lik sağkalım oranına sahip.

Böylece, Dr. Söğüt hem iyi durumdaki hastalarda daha başaralı (Dr. Kavak'ın %87.6 lık sağkalım oranına karşı %88.2 ile) hem de ağır durumdaki hastalarda daha başarılı (Dr. Kavak'ın %70.8 lik sağkalım oranına karşı %71.4 ile). Bunula beraber, Dr. Söğüt'ün genel sağkalım oranı daha düşük (Dr. Kavak'ın %80 ine karşı %76 ile). Bu nasıl olabiliyor?

Bu problem *Simpson paradoksu* için bir misaldir. Birden çok sayıdaki gruplar, tek bir grup oluşturmak üzere katıştırıldığı zaman, eğer bir karşılaştırma ters yönde cereyan ediyor ise buna Simpson paradoksu denir. *Karışan değişken* adı verilen, bir başka değişkenin etkisi, gruplar katıştırıldığı zaman maskelenmiş olur. Şu anki misalimizde, daha derin bir inceleme gösteriyor ki Dr. Söğüt'ün ameliyat etmiş olduğu ağır durumdaki hastaların sayısı Dr. Kavak'tan çok daha fazladır ve ağır durumdaki bu hastaların elbette ki sağkalım oranı daha düşüktür. Böylece, Dr. Söğüt her iki hasta grubunda ayrı ayrı daha başarılı olsa bile, genel başarı oranı daha düşüktür. İlk tablomuz mevcut bir karışan değişkenin etkisini gizlemektedir ki bu karışan değişken hastaların durumuna ilişkindir.

Risk ve İhtimaller

Rastgele seçilecek bir bireyin bir nitel değişkenin belirli bir kategorisine düşme şansını ifade etmek için yalnızca iki temel yol vardır. Birincisi, bir kategoriyi toplamın *orantısı* olarak ifade etmek; ikincisi, bir kategoriyi diğerine nispetle *ihtimaller* (/*odds*) formunda karşılaştırmak.

Farzedelim ki bir popülasyon 1000 bireye sahip ve bunların 400 ü belirli bir hastalığın genini taşıyor. Aşağıdakilerin hepsi, bu orantıyı ifade etmenin eşit yollarıdır:

- Tüm bireylerin *yüzde* kırkı (%40) geni taşıyor.
- Geni taşıyanların *orantısı* 0.40 tır.
- Geni taşıma *riski* 0.40 tır.

Mamafih, bunu ihtimaller ile ifade etmek farklı bir hesaplama gerektirir. İhtimaller olarak verilmiş eşdeğer ifade şudur:

- Geni taşıma ihtimalleri 6 ya 4 tür (veya 3 e 2, veya 1 e $2/3$, veya ≈ 0.67)

İhtimaller ifade edilirken arzu edilen kategoriye düşenlerin ve düşmeyenlerin sayıları genellikle olanaklı en küçük tam sayılara indirgenir. Nitekim, ihtimalerin 1 e 2/3 olduğunu söylemek yerine 3 e 2 olduğunu söylemeyi tercih ederiz.

Formal tanımlarını vermek icabederse; **risk**, arzu edilen kategorinin orantısıdır; **ihtimaller**, arzu

edilmeyen kategoriye düşen birey sayısına karşı, arzu edilen kategoriye düşen birey sayısıdır (veya oranıdır).

Risk doğrudan arzu edilen kategorinin orantısına eşittir. İhtimalleri ise orantıdan kolayca hesaplayabiliriz. Eğer kategorinin orantısı p ise, rastgele seçilecek bir bireyin bu kategoriye düşme ihtimalleri 1 e $p/(1-p)$ dir. Tersinden bakarsak, eğer arzu edilen kategoriye düşme ihtimalleri b ye a ise, o zaman bu kategorinin orantısı $a/(a+b)$ dir. Mesela, belli bir geni taşıma orantısı 0.4 ise, o zaman rastgele seçilecek bir bireyin bu geni taşıma ihtimalleri 1 e $0.4/0.6$ veya 1 e $2/3$ veya 3 e 2 dir. Tersinden bakarsak, geni taşıma ihtimalleri 3 e 2 ise, o zaman geni taşıma orantısı $2/(2+3) = 2/5 = 4/10 = 0.40$ tır.

Nispi Risk

Yalnızca iki kategoride değer alabilen bir açıklayıcı değişkenin rastgele bir çıkımının nispi riski kategorilere ait risklerin biribirine oranıdır. Hakeza, *risk oranı* da denir. Nispi risk çoğu zaman bir çarpım olarak ifade edilir. Mesela, nispi riskin 3 olması, bir grubun bir hastalığı geliştirme riskinin diğer grubun aynı hastalığı geliştirme riskinin üç katı olduğunu söyler. Nispi risk 1 olduğu zaman, açıklayıcı değişkenin her iki kategorisi için risk aynıdır.

Misal 6 Pagano ve Gauvreau 1993 yılında yayınladıkları bir bilimsel-araştırma makalesinde katılımcı kadınlara ait yoklama verisini rapor etmişlerdir. Öğrenme-çalışmasında açıklayıcı değişken bir kadının ilk çocuğunu 25 yaşından sonra doğurup doğurmadığı ve yanıt değişkeni meme kanseri geliştirip geliştirmediğidir.

İlk çocuk 25 yaşından sonra mı?	Meme kanseri var	Meme kanseri yok	Toplam
Evet	31	1597	1628
Hayır	65	4475	4540
Toplam	96	6072	6168

Bir kadının ilk çocuğunu 25 yaşından sonra doğurup doğurmadığına göre meme kanseri geliştirme nispi riskini hesaplamak için, önce her grup için meme kanseri geliştirme riskini bulmalıyız:

- İlk çocuğunu 25 yaşından sonra doğuran kadınlar için risk $= 31/1628 = 0.019$
- İlk çocğunu 25 yaşından önce doğuran kadınlar için risk $= 65/4540 = 0.0143$
- Nispi risk $= 0.019/0.0143 = 1.33$

Diğer bir deyişle, ilk çocuğunu 25 yaşından sonra doğuran kadınlar için meme kanseri geliştirme riski 1.33 kat daha yüksektir.

Nispi riski hesaplarken, düşük riske sahip grubu paydaya koymak müşterek bir uygulamadır çünkü bu yöndeki sonuçların interpretasyonu pek çok insan tarafından daha kolay yapılabilir. Misal 6 için diğer yönden hesaplanan nispi risk 0.75 tir. Yani, ilk çocuğunu 25 yaşından önce doğuran kadınlar için meme kanseri geliştirme riski 0.75 katı yüksektir. Gördüğün gibi, bu nispi risk istatistiğini okumak daha zor olabilir.

Risk Artımı

Bazen, bir çarpım olarak ifade etmek yerine, riskteki artım yüzde olarak ifade edilebilir. Riskteki yüzde artım şöyle hesaplanır:

$$\text{risk artımı} = \frac{\text{riskteki değişme}}{\text{bazal risk}} \times 100$$

Risk artımını hesaplamanın eşdeğer bir yolu da şudur:

$$\text{risk artımı} = (\text{nispi risk} - 1) \times 100$$

Misal 7 İlk çocuğuğunu 25 yaşından önce doğuran kadınların meme kanseri riski ile

karşılaştırıldığında, ilk çocuğunu 25 yaşından sonra doğuran kadınların meme kanseri riskindeki değişme $(0.0190 - 0.0143) = 0.00047$ dir. İlk çocuğunu 25 yaşından önce doğuran kadınların riskini baz aldığımızdan ve bu risk 0.0143 olduğundan dolayı, bu değişme $(0.0047/0.0143) \times 100 = \%32.9$ luk ya da yaklaşık $\%33$ lük bir artımı ifade eder. Diğer bir deyişle, ilk çocuğunu 25 yaşından sonra doğurmak, kadınların meme kanseri şansını $\%33$ artırıyor. Dikkat ettiysen, bu aynı zamanda $(\text{nispi risk} - 1) \times 100$ dür.

İhtimaller Oranı

Hastalıkların ve diğer sağlık risklerinin nedenlerini ve progresyonunu öğrenmeye-çalışan epidemiyolojistler çoğu zaman riskleri karşılaştırmak için nispi risk yerine *ihtimaller oranı*nı (/*odds ratio*) kullanırlar. Eğer bir hastalığın riski düşükse, bu iki ölçü yaklaşık aynı olacaktır. Nispi riski anlamak daha kolaydır, fakat ihtimaller oranı istatistiksel çalışma için daha uygundur. Binaenaleyh, sağlıkla alakalı meseleleri kapsayan bilimsel dergilerdeki makalelerde çoğu zaman ihtimaller oranının rapor edildiğini göreceksin.

İhtimaller oranını hesaplamak için, önce açıklayıcı değişkenin iki kategorisinden her biri için hastalığa yakalanma ve yakalanmama ihtimallerini hesaplarız. Sonra bu ihtimalleri birbirine oranlarız. Şimdi Misal 6 daki meme kanseri ile ilgili veri için ihtimaller oranını hesaplayalım:

- İlk çocuğunu 25 yaşından sonra doğuran kadınlar için ihtimaller 1 e $31/1597 = 0.0194$.
- İlk çocuğunu 25 yaşından önce doğuran kadınlar için ihtimaller 1 e $65/4475 = 0.0145$.
- İhtimaller oranı $0.0194/0.0145 = 1.34$.

İhtimaller oranının 1.34 olarak hesaplanan değerinin, nispi riskin 1.33 olarak hesaplanan değerine çok benzer çıktığını görebilirsin. Evvelce belirtildiği üzere, hastalık riski her bir kategoride düşük olduğu müddetçe durum bu minvaldedir.

İhtimaller oranını hesaplamanın daha kolay bir yolu var, fakat eğer sadece formülü görmüş olsaydın, ihtimaller oranının iki ihtimallerin oranı olduğunu anlayamayabilirdin. Bir çapraz tablodan ihtimaller oranını hemencecik hesaplamak için:

1. Tablonun üst sol ve alt sağ hücrelerindeki sayıları çarp.
2. Tablonun üst sağ ve alt sol hücrelerindeki sayıları çarp.
3. 1. sonucu 2. sonuca böl.

Misal 6 daki çapraz tablodan ihtimaller oranını şöyle hesaplarız:

$$\text{ihtimaller oranı} = \frac{31 \times 4475}{1597 \times 65} = 1.34$$

Tablonu nasıl yapılandırıldığına bağlı olarak, pay ve paydanın yerini değiştirmen gerekebilir. Nispi riski hesaplarken öğrendiğin gibi, ihtimaller oranını 1 den büyük olacak şekliyle yapılandırmak gelenekselleşmiştir. Tek fark, açıklayıcı değişkenin hangi kategorisinin oranın payı ve hangisinin paydası olarak kabul edileceğidir.

Mini Proje 05

MP 01 de topladığın veri kümesinde yer alan iki nitel değişken arasındaki ilişkiyi betimle. Sonuçların interpretasyonunu yap.

BÖLÜM 06 | İLİŞKİLERİ BETİMLEME: İKİ NİCEL DEĞİŞKEN

Giriş

Aralarındaki ilişkiyi öğrenmeye-çalıştığımız değişkenlerin her ikisi de pek çok zaman nicel değişkenler olabilmektedir. Mesela, kadınların gebelikte içtikleri tütün miktarı ve bebeklerinin doğum ağırlıkları arasında bir ilişki var mı? Et fiyatı ve kişi başına et tüketimi arasında bir ilişki var mı? Bitkiye verilen gübre miktarı ve alınan ürün miktarı arasında bir ilişki var mı?

Her iki değişkenin de nicel değişkenler olması, sayısal değerler almaları demektir. Sayısal değerlerin mevcudiyeti ise matematiksel araçların daha doğrudan kullanılması manasına gelmektedir. Bu halde, iki soru hemen kendini gösterir. İlki, mevcudiyetini gördüğümüz bir ilişkinin gücünü ve yönünü nasıl ölçebiliriz? İkincisi, Gözlenen bir ilişkiyi bir fonksiyon ile nasıl ifade edebiliriz? Mesela, bir emlak borsacısı, sadece faiz oranı ve satılan yeni konutların sayısı arasında bir ilişki olup olmadığını belirlemeyi değil, aynı zamanda belirli bir faiz oranı verildiği zaman, satılan konut sayısını tahmin etmekte kullanabileceği bir matematiksel ifadeye sahip olmayı da arzu eder.

İki nicel değişken arasındaki olanaklı bir ilişkinin görsel bir betimlemesini bir *saçılım grafiği* ile yapabiliriz. Daha sonra, *korelasyon katsayısı* adını verdiğimiz bir ölçü ile doğrusal bir ilişkinin gücünü ve yönünü sayısal olarak betimleyebiliriz. Her iki halde de, bir ilişkinin varlığının kanıtı, nedenselliğin de kanıtı değildir.

İki değişken arasında mevcudiyetini gözlediğimiz bir ilişkinin biçimini en küçük kareler regresyon doğrusu ile nasıl betimleyebileceğimizi öğrenmeye-çalışacağız.

Bu bölümü herhangi bir değişken ve zaman arasındaki ilişkiyi (veya bir fenomenin zamana göre değişimini) biraz daha tafsilatlı inceleyerek bitireceğiz.

Saçılım Grafiği

X ve Y ile temsil ettiğimiz iki nicel değişken arasındaki ilişkiyi, bu iki değişene ait (x, y) değer çiftlerinin bir grafiği ile betimleyebiliriz. İlişkinin gücü ve yönü, ilişkinin biçimi ve ilişkinin biçiminden sapmalarla ilgileniriz. Bu bölümde iki nicel değişken arasındaki ilişkinin bir doğrusal fonksiyon, yani düzlemde bir doğru ile nasıl açıklanabileceğini inceleyeceğiz.

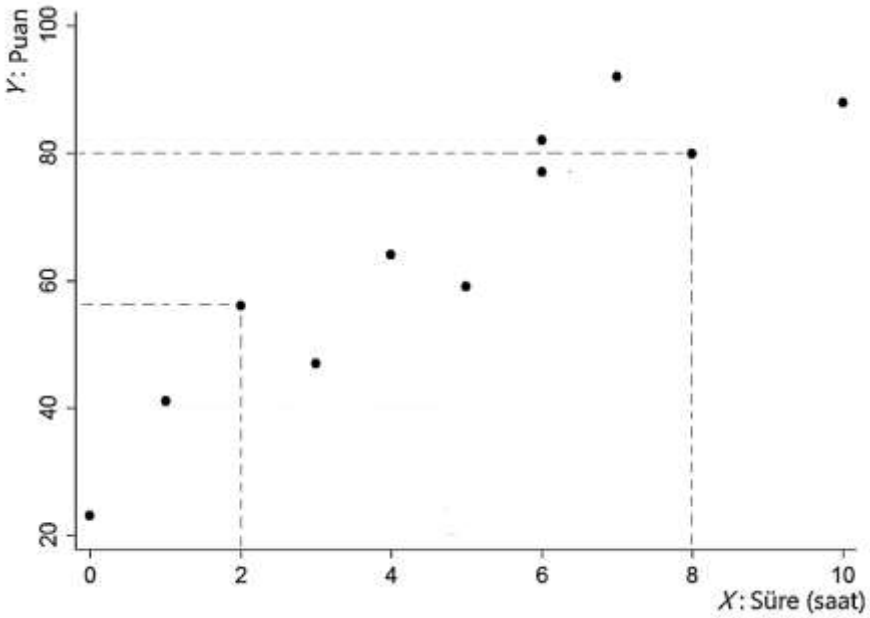
Bir istatistiksel öğrenme-çalışması bağlamında, genellikle, aralarındaki ilişkiyi incelediğimiz değişkenlerden birini *yanıt değişkeni* (Y) ve diğerini de *açıklayıcı değişken* (X) olarak kabul ederiz. Bununla beraber, yanıt değişkenine *bağımlı değişken* ve açıklayıcı değişkene *bağımsız değişken* de denilmektedir. Fakat,aklımızdan çıkarmamalıyız ki değişkenlerin bu şekilde adlandırılmaları her zaman bir neden-ve-etki ilişkisini kastetmek zorunda değildir. Bu bağlamda gözlem üniteleri üzerinde (x, y) değer çiftleri olarak üretilen veri genel olarak *iki-değişkeli* veri olarak adlandırılır.

Elimizde belli bir dersi alan öğrencilerden 11 inin verilen bir ödevi yapmak için harcadıkları süreleri ve yaptıkları ödevlerden aldıkları puanları gösteren aşağıdaki veri kümesi mevcut. Öğrencilerin ödev için harcadıkları süre ve aldıkları puan arasındaki ilişkiyi öğrenmeye-çalışıyoruz.

Öğrenci:	i:	1	2	3	4	5	6	7	8	9	10	11
Süre (saat):	x:	6	10	2	4	6	7	0	1	8	5	3
Puan:	y:	82	88	56	64	77	92	23	41	80	59	47

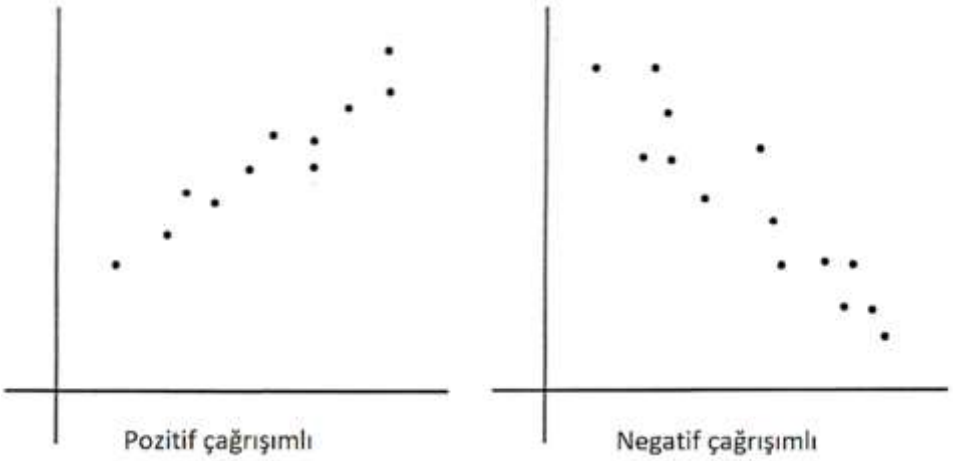
Aşağıdaki saçılım grafiğinde, X ve Y değişkenlerine ait (x_i, y_i) değer çiftlerinin koordinat düzleminde noktalar halinde saçılımını görüyorsun. Mesela üçüncü öğrenciye ait değer çiftinin $(2, 56)$ ve dokuzuncu öğrenciye ait değer çiftinin $(8, 80)$ düzlemdeki yerleri kesikli çizgilerle belirtilmiştir.

İki değişken arasında bir ilişki var mı? Varsa, nasıl bir ilişki bu? Pozitif mi, negatif mi? Güçlü mü, zayıf mı?

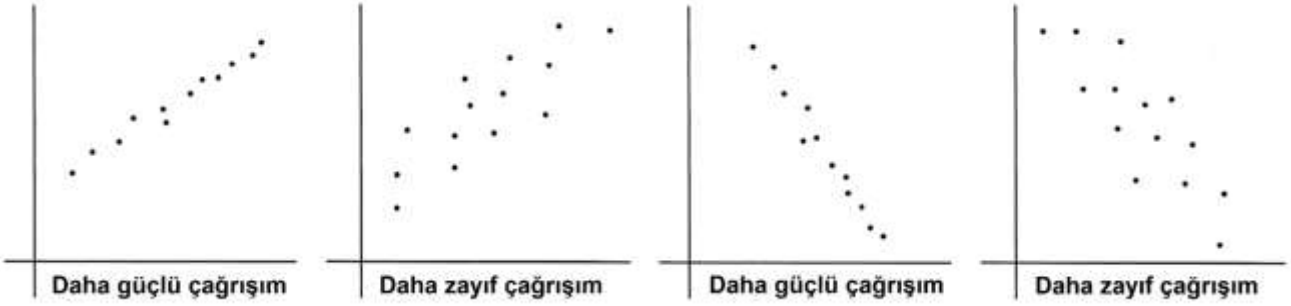


Saçılım grafiğinde biçimler

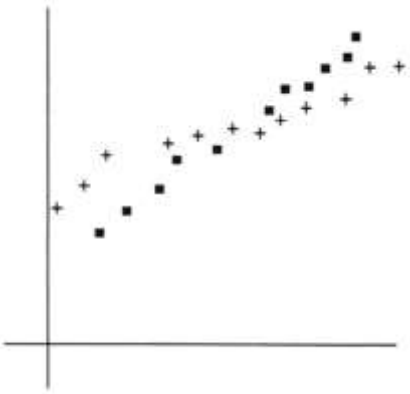
Bir değişkenin daha büyük değerleri, ikinci bir değişkenin daha büyük değerleri ile **çağrışımlı** ise, bu iki değişkene **pozitif çağrışımlı** denir. Birinin daha büyük değerleri, ikincinin daha küçük değerleri ile çağrışımlı ise, bu değikenlere **negatif çağrışımlı** denir.



Çağrışımın gücü, düzlemdeki noktaların bir doğru oluşturacak şekilde ne kadar sıkıca bir araya geldiğiyle değerlendirilir. Noktalar düzlemde bir doğru oluşturmaya ne kadar yakınsa, iki değişken o kadar **güçlü çağrışımlıdır**.

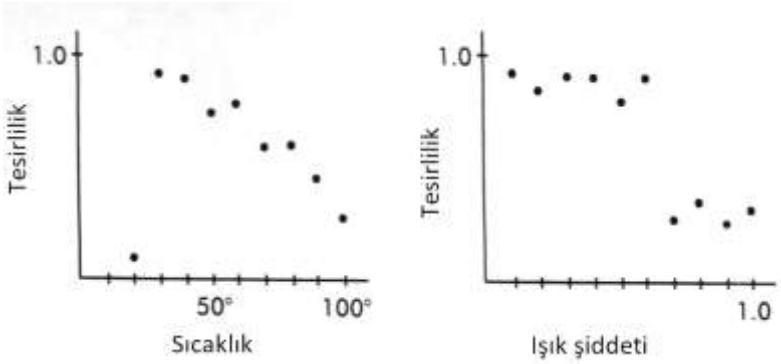


Bazen, saçılım grafiğindeki noktaları, başka bir nitel değişkenin farklı kategorilerine düştüklerini belirtmek üzere farklı renklerle veya sembollerle gösterebiliriz. Mesela, aşağıdaki saçılım grafiğinde kadınlar birer artı işareti ve adamlar birer kare ile etiketlenmek suretiyle biribirlerinden ayrılmışlardır.



Bir saçılım grafiğinin genel biçimini analiz ederken, aynı zamanda *öbeklerin* ve *dışlakların* mevcudiyetini de kontrol etmek ve bunları nazar-ı dikkate almak önemlidir.

Misal 1 Belirli bir antibiyotiğin tesirliliğine, sıcaklığın ve ışığın etkisini test etmek için, bir deney yürütülür. Antibiyotik viyallerinin bir kümesi aynı ışıklandırmayla farklı sıcaklıklarda depolanırken, ikinci bir kümesi de aynı sıcaklıkta farklı ışıklandırmalarla depolanır.



Birinci saçılım grafiğine baktığımızda, bir doğrusal biçim görüyoruz ve bu biçimin hayli uzağında bir dışlak mevcut. Bunun olanaklı bir açıklaması, antibiyotiğin daha düşük sıcaklıklarda daha tesirli olduğu, fakat belirli bir sıcaklığın altında tesirini kaybettiği olabilir.

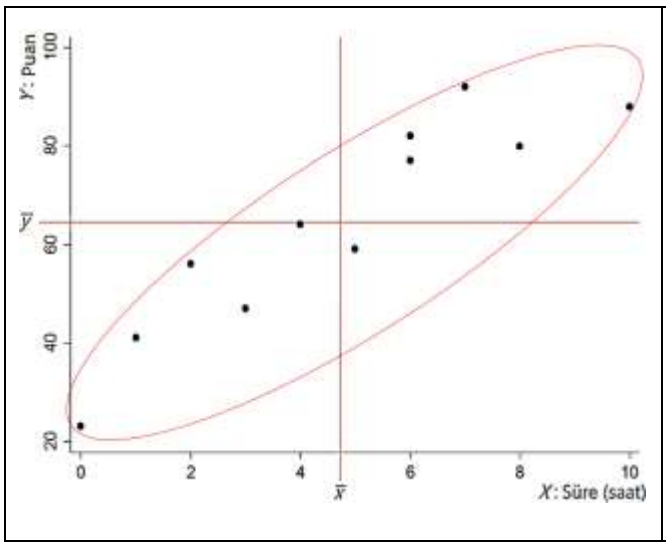
İkinci saçılım grafiğine baktığımızda, iki öbek görüyoruz. Bunun manası açıkça şu ki belli bir ışık şiddetinin altında antibiyotik bir tesir gösterirken, o ışık şiddetinin üstünde başka bir tesir gösteriyor. Her iki öbekte de ışık şiddeti ve tesirlilik arasında bir çağrışım görülmüyor.

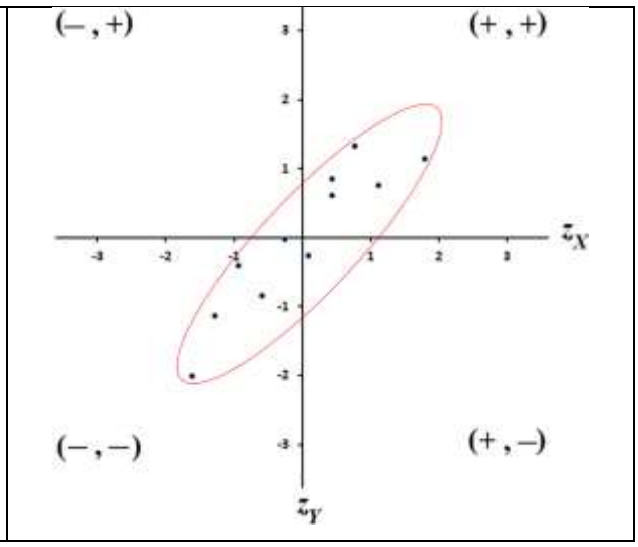
Korelasyon ve Doğrusallık

Her ne kadar bir saçılım grafiği bize iki nicel değişken arasındaki ilişkinin yönünü ve gücünü görsel olarak belirtse bile, pek çok durumda, bir ilişkinin spesifik gücünü görsel olarak değerlendirebilmek hayli zordur. Bu sebeple, iki nicel değişken arasındaki bir doğrusal ilişkiyi betimlemek için, **korelasyon katsayısı** (veya kısaca **korelasyon**) adı verilen bir matematiksel ölçü kullanırız. Korelasyon, bir rastgele örneklemeyle elde edilen iki değişkene ait verinin aritmetik ortalamaları ve standart sapmaları cinsinden şöyle ifade edilir:

$$r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

Korelasyon ifadesine dikkatle bakarsan, görürsün ki aslında (x_i, y_i) değer çiftlerinin mütekabil *z-puan*larının çarpımlarının aritmetik ortalamasıdır. Örnekten hesaplanan bir korelasyon katsayısı (ki r ile temsil ederiz) örneğin çekildiği popülasyona ait korelasyon katsayısının bir kestirimidir. Popülasyona ait korelasyon katsayısını ρ (Yunan alfabesindeki küçük "rho" harfi) ile temsil ederiz. Popülasyona ait veri için korelasyonu hesaplarken, s_x ve s_y yerine σ_x ve σ_y kullanılacağı için, $n-1$ yerine de N kullanmak lazımdır.





Verilen ödevi yapmak için öğrencilerin harcadıkları sürelerin aritmetik ortalaması 4.6 saat ve standart sapması 3.2 saat, yaptıkları ödevlerden aldıkları puanların aritmetik ortalaması 62.7 ve standart sapması 21.7 dir. Buna göre, öğrencilerin ödev yapmak için harcadıkları süre ve aldıkları puan arasındaki ilişkinin bir sayısal betimleyicisi olarak korelasyon katsayısını şöyle hesaplarız.

$$r = \frac{1}{11-1}\left[\left(\frac{6-4.6}{3.2}\right)\left(\frac{82-62.7}{21.7}\right) + \left(\frac{10-4.6}{3.2}\right)\left(\frac{88-62.7}{21.7}\right) + \left(\frac{2-4.6}{3.2}\right)\left(\frac{56-62.7}{21.7}\right)\right.$$
$$+ \left(\frac{4-4.6}{3.2}\right)\left(\frac{64-62.7}{21.7}\right) + \left(\frac{6-4.6}{3.2}\right)\left(\frac{77-62.7}{21.7}\right) + \left(\frac{7-4.6}{3.2}\right)\left(\frac{92-62.7}{21.7}\right)$$
$$+ \left(\frac{0-4.6}{3.2}\right)\left(\frac{23-62.7}{21.7}\right) + \left(\frac{1-4.6}{3.2}\right)\left(\frac{41-62.7}{21.7}\right) + \left(\frac{8-4.6}{3.2}\right)\left(\frac{80-62.7}{21.7}\right)$$
$$\left. + \left(\frac{5-4.6}{3.2}\right)\left(\frac{59-62.7}{21.7}\right) + \left(\frac{3-4.6}{3.2}\right)\left(\frac{47-62.7}{21.7}\right)\right]$$
$$= \frac{1}{10}[0.34 + 1.87 + 0.35 + 0.24 + 0.94 + 2.951.310.77 - 0.02 + 0.45] = 0.92$$

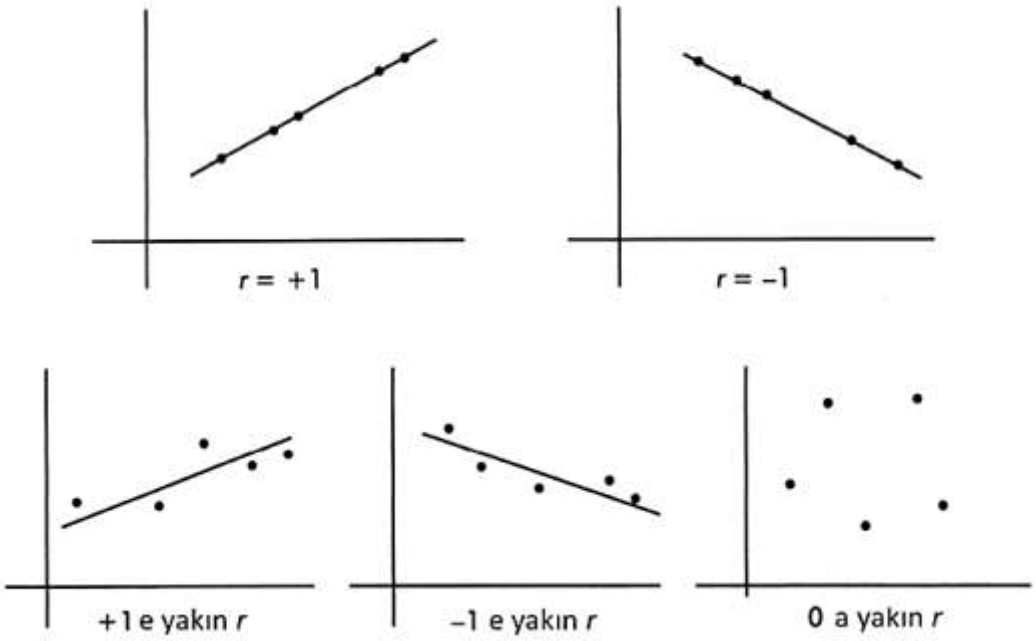
Mamafih, korelasyon günümüzde hesap makineleri ve kompüter yazılımları ile kolayca hesaplanabilmektedir. Burada korelasyonun matematiksel ifadesini, sadece korelasyon fikrini daha iyi anlayabilmeniz bakımından inceliyoruz. Yoksa, pratikte her seferinde r yi elde hesaplamak için vaktimizi harcamayız.

Korelasyonun formülüne bakınca, hangi değişkenin X olarak, hangi değişkenin Y olarak adlandırılacağının bir ayrımını göremiyoruz. Aynı zamanda, formül z-puanlarına dayanmaktadır ve ölçüm birimini değiştirmek, korelasyonu değiştirmeyecektir. Nihayetinde, dışlakların aritmetik ortalamalar ve standart sapmalar üzerindeki güçlü tesirinden dolayı, korelasyonun da aşırı değerlerden hayli etkilenebileceğini söylememiz gerekir.

Korelasyonun değeri

Daima, r nin değeri –1 ve +1 arasına düşer. Mükemmel negatif korelasyon olduğu zaman, $r = -1$ ve mükemmel pozitif korelasyon olduğu zaman, $r = +1$. Eğer korelasyon sıfır çıkarsa, bunun manası iki değişken arasında bir doğrusal ilişkinin mevcut olmadığıdır, yoksa bir ilişkinin mevcut olmadığı değil. Mamafih, bir doğrusal ilişkinin mevcut olmadığı durumda, iki değişken arasında pekâlâ güçlü bir *doğrusal-olmayan* ilişki mevcut olabilir. Korelasyon katsayısı, r'nin 0'a yakın değerleri çok zayıf bir doğrusal ilişkiyi belirtir. Korelasyon katsayısı, r'nin değeri 0'dan uzaklaşarak -1'e veya $+1$'e doğru yaklaştıkça doğrusal ilişkinin gücü artar.

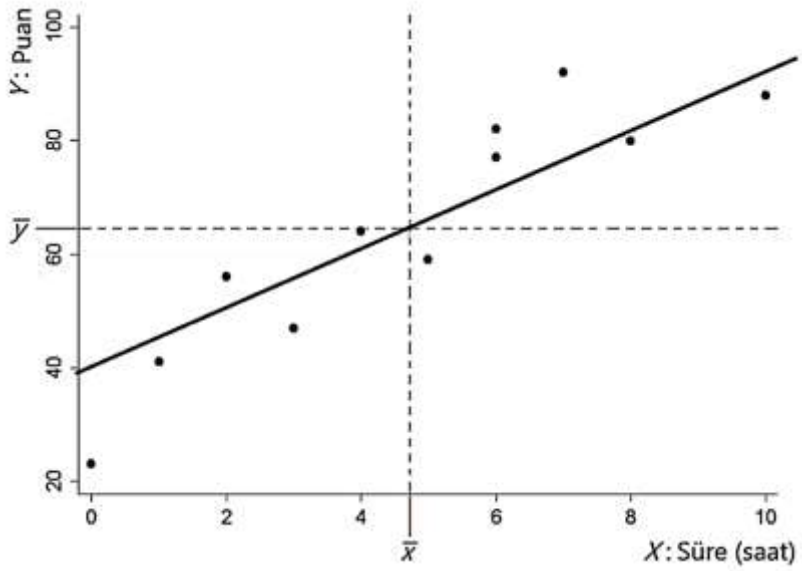
Korelasyon nedensellik demek değildir!



Eğer iki değişken arasında nedensel bir ilişki mevcut değilse, ya tesadüfen ya da belli bir üçüncü görünmeyen faktörün (müşterek yanıt değişkeni veya gizli değişken) valığından dolayı bu iki değişkenin çağrışımlı oldukları neticesine varılabilir. Bu çağrışıma *sahte ilişki* veya *sahte korelasyon* adı verilir.

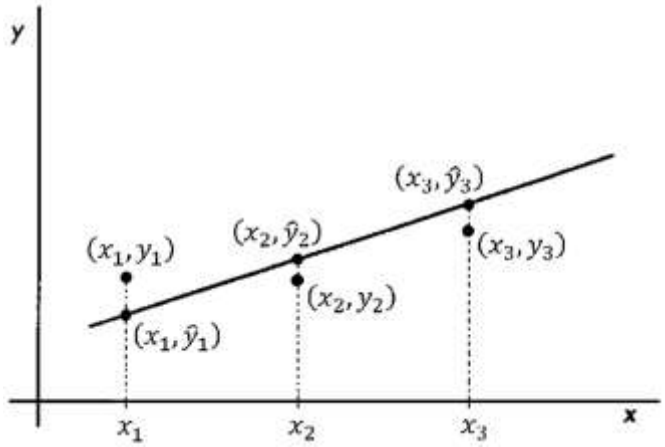
En Küçük Kareler Regresyon Doğrusu

Koordinat düzlemine saçılan noktalar boyunca çizilebilecek en iyi uyum doğrusu hangisidir? Çizilmesi olanaklı pek çok doğrudan (yani düz çizgiden) öyle birini seçmeliyiz ki iki değişken arasındaki ilişkinin en iyi betimleyicisi olabilsin. Aşağıdaki saçılım grafiğinde, X ve Y değişkenlerine ait (x_i, y_i) değer çiftlerinin koordinat düzleminde noktalar halinde saçılımını ve bu noktalar arasından geçirilen *en iyi uyum doğrusunu* görüyorsun.



Değişkenler arasındaki ilişkinin yeterince güçlü olup olmadığını değerlendirebilmek ve eğer yeterince güçlü bir ilişki var ise, bu ilişkiyi betimleyen doğrunun hangisi olduğunu bulmak için "en iyi uyum doğrusu" teriminin ne manaya geldiğini bilmemiz gerekir.

Bir değişkenin varyansını hesaplarken kullandığımız fikri anımsamanızı istiyorum, aritmetik ortalamadan farkların karelerini toplamıştık. En iyi uyum doğrusu ile, benzer bir fikri kastediyoruz. Gözlenen değerler ve doğru tarafından tahmin edilen değerler arasındaki dikey farkların karelerinin toplamını en küçük yapan doğruya, *en iyi uyum doğrusu* diyoruz.

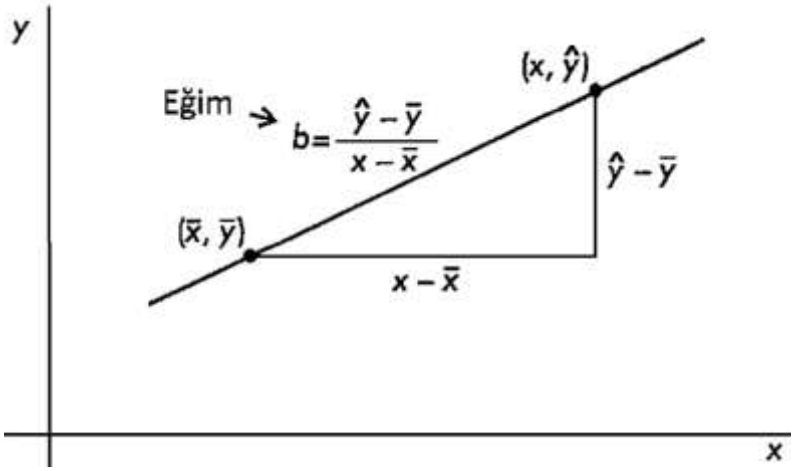


Bir değişkeni temsil eden harfin tepesindeki şapka, o değişkenin bir **tahmini** manasına gelir.

Yukarıdaki figürde, X ve Y değişkenlerinin üç örnekleme ünitesi ($n = 3$) üzerinde gözlenen (x_i, y_i) değer çiftlerinin düzlemdeki saçılımını ve bu noktalar boyunca çizilen bir doğru vasıtasıyla her x_i gözlenen değeri için tahmin edilen $\hat{y}_i$ değerini görüyoruz. En küçük yapmak istediğimiz farkların kareleri toplamını şöyle yazabiliriz:

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + (y_3 - \hat{y}_3)^2$$

Örnek hacmi genişledikçe, yani n büyüdükçe, düzleme saçılan noktaların sayısının ve kareli farkların sayısının da aynı oranda artacağı aşikar. Mantıklı, sezgisel ve doğru ki en iyi uyum doğrusu her halükarda $(\bar{x}, \bar{y})$ noktasından geçecektir. Yani, X ve Y değişkenlerinin aritmetik ortalamalarının düzlemde kesiştiği noktadan. O zaman, belli bir noktadan geçen belli bir eğime (b) sahip bir doğrunun matematiksel ifadesine dayanarak,



$$\hat{y} - \bar{y} = b(x - \bar{x})$$

veya

$$\hat{y} = \bar{y} + b(x - \bar{x})$$

ve kesim noktası

$$a = \bar{y} - b\bar{x}$$

olmak üzere,

$$\hat{y} = a + bx$$

eşitliklerini yazabiliriz. Sonki eşitlik regresyon doğrusunu ifade etmek için çok daha müştereken kullanılanıdır.

Matematiksel detaylarına girmeden, iki-değişkeli bir örnek veri için, en küçük kareler regresyon doğrusunun eğimini, korelasyon katsayısı (r) ve her iki değişkenin standart sapmaları (s_y ve s_x) cinsinden şöyle hesaplarız:

$$b = r\frac{s_y}{s_x}$$

Yani, X teki her bir standart sapmalık değişime, Y de r standart sapmalık bir değişim karşılık gelir. Eğer Y değişkenine ait *z-puan*larının, X değişkenine ait *z-puan*larına karşı garafiğini çizersek, regresyon doğrusunun eğimi tam olarak r ye eşittir. Bu durumda, doğrusal denklem,

$$\hat{z}_y = rz_x$$

halini alır. *Z-puan*larına ait regresyon doğrusu daima orijinden geçeceği için kesim noktası da daima sıfırdır.

Standartlaştırılmış veri (yani *z*-puanları) ile hesaplandığında, korelasyon katsayısı (r) regresyon doğrusunun eğimidir.

Bu en iyi uyum doğrusu, yani, gözlenen değerler ve doğru vasıtasıyla tahmin edilen değerler arasındaki farkların karelerinin toplamını en küçük yapan doğru, *en küçük kareler regresyon doğrusu* diye veya daha basitçe *regresyon doğrusu* diye adlandırılır. Her iki değişkene ait gözlem değerlerinden meydana gelen veri kümesini bir hesap makinasına veya bir kompütere girdikten sonra bir istatistiksel yazılım paketi kullanmak suretiyle regresyon doğrusu kolayca hesaplanabilir.

Alıştırma 1 İki-değişkeli bir veri kümesinde; $\bar{y} = 32$ g, $\sigma_y = 12$ g, $\bar{x} = 1.2$ cm, $\sigma_y = 0.69$ cm ve doğrusal regresyon denklemi $\hat{y} = 14 + 15x$ olarak hesaplanmış. Korelasyon katsayısını bul. ($r = 0.86$)

Eğer bir saçılım grafiği doğrusal-olmayan bir ilişkiye işaret ediyor ise, bir doğrunun (yani bir düz çizginin) uyumunu sağlamaya çabalamamak gerekir.

Uyum İyiliği

Bir regresyon doğrusunun uyum iyiliğini ölçmek için mükemmel bir ölçü henüz yok. Biz burada bu maksada matuf bir kaç kavram üzerinde duracağız. Öyle ki bunlar belirleme katsayısı, model kalıntılarının incelenmesi, dışlakların belirlenmesi ve tesirli noktalar olacak.

Belirleme Katsayısı

Acaba *X* değişkeninin aldığı değerlere tekabül eden *Y* değişkeninin değerlerini regresyon doğrusu ile tahmin ettiğimiz zaman, *y* değerlerindeki değişimi ne kadar azaltmış oluruz? Tahmin edilen $\hat{y}$ değerlerinin varyansının, gözlenen *y* değerlerinin varyansına oranına *belirleme katsayısı* denir ve r^2 ile gösterilir. Yani, r^2, *x* leri bildiğimiz zaman, *y* lerin varyansının tahmin edilebilen kısmının orantısıdır. Başka bir deyişle, r^2 bize, *Y* deki değişimin *X* teki değişim tarafından açıklanan yüzdesini verir.

Korelasyon katsayısı, *r*, daima –1.0 ve +1.0 arasında bir kesir olarak verilirken, belirleme katsayısı, r^2, daima bir yüzde olarak verilir. Eğer bir r^2 nin değeri %100 ise, *Y* deki bütün değişimin *X* teki değişim tarafından açıklandığı mükemmel bir uyumu gösterir. Belirleme katsayısının arzulanan büyüklüğü uygulama alanına bağlıdır. Bilimsel deneylerde, çoğu zaman %90 ve üzerinde bir r^2 hedeflenirken, gözlemsel öğrenme-çalışmalarında %10 ve %20 arasında bir r^2 enformasyonlu olarak addedilebilir. Nazar-ı dikkatini çekmek isterim ki 0.6 lık bir korelasyon 0.3 lük bir korelasyonun iki katıyken, bunlara tekabül eden sırasıyla %36 lık r^2 %9 luk r^2 nin dört katıdır.

Misal 2 Bir sigorta şirketi, satış temsilcilerinden 15 i ile bir yoklama yapar. Her potansiyel müşteriye harcadıkları ortalama zaman (dakika) ve bir haftada sattıkları poliçe sayısı kaydedilir. Müşteri başına harcanan zamanı *X* ve satılan poliçe sayısını *Y* ile temsil etmek suretiyle, veri şöyledir:

x_i	25	23	30	25	20	33	18	21	22	30	26	26	27	29	20
y_i	10	11	14	12	8	18	9	10	10	15	11	15	12	14	11

Bu iki-değişkeli veriyi saçılım grafiğine dök ve en iyi uyum doğrusunu bul.

Cevap: Verinin saçılım grafiği iki değişken arasındaki ilişki hakkında bize sezgisel bir görsel izlenim verecektir. Veri çiftlerini, {(*x*, *y*): (25, 10), (23, 11), (30,14), (25,12), (20, 8), (33, 18), (18, 9), (21, 10), (22, 10), (30, 15), (26, 11), (26, 15), (27, 12), (29, 14), (20, 11), (20, 11)}, düzlemde işaretlediğimiz zaman, şu saçılım grafiğini elde ederiz:

Bu saçılım grafiği, doğrusal olduğu görünen bir ilişkinin mevcudiyetine işaret etmektedir. Yani, noktalar kabaca bir doğrusal hat boyunca yayılmaktadır. Dahası, bu doğrusal ilişki pozitiftir, yani, değişkenin biri artarken diğeri de artmaktadır (doğru yukarı eğimlidir).

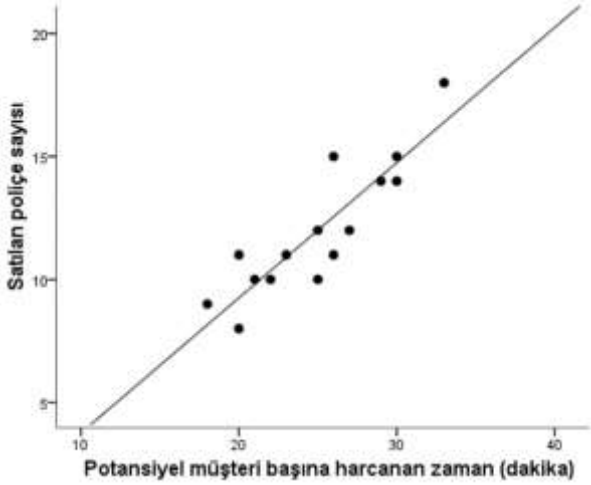
Bir hesap makinası veya kompüter kullanarak korelasyon $r = 0.8836$ olarak hesaplanmıştır. Determinasyon katsayısı $r^2 = 0.78$ olarak hesaplanmıştır ki satılan poliçe sayısına ait değişimin %78 inin, potansiyel müşteri başına harcanan zamana ait değişim tarafından açıklandığını ifade eder. Regresyon doğrusu,

$$\hat{y} = -1.73 + 0.5492x$$

olarak kestirilir. Regresyon doğrusunu şöyle de yazabiliriz:

$$\text{Tahmin edilen satılan poliçe sayısı} = -1.73 + 0.5492(\text{Harcanan zaman})$$

Hesapladığımız regresyon doğrusunu, saçılım grafiğine ekleyelim.



Böylece, mesela, potansiyel müşteri başına 24 dakika harcayan bir müşteri temsilcisinin bir haftada,

$$-1.73 + 0.5492(24) = 11.45$$

adet poliçe satacağını tahmin edebiliriz. Aynı zamanda şuna dikkatini çekmek isterim ki potansiyel müşteri başına harcanan her bir dakikalık ilave zaman, aritmetik ortalama 0.5492 adet daha fazla poliçe satışıyla sonuçlanacaktır.

Eksenlerin etiketlenmeleri ve daima ölçeklendirilmeleri gerekir.

Misal 3 Aşağıda sunulan veri kümesi pazara sürülen aynı tipteki altı deterjan ürünü için yapılan reklam harcamlarını ve toplam satışları kapsamaktadır:

Reklam harcaması (1000 TL):	x:	2.3	5.7	4.8	7.3	5.9	6.2
Toplam satış (1000 TL):	y:	77	105	96	118	102	95

Eğer bir ürünün reklam kampanyası için 5000 TL harcanırsa, toplam satışını tahmin et ve regresyon doğrusunun eğiminin interpretasyonunu yap. Eğer reklam için 100000 TL harcanırsa ne olur?

Cevap: Bir hesap makinası veya kompüter kullanarak, regresyon doğrusunun denklemi şöyle hesaplanır,

$$\hat{y} = 59.691 + 7.293x\ .$$

Regresyon denklemini daha kolay kavranabilir hale getirmek için, x i ve $\hat{y}$ yı daha mütenasipçe isimlendirebiliriz. Mesela, şöyle yazılabilir,

$$\widehat{\text{Satış}} = 59.691 + 7.293(\text{Harcama})\ .$$

Regresyon doğrusu, eğer reklam kampanyasına 5000 TL harcanırsa, toplam satışı,

$$59.691 + 7.293(5) = 96.156$$

$\times$ 1000 TL, yani, 96156 TL olarak tahmin eder.

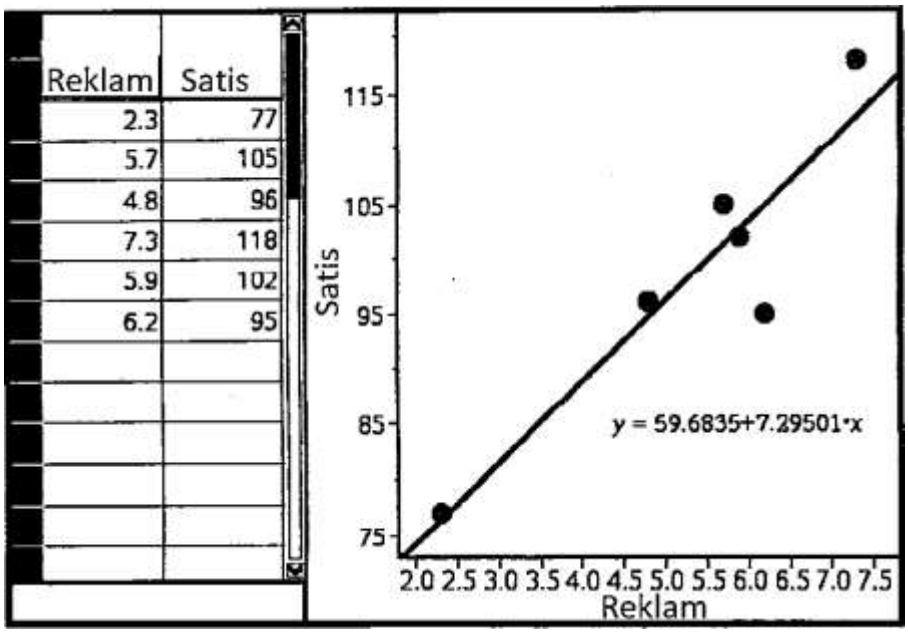
Regresyon doğrusunun eğimi bize şunu söylüyor ki reklama harcanacak her ilave 1000 TL, pazara sürülen ürünün satışında ortalama 7293 TL lik bir artışla sonuçlanacaktır.

Eğer reklama 100000 TL harcanırsa, toplam satış,

$$59.691 + 7.293(100) = 788.991$$

$\times$ 1000 TL, yani, 788991 TL olarak tahmin edilir. Peki, bu cevaba ne kadar güvenebiliriz? Pek fazla güvenemeyiz. Çünkü, regresyon doğrusunu, veri değerlerinin gidiminin çok dışına düşen bir değeri tahmin etmek için kullanmaya çalışıyoruz. Bu prosedüre *ekstrapolasyon* denir ve kullanılırken çok dikkat edilmesi gerekir. Çünkü, ilişkinin aynı biçimde, elimizdeki verinin gidiminin dışında da devam edeceğinin bir garantisi yok.

Şunun iyice anlaşılmasını sağlamak gerekir ki regresyon doğrusunu, verilen bir *x* değeri için bir *y* değeri tahmin etmek üzere kullandığımız zaman, verilen bu *x* değeri için, aslında, aritmetik ortalama *y* değerini tahmin ederiz. Verilen bir *x* değeri için çok sayıda olanaklı *y* değeri vardır ve biz bu olanaklı *y* değerlerinin aritmetik ortalamasını tahmin etmiş oluruz. Böylece, eğer reklam kampanyası için pek çok kere 5000 TL harcansaydı, bu harcamalar çeşitli toplam satış rakamlarıyla sonuçlanabilirdi, fakat, bu satışların tahmin edilen aritmetik ortalaması 96156 TL dir. Bu veriye ait kompüter çıktısı şudur:



Bir regresyon denklemini yazarken, manalı değişken isimleri kullanmak, ilişkinin daha kolay kavranılabilmesi bakımından yararlıdır.

Misal 4 Burger King hamburgerleri üzerinde yapılan bir öğrenme-çalışmasında, satılan

hamburgerlerdeki protein miktarları (gr) ve yağ miktarları (gr) ölçüldü. Veriye ait istatistikler şöyledir:

Protein (gr): $\bar{x} = 17.2$, $s_x = 14.0$
Yağ (gr): $\bar{y} = 23.5$, $s_y = 16.4$
Korelasyon: $r = 0.83$

Bu öğrenme-çalışmasına dayanarak, 31 gr protein içeren bir double hamburgerde ne kadar yağ vardır?

Cevap: Bu değişkenlere ilişkin regresyon denklemini tesis edelim:

$$b = r\frac{s_y}{s_x} = 0.83\frac{16.4}{14.0} = 0.972$$

$$a = \bar{y} - b\bar{x} = 23.5 - 0.972(17.2) = 6.777$$

$$\hat{y} = 6.777 + 0.972x$$

Regresyon denklemini çok daha anlaşılabilir bir formda, şöyle yazalım,

$$\widehat{\text{Yağ}} = 6.777 + 0.972(\text{Protein})\ .$$

Böylece,

$$6.777 + 0.972(31) = 36.909 \text{ gr.}$$

Bu sorunun cevabına yönelik başka bir yaklaşım da şöyle olurdu. Verilen protein miktarına karşılık gelen *z*-puanı,

$$\frac{31 - 17.2}{14.0} = 0.986$$

olduğuna göre, bu öğrenme-çalışmasında rapor edilen aritmetik ortalama protein miktarından 0.986 standart sapma daha büyük bir değerdir. Korelasyon katsayısı $r = 0.83$ olduğuna göre, tahhmini yağ miktarı, aritmetik ortalama yağ miktarından $0.83(0.986) = 0.818$ standart sapma daha büyük bir değerdir. Öyle ise,

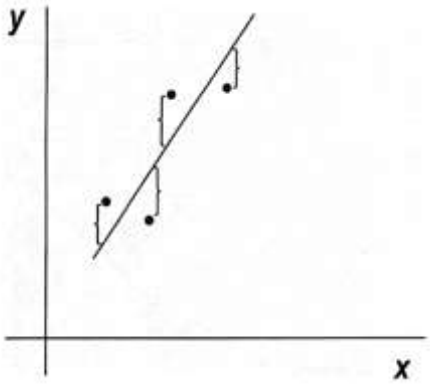
$$23.5 + 0.818(16.4) = 36.915.$$

Gözlenen *x* değerlerinin gidimi dışında ekstrapolasyon yapmak hususunda çok dikkatli olmak gerekir.

Kalıntı Grafikleri

Gözlenen bir değer ve tahmin edilen bir değer arasındaki fark, *kalıntı* olarak adlandırılır. Regresyon doğrusu, saçılım grafiği üzerine çizildiği zaman, bir noktanın kalıntısı o noktadan regresyon doğrusuna kadar olan dikey mesafedir. Regresyon doğrusu öyle bir doğrudur ki kalıntıların kareleri toplamını en küçük yapar.

Her bir veri noktası için kalıntıyı hesaplayabilmelisin.



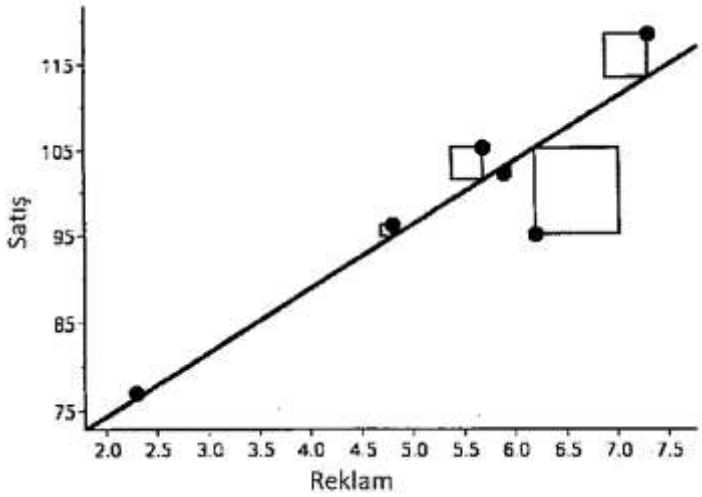
Misal 5 Misal 3 teki regresyon doğrusundan tahmini değerleri hesaplayıp, kalıntıları elde etmek üzere, gözlenen değerlerden çıkartalım:

x_i:	2.3	5.7	4.8	7.3	5.9	6.2
y_i:	77	105	96	118	102	95
$\hat{y}_i$:	76.5	101.3	94.7	112.9	102.7	104.9
$y_i - \hat{y}_i$:	0.5	3.7	1.3	5.1	-0.7	-9.9

Kalıntıların toplamına dikkatini çekmek istiyorum,

$$0.5 + 3.7 + 1.3 + 5.1 - 0.7 - 9.9 = 0.0.$$

Aşağıdaki saçılım grafiği üzerinde gözlenen noktaları, regresyon doğrusunu ve "kalıntıların karelerini" geometrik olarak görebilirsin. Kalıntıların kareleri toplamı minimumdur.



Misal 5 te verilen eşitlik genel olarak doğrudur, yani, *kalıntıların toplamı ve dolayısıyla aritmetik ortalaması daima sıfırdır*.

Kalıntıları matematiksel olarak göstermek için şu notasyon takip edilir:

$$e_i = y_i - \hat{y}_i$$

ve böylece,

$$\sum_{i=1}^{n} e_i = 0.$$

Kalıntıların standart sapmasını da şöyle hesaplarız:

$$s_e = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{\sum e_i^2}{n-2}}.$$

Noktaların regresyon doğrusu etrafında nasıl saçıldıklarının bir ölçüsünü bize s_e verir. Mesela, Misal 5 te Misal 3 teki veriye uydurduğumuz doğrusal regresyon modelinin kalıntılarını hesaplamıştık. Buna göre, kalıntıların standart sapması

$$s_e = \sqrt{\frac{(0.5)^2 + (3.7)^2 + (1.3)^2 + (5.1)^2 + (-9.9)^2}{5-2}} = 6.822\,.$$

Kalıntıların standart sapmasının y değerlerinin standart sapmasından (ki 14.977) daha küçük olduğuna dikkatini çekmek isterim. Demek ki y değerleri regresyon doğrusu üzerindeki y-tahminlerine, kendi aritmetik-ortalamalarına nazaran daha yakınlar (elbette dikey farkları bakımından). Bu da demektir ki gözlenen y_i değerleri için $\hat{y}_i$ değerleri, $\bar{y}$ değerlerine nazaran, daha iyi birer tahmindir.

Kalıntıların grafiğini çizmekle, değişkenler arasındaki ilişki hakkında daha ileri düzeyde enformasyon elde edebiliriz. Bilhassa, belli biçimdeki bir kalıntı grafiği, veriye, doğrusal modelden ziyade bir doğrusal-olmayan modelin daha iyi uyacağının bir göstergesi olabilir. Kalıntıların normal (yani çan eğrisi şeklinde) dağılıp dağılmadıklarının yanısıra, pozitif ve negatif değerlerin dengesine ve aynı zamanda mütekabil y değerlerine kıyasla kalıntıların büyüklüklerine de bakmak gerekir.

Kalıntıların x değerlerine veya $\hat{y}$ değerlerine karşı grafikleri çizilebilir. Çünkü, $\hat{y}$ değerleri x değerlerinin birer doğrusal transformasyonudur. Bu sebeple, garfiğin ölçeği ve eğim negatif olduğu zaman bir sol-sağ tersliği dışında, grafikler özdeştir.

Aynı zamanda, şunu anlamak önemli ki bir doğrusal model mütenasip görünebilir, fakat, düşük bir korelasyonla, zayıf olabilir. Ve, öbür taraftan, bir doğrusal model en iyi model olmayabilir (kalıntı grafiğinin delaletiyle), fakat, yüksek bir r^2 ile, hala çok iyi bir model olabilir.

Misal 6 Belirli bir boyanın kuruma süresi, bu boyaya katılan belli bir maddenin yüzdesine bağlı olarak değişmektedir.

Katkı (%):	x:	1	2	3	4	5	6	7	8	9	10
Kuruma süresi (saat):	y:	4	2.1	1.5	1	1.2	1.7	2.5	3.6	4.9	6.1

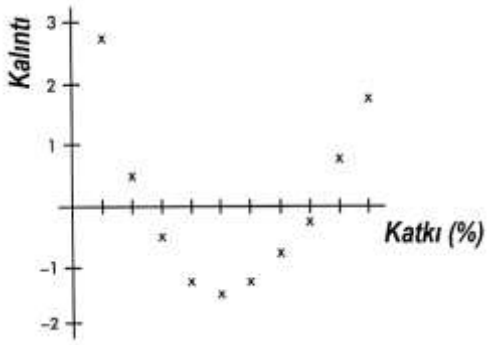
Bir istatistiksel yazılım paketi marifetiyle regresyon doğrusu,

$$\hat{y} = 1.062 + 0.327x$$

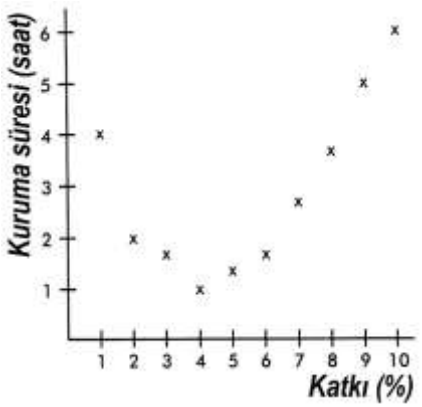
olarak kestirilmiştir. Bu doğrusal regresyonun kalıntıları ise aşağıdaki üzeredir:

x_i:	1	2	3	4	5	6	7	8	9	10
$y_i - \hat{y}$:	2.61	0.38	-0.54	-1.37	-1.5	-1.32	-0.85	-0.08	0.89	1.77
$(y_i - \hat{y})^2$:	6.81	0.14	0.29	1.88	2.25	1.74	0.72	0.01	0.79	3.13

Aşağıdaki kalıntı grafiğinde bariz bir biçim gözlemekteyiz:



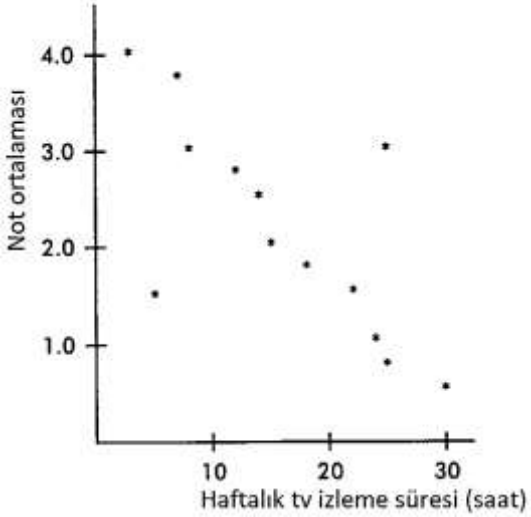
Bu biçim işaret ediyor ki doğrusal-olmayan bir model doğrusal bir modelden daha iyi bir uyum gösterecektir. Orijinal verinin saçılım grafiği, durumu açıkça göstermektedir:



Dışlaklar

Bir saçılım grafiğine baktığımız zaman, dışlakları genel biçimin dışına düşen noktalar olarak teşhis ederiz. Yani, eğer bir noktanın kalıntısı kalıntılar kümesinde bir dışlak ise, o noktanın kendisi de bir dışlaktır.

Misal 7 Bir grup üniversite öğrencisinin haftalık tv izleme sürelerine (saat) karşılık, not ortalamalarının çizilen saçılım grafiği şöyledir:

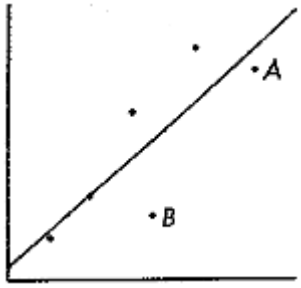


Saçılım grafiğine baktığımızda, görüyoruz ki iki dışlak mevcut: biri, haftalık tv izleme süresi 5 saat olan ama not ortalaması 1.5 olan bir öğrenci, diğeri ise, haftalık tv izleme süresi 25 saat olan ama not ortalaması 3.0 olan bir öğrenci. Yine, görüyoruz ki haftalık tv izleme süresi 30 saat olan öğrenci televizyon izleme süresi değişkeni için ve not ortalaması 0.5 olan öğrenci de not ortalamaları değişkeni için birer dışlak olarak addedilebilecekleri halde, (30, 0.5) noktası regresyon bağlamında bir dışlak olarak addedilmez, çünkü doğrusal biçimin dışına düşmemektedir.

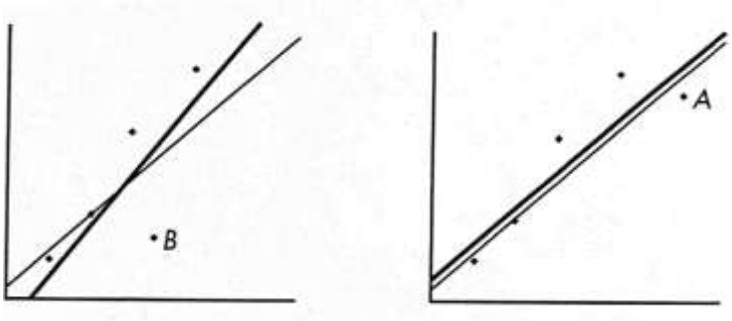
Tesirli Noktalar

Veri kümesinden atılmaları halinde, regresyon doğrusunu barizce değiştirebilecek noktalara *tesirli gözlem* denir. Bazen, bu tanım, aşırı *x* değerlerine sahip noktalara kısıtlanır. Bir tesirli puan küçük bir kalıntıya sahip olabilir, ancak, hala regresyon doğrusu üzerinde kendisinden daha büyük olan fakat ortalama büyüklükte *x* değerlerine sahip kalıntılardan daha fazla etkisi olabilir.

Misal 8 Aşağıda altı noktası olan saçılım garfiğini ve regresyon doğrusunu mülahaza edelim:



Şimdi A noktasını ve B noktasını ayrı ayrı veri kümesinden atalım ve saçılım garfiklerini ve koyu birer çizgiyle regresyon doğrularını yeniden çizelim:



Görüyoruz ki A noktasının atılmasından regresyon doğrusu oldukça etkilendi fakat B noktasının atılmasından öyle fazla etkilenmedi.

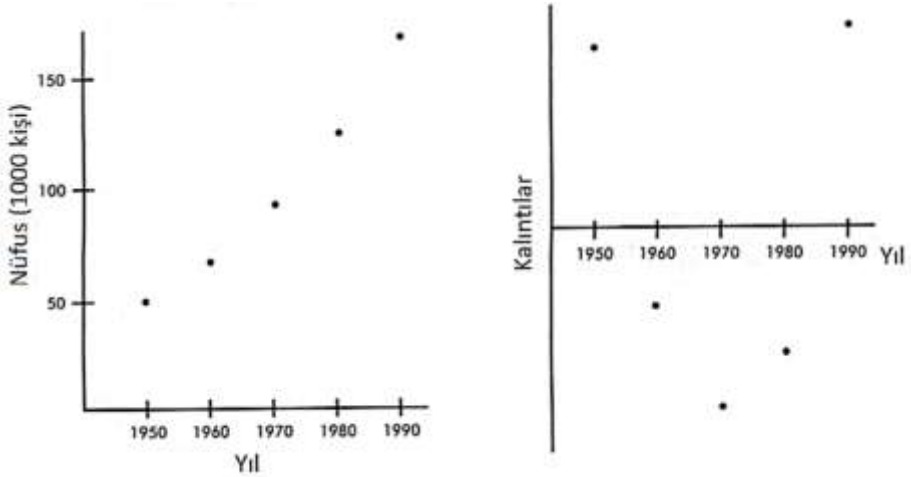
Doğrusallığın Sağlanması: Transformasyonlar

Çoğu zaman, bir doğrusal biçim, iki değişken arasındaki ilişkiyi betimlemenin en iyi yolu değildir. Bu problem, saçılım grafiği belirgin bir eğrisel biçim gösterdiğinde, açıkça farkedilir. Veya, kalıntılar rastgeleliğin dışında bir belirgin biçim gösterdiklerinde bu problem farkedilir. Hal böyle olunca, doğrusal-olmayan model, değişkenlerden birini veya her ikisini birden transformasyona uğratmak suretiyle açığa çıkarıldıktan sonra doğrusal bir ilişki biçimiyle kaydedilir. Faydalı transformasyonlar arasında $\log y$, $\log x$, $\log y$ ve $\log x$, $\sqrt{y}$ veya $1/y$ transformasyonlarını sayabiliriz. Esasında, veriyi transformasyona uğratmak suretiyle yeni değişkenler türetmiş oluruz.

Misal 9 Aşağıdaki yılları ve mütekabil nüfusları mütalaa edelim:

Yıl:	x:	1950	1960	1970	1980	1990
Nüfus (1000 kişi):	y:	50	67	91	122	165

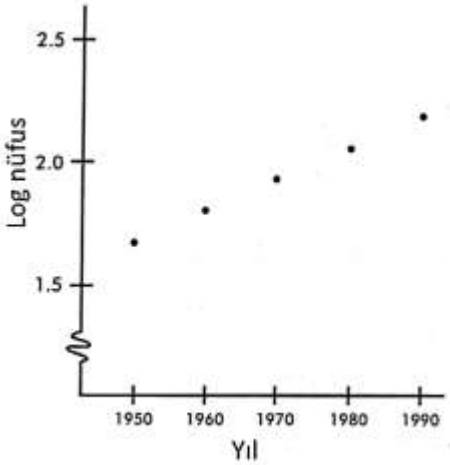
Verinin saçılım grafiği ve kalıntı garafiği gösteriyor ki doğrusal-olmayan bir ilişki daha güçlü bir model olabilir.

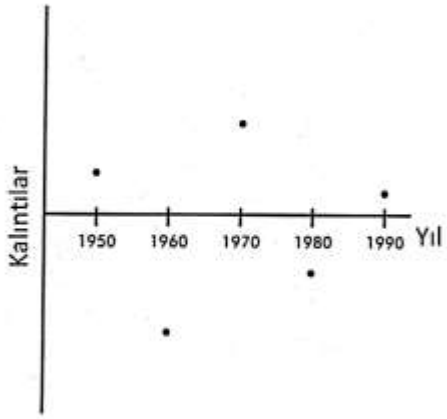


Değerlerin logaritmalarını almak suretiyle, Y değişkenini log Y değişkenine transforme edersek,

x_i	1950	1960	1970	1980	1990
$\log y_i$	1.70	1.83	1.96	2.09	2.22

Şimdi, saçılım grafiği ve kalıntı grafiği daha güçlü bir doğrusal ilişkiyi gösteririyor.





Transformasyona uğrattığımız verinin regresyon denklemi:

$$\widehat{\log y} = -23.65 + 0.013x$$

veya

$$\widehat{\log \text{nüfus}} = -23.65 + 0.013(\text{yıl})$$

Böylece, mesela, 2000 yılı için tahmin edilen nüfusu şöyle hesaplarız,

$$\widehat{\log \text{nüfus}} = -23.65 + 0.013(2000) = 2.35$$

ve

$$\text{Nüfus} = 10^{2.35} = 223.87 \text{ bin veya } 223870.$$

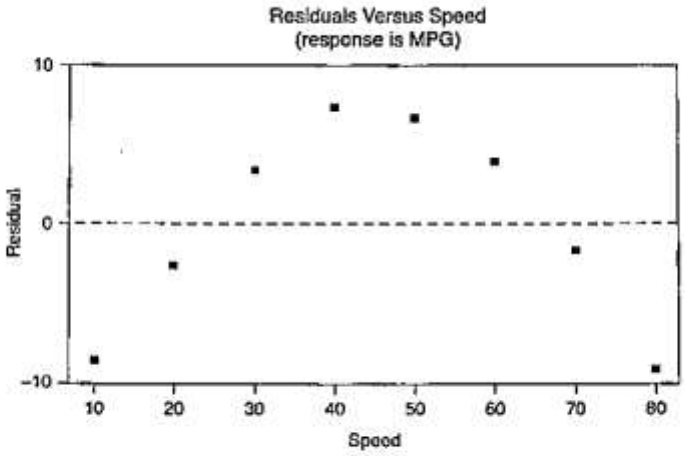
Öte taraftan, $\widehat{\log y} = -23.65 + 0.013x$ "doğrusal" denklemi, $\hat{y} = 10^{-23.65+0.013x}$ olarak yeniden ifade edilebilir.

İstatistiksel Yazılım Paketi Çıktısının İnterpretasyonu

İstatistiksel yazılım paketlerinin verdiği çıktının interpretasyonunu yapabilme yeteneği, sadece bu dersin sınavlarını geçebilmek için değil, aynı zamanda, iş ve bilim dünyasına ait istatistiksel raporları anlamak için de önemlidir. Ancak, şunu üzülerek belirtmek durumundayım ki kendi dilimizde yazılmış ve müştereken kullandığımız bir istatistiksel yazılım paketi henüz mevcut değil. Bu yüzden, karşılaşacağınız kompüter çıktılarının (ve pek tabii ki bu çıktıyı veren yazılım paketlerinin) dili çoğunlukla İngilizcedir.

Misal 10 Yeni model bir otomobilin süratine (mil/saat) karşı galon başına katettiği mesafeye (mil) en küçük kareler regresyon doğrusu uydurulmuştur. Kalıntıların grafiği ve regresyon analizine ait kompüter çıktısının bir kısmı aşağıdadır ("MPG": "miles per gallon" : galon başına katedilen mesafe ve "speed": sürat):

Regression Analysis: MPG Versus Speed




```
The regression equation is
MPG = 38.9 - 0.218 Speed

Predictor        Coef     SE Coef        T       P
Constant       38.929       5.651     6.89   0.000
Speed         -0.2179      0.1119    -1.95   0.099

S = 7.252   R-Sq = 38.7%   R-Sq(adj) = 28.5%

Analysis of Variance

Source            DF        SS        MS        F     P
Regression         1    199.34    199.34     3.79 0.099
Residual Error     6    315.54     52.59
Total              7    514.88
```

a. Regresyon doğrusunun eğiminin bu bağlamda interpretasyonunu yap.
Cevap: Regresyon doğrusunun eğimi –0.2179 ve süratteki (speed) her bir mil/saat lik artış için galon başına katedilen mesafenin (MPG) ortalamada 0.2179 mil/galon azaldığını belirtiyor.

b. Otuz mil/saat lik süratte aritmetik ortalama tahmin edilen galon başına katedilen mesafe (mil/galon) nedir?
Cevap: Saatte 30 millik süratte, aritmetik ortalama tahmin edilen galon başına katedilen mesafe,

$$38.929 - 0.2179(30)$$

ya da yaklaşık 32.4 mildir.

c. Saatte 30 millik süratte galon başına katedilen fiili mesafe neydi?
Cevap: Saatte 30 millik sürat için kalıntı yaklaşık +3.5 tur ve

$$\text{kalıntı} = \text{gözlem} - \text{tahmin}$$

olduğuna göre, gözlenen galon başına katedilen mesafeyi

$$32.4 + 3.5$$

ya da yaklaşık 36 mil olarak kestiririz.

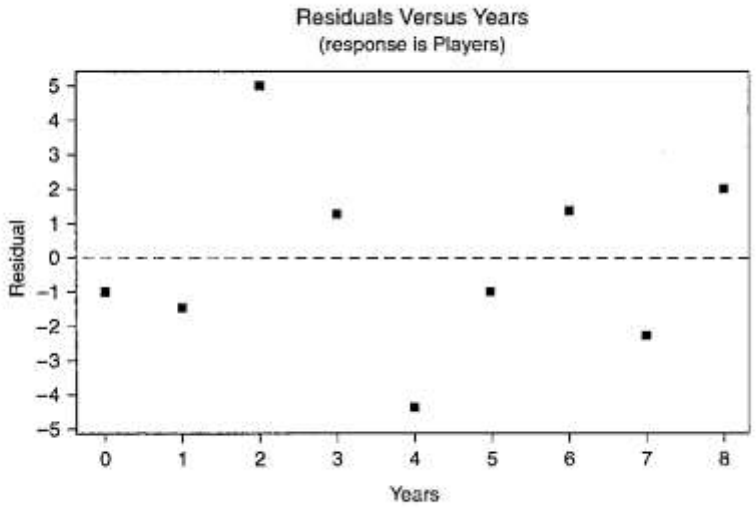
d. Düz çizgi (yani doğru) en mütenasip model midir? Açıkla.
Cevap: Kalıntıların gösterdiği hayli belirgin eğri biçim söylüyor ki bir doğrusal-olmayan model daha mütenasip olabilirdi.

e. "S = 7.252" ne manaya gelir?
Cevap: Kalıntıların standart sapması, $s_e = 7.252$, kalıntılar için tipik bir değerdir (yani kalıntıların işaretsiz aritmetik ortalaması) ve noktaların (yani gözlem değerlerinin) regresyon doğrusu etrafında nasıl yayıldığının bir ölçüsünü verir.

Misal 11 Bir mahalli minikler beyzbol liginde 19955 – 2003 yılları boyunca oynayan çocukların sayılarına en küçük kareler regresyon doğrusu uydurulmuştur. Kalıntıların grafiği ile beraber

regresyon analizinin kompüter çıktısının bir kısmı şöyledir:

Regression Analysis: Number of Players Versus Years Since 1995



Predictor	Coef	SE Coef	T	P
Constant	123.800	1.798	68.84	0.000
Years	12.6333	0.3778	33.44	0.000

S = 2.926 R-Sq = 99.4% R-Sq(adj) = 99.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	9576.1	9576.1	1118.45	0.000
Residual Error	7	59.9	8.6		
Total	8	9636.0			

a. Düz çizginin (yani doğrunun) veri için mütenasip bir model olduğu görülüyor mu? Açıkla.
Cevap: Evet. R-kare büyük bir değere, %99.4, sahip ve kalıntıların grafiği herhangi bir biçim sergilemiyor.

b. Regresyon doğrusunun, bu bağlamdaki denklemi nedir?
Cevap:

$$\widehat{\text{Oyuncu sayısı}} = 123.8 + 12.6(1995 \text{ ten sonraki yıllar})$$

c. Regresyon doğrusunun eğiminin bu problem bağlamında interpretasyonunu yap.
Cevap: Regresyon doğrusunun eğimi 12.6 ve 1995 ten sonra geçen her bir yılda oyuncuların sayısının ortalama 12 veya 13 çocuk arttığını belirtiyor.

d. Regresyon doğrusunun *y*-kesiminin bu problem bağlamında interpretasyonunu yap.
Cevap: *Y*-kesiminin değeri 123.8 ki 1995 teki oyuncu sayısına ait tahmindir. Böylece, bu mahalli minikler ligindeki oyuncuların sayısının 1995 te 124 çocuk civarında olduğunu söyleyebiliriz.

e. Bin dokuz yüz doksan yedi yılı için tahmin edilen oyuncu sayısı nedir?
Cevap: 1997 için, $x = 2$ olduğuna göre, tahmin edilen oyuncu sayısı,

$$123.8 + 12.6(2) = 149.$$

f. Bin dokuz yüz doksan yedi yılında gözlenen oyuncu sayısı neydi?
Cevap: Kalıntıların grafiğinde 1997 yılına ($x = 2$) ait kalıntı +5 tir. Buna göre,

$$\text{gözlenen} - \text{tahmin edilen} = 5$$

olması gerekeceğinden dolayı, 1997 yılındaki oyuncuların sayısı,

$$5 + 149 = 154$$

olmalıdır.

g. Hangi yıllarda, eğer mevcutsa, oyuncuların sayıları bir önceki yıla göre azalmıştır?
Cevap: Eğer herhangi bir yıla ait kalıntının değeri bir sonraki yıla ait kalıntının değerinden

12.6 çocuk daha büyük olursa o yıldaki oyuncu sayısı azalmış olacaktır. Böyle bir şey bu halde mevcut değil. Böylece, minikler ligindeki oyuncuların sayısı hiç bir zaman azalmıyor.

Bir hesap makinesi kullanarak, bir regresyon doğrusunu sadece bulmak yetmez; bulduğun doğruyu anlaman gerekir (mesela, verinin ait olduğu değişkenler bağlamında, doğrunun eğiminin ve kesiminin interpretasyonunu yapabilmelisin).

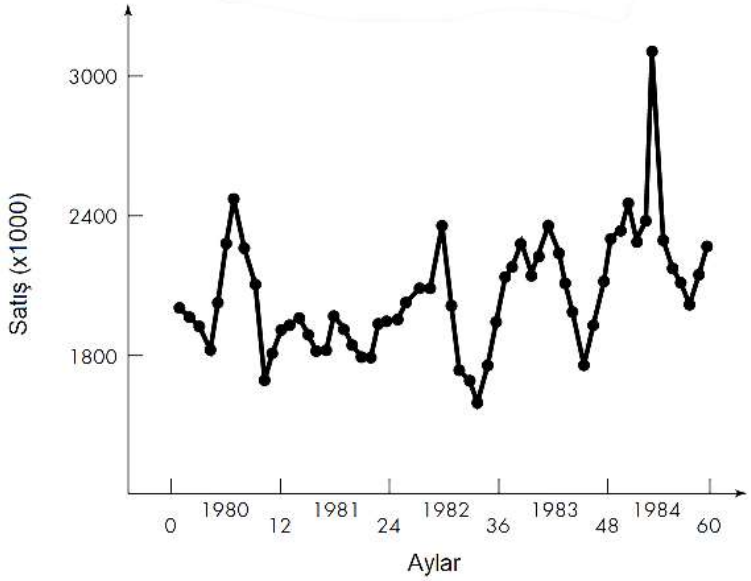
Zaman Serileri

Mesela hisse senedi fiyatları, gayrimenkul fiyatları veya döviz kurları gibi bazı değişkenlerin zamana göre değişimini betimleyen grafikleri görmüşsündür. Basitçe tanımlarsak, **zaman serisi** bir değişkenin değerinin zaman boyunca kaydedilmesidir ve çoğunlukla eşit zaman aralıklarında ölçülür. Bilhassa ekonomik indikatörlerin pek çoğu periyodik ölçülen zaman serileridir. Ölçüm periyodu yıl, ay, gün, saat, dakika ve hatta saniye olabilir. Bir zaman serisi analizinde araştırmacının ilgisi daha ziyade ölçülen değişkenin gelecekte alacağı değeri tahmin etmeye odaklıdır.

Zaman boyunca sunulan veriyi anlamak için, bir zaman serisinin çıkışlarına ve inişlerine katkıda bulunan çeşitli bileşenlerin nasıl mülahaza edileceğini bilmek önemlidir. Aksi halde, geçici dönemsel bir çıkışı devamlı artan bir trend zannedebilirsin ve hiç te akıllıca olmayan bir ekonomik karar verebilirsin.

Zaman Serisi Grafiği

Aşağıda gördüğün bir zaman serisi grafiğidir. Veri büyük bir kot pantolonu markasının Ocak 1980 den Aralık 1984 e kadarki 5-yıllık bir periyotta Avrupa'da yaptığı aylık satışları temsil ediyor.



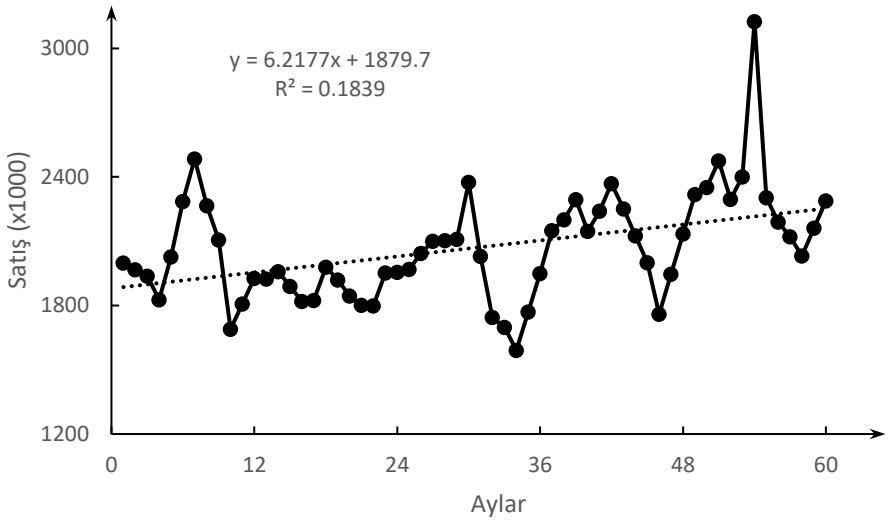
Veri noktalarının bir çizgiyle biribirine bağlandığını görüyorsun. Böyle yaparak, zaman boyunca gerçekleşen çıkışları ve inişleri daha kolay takip edebiliyoruz. Veri 1000 adet satış cinsinden ölçülmüş. Birinci ay Ocak 1980 ve 60 ıncı ay Aralık 1984.

Zaman Serisinin Bileşenleri

Pek çok zaman serisi şu dört temel bileşene sahiptir: *uzun-vadeli trend, sezonsal bileşenler, düzensiz dönemler* ve *rastgele oynamalar*.

Uzun-vadeli Trend

Bir zaman serisinin ölçtüğü değişkenin çıkışları ve inişleri zaman boyunca bir kararlılık gösteriyorsa, bu kararlığa **trend** adı verilir. Grafiğine bakmak suretiyle, bir zaman serisinin trendini anlayabiliriz. Mesela yukarıdaki grafikte sunulan kot pantolonu satışlarının artan bir trendi olduğunu açıkça görebilmiş olmalısın.



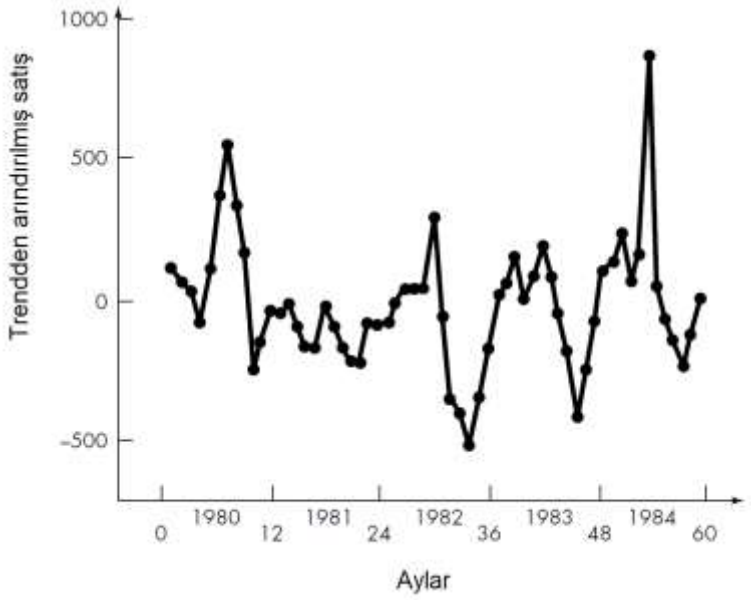
Eğer uzun-vadeli trend doğrusal ise, bir regresyon doğrusu hesaplamak suretiyle, bu trendi kestirebiliriz. Bu halde, zaman periyodu regresyon denklemimizde açıklayıcı değişken ve zaman serisi değişkeni de yanıt değişkeni olacaktır. Daha sonra, seride daha başka hangi ilginç hususiyetlerin mevcut olduğunu görebilmek için, trendi seriden giderebiliriz. Bunu yaptığımızda, elde edeceğimiz sonuç **trendsizleştirilmiş zaman serisi** olacaktır.

Kot pantolonu satışları verisine ait regresyon doğrusu şudur:

$$\widehat{\text{satış}} = 1880 + 6.22(\text{ay})$$

Dikkatini çekmek isterim ki 1 inci ay Ocak 1980 ve 60 ıncı ay Aralık 1984 tür. Regresyon doğrusu bize gösteriyor ki satışların aylık ortalama 6.22 lik artan bir trendi var. Ölçü birimi 1000 adet olduğu için, aylık gerçek artış ortalama $6.22(1000) = 6220$ adettir.

Trendsizleştirilmiş zaman serisinin grafiği ise şudur:



Bu grafiği ilk grafikle karşılaştıralım. Bu grafikte geriye kalan oynaklık ilk grafiktekiyle benzer bir karaktere sahip. Fakat, ilk grafikte yukarı giden trend burada artık yok.

Bir trend ihtiva eden bir zaman serisi *durağan değildir*. Belirli bir trend modellenip, zaman serisinden giderildikten sonra, o veri kümesinin *trend-durağan* olduğu söylenir. Belirsiz trendler serinin farklarını almak suretiyle giderilebilir, bu halde, o veri kümesinin *fark-durağan* olduğu söylenir

Alıştırma 2 Zaman serilerinde trendsizleştirmenin basit veya daha karmaşık çeşitli teknikleri

mevcuttur. Bu tekniklerin neler olduğunu başka kaynaklardan kısaca öğrenmeye-çalışabilirsin. Kot pantolonu satışları verisi için hangi trendsizleştirme tekniği kullanılmış olabilir?

Sezonsal Bileşenler

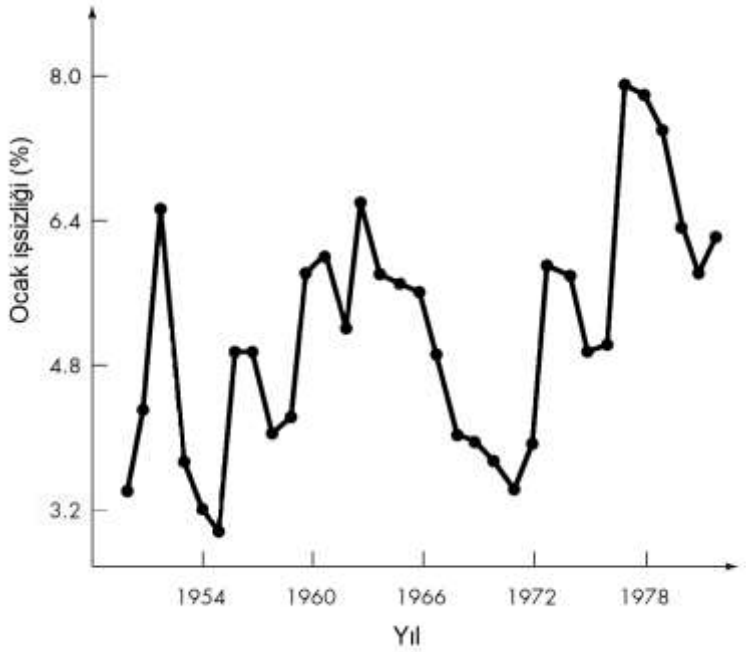
Ekonomik veriye ait veya insanların davranışları ile alakalı pek çok zaman serisi **sezonsal bileşenlere** sahiptir. Başka bir deyişle, bu zaman serileri mesela her yıl belli aylarda veya sezonlarda yüksek, diğerlerinde düşüktür. Mesela konut satışları daha sıcak aylarda çok daha yüksektir. Mesela işsizlik oranı Ocakta yükselme eğilimindedir çünkü dışarı işleri en azdır.

Ekonomik indikatörlerin pek çoğu sezonsal oynamalara tabidir. Zaman serileri, yine, *sezonsal olarak düzeltilmiş* olarak enformasyon kullanıcılarına sunulur. Ekonomistlerin, zaman serilerini sezonsal olarak düzeltmek için istihdam ettikleri sofistike teknikler vardır. Bunlara bu dersin kapsamında değinmiyoruz.

Kot pantolonu satışlarına ait zaman serisine yeniden dönersek, hem orijinal seride hem de trendsizleştirilmiş seride bir sezonsal bileşenin varlığını görebiliriz. Her yılın Haziran ve Temmuz aylarında satışların pik yaptığı ve Ekim ayında dip yaptığı aşikar. Eğer enformasyon kullanıcısı bir zaman serisine dahil olan sezonsal enformasyona sahip değilse, yalnızca genel trende matuf tahminlerinde büyük yanılabilir.

Düzensiz Dönemler ve Rastgele Oynamalar

Çoğu durumda bir zaman serisinin geriye kalan bu iki bileşeni arasında ayrım yapmak kolay değildir. Bilhassa dönemler düzenli değilse. *Düzensiz (fakat düzgün) dönemler* genellikle ekonomik sistemlerin takip ettiği bir eğilimdir ve *rastgele oynamalar* açıklanabilir değildir. Müteakip grafik 1950 den 1982 ye kadar her Ocak ayındaki sezonsal olarak düzeltilmiş ABD işsizlik oranlarını gösteriyor.



Bir kaç yıl boyunca çıkan ve inen işsizlik oranlarından meydana gelen belgili *düzensiz dönemlere* dikkatini çekmek isterim. Bunların bazıları belki bazı sosyal ve politik faktörler tarafından en azından kısmen açıklanabilir. Mesela, ABD'nin dahil olduğu Vietnam Savaşı 1960 ların ortasından 1970 lerin başına kadar devam etmişti. Zorunlu askerliğin sona ermesiyle beraber pek çok genç adam iş piyasasına dahil olmuştu. Zaman serisi grafiğinde Vietnam savaşı yıllarını kapsayan bir iniş dönemini görebiliyoruz.

Bir zaman serisindeki **rastgele oynamalar** diğer üç bileşen çıkarıldığında geriye kalan şey olarak

tanımlanır. Tüm ölçümlerde mevcut olan doğal değişkenliğin bir parçasıdır. Kot pantolonu satışlarına ait zaman serisinde yaptığımız üzere, uzun-vadeli trendi, sezonsal bileşenleri ve düzgün düzensiz dönemleri hesaba katmış olsak dahi, aylık kot pantolonu satışlarını hala mükemmelce açıklayabilmeyiz. Geriye kalan, açıklanabilmez bileşenler rastgele oynamalar olarak adlandırılır.

Ekonomik zaman serilerinin bazıları enflasyona veya sezonsal bileşenlere göre düzeltilmeden yayımlanır. Bu sebeple, bir ekonomik indikatöre ait bir raporu okurken, zaman serisinin enflasyona ve/veya sezonsal bileşenlere göre düzeltilip düzeltilmediğini kontrol etmelisin.

Zaman serileri yalnızca ekonomik indikatörlere mahsus değildir, elbette. Başka bilim alanlarında da pek çok değişkene ait zaman serileri kullanılır. Mesela, okyanus dalgalarının boyu, güneş lekelerinin sayısı, beynin elekriksel aktivitesi, yağış miktarı, abone sayısı, nabız sayısı, sıcaklık, elektrik tüketimi ve benzeri değişkenler zaman serileriyle ölçülür ve analiz edilir.

Bir zaman serisi verisini okurken, şu soruları sormalısın:

1. Zaman periyotları eşit aralıklı mı?
2. Ekonomik zaman serisi enflasyona göre düzeltilmiş mi?
3. Değerler sezonsal olarak düzeltilmiş mi?
4. Tipik uzun-vadeli davranışı temsil edecek kadar, seri yeterli bir zaman dilimini kapsıyor mu?
5. Yukarı-yönde veya aşağı-yönde bir trend var mı?
6. Giderilmeyen başka sezonsal bileşenler var mı?
7. Düzgün dönemler var mı?

Bu sorulara vereceğin cavaplara istinaden, rapor edilmiş olan veriye ait bir düzeltme hesaplamak veya en azından bir yaklaşımda bulunmak ihtiyacını duyabilirsin.

Mini Proje 06

MP 01 de topladığın veri kümesinde yer alan iki nicel değişken arasındaki ilişkiyi betimle. Sonuçların interpretasyonunu yap.

BÖLÜM 07 | VERİ TOPLAMA METOTLARI

Giriş

Gerçek dünyada, zaman ve maliyet kısıtları, çoğu durumda tüm popülasyonu analiz etmeyi olanaksız hale getirir. Türkiye İstatistik Kurumu aylık işsizlik oranını ilan etmeden önce sana veya ailene soruyor mu? Televizyon yapımcısı, bir pilot programın devam edip etmeyeceğine karar vermeden önce, her hanenin seyir tercihine bakıyor mu? Bir ilaç şirketi, yeni geliştirdiği bir kalp ilacını piyasaya sürmeden önce, tüm hastalar üzerinde deniyor mu? İstatistik bilimini öğrenme-çalışmamızda, bir *örneği* ele alarak, popülasyonun karakteristiklerini nasıl kestireceğimizi öğreniyoruz. Mesela, ilerleyen bölümlerde, örneklerin aritmetik ortalamalarına ve orantılarına bakarak popülasyonların aritmetik ortalamalarını ve orantılarını nasıl kestireceğimizi göreceğiz. Veya regresyon doğrusunun parametrelerini.

Daha geniş bir popülasyona dair neticelere varmak için, çektiğimiz örneğin popülasyonu hakkıyla temsil ettiğinden emin olmamız gerekir. Veriyi, kompüter programları yardımıyla analiz etmek, çoğu zaman, veriyi toplamaktan daha kolaydır. Fakat unutmamalıyız ki "çöp girerse çöp çıkar." Eğer veri iyi toplanmamışsa, hiç bir işe yaramaz. Ne yazık ki gazetelerden, radyolardan, televizyon kanallarından ve internet sitelerinden her gün üzerimize yağdırılan istatistiklerin pek çoğu yetersiz dizayn edilmiş veri toplama prosedürlerine dayanmaktadır.

Nüfus Sayımı

Popülasyona ait tüm bireyleri gözlemeye **nüfus sayımı** denir. Bilinen manasıyla, popülasyonun her üyesiyle resmen temas kurulmasını ve yaş, medeni durum, ırk, cinsiyet, meslek, gelir, öğrenim durumu gibi değişkenlere ait verinin toplanmasını kapsar. Türkiye İstatistik Kurumu (TÜİK) ülkede yaşayan herkese ait enformasyonu belli periyotlarda elde etmeye çalışır. Türkiye Cumhuriyeti devletinin tüm gayretiyle, muazzam bir veri yığınına ulaşılmasına rağmen yine de nüfus sayımı hiç bir zaman tam değildir. Mesela, evsizlerin çoğu ıskalanır ya da geçici olarak kaldıkları ayrı ayrı yerlerde sayılabilirler ve her zaman soruları yanıtlamayan bir takım aileler geride kalır. Amerika Birleşik Devletleri'nde 2000 yılında yapılan nüfus sayımında $3.3\ milyon$ kişinin ıskalandığı tahmin edilmektedir ki bu toplam popülasyonun $\%1.2$'sidir.

İster özel sektörde isterse devlet sektöründe olsun, pek çok öğrenme-çalışmasında, tüm popülasyonun gözlenmesi akıl dışıdır. Çünkü, zaman ve maliyet daima baş rolü oynar. Dahası, tüm veriyi elde etme çabasının üstünkörülüğe yol açacağı da bilinir. Nihayetinde, ve en önemlisi, iyi dizayn edilmiş ve iyi yürütülmüş bir *örnek yoklaması*, popülasyonun tümüyle gözlendiği yetersiz dizayn edilmiş bir öğrenme-çalışmasından çok daha üstündür. Mesela, kötü ifade edilmiş bir soru, popülasyondaki herkes tarafından yanıtlansa bile elde edilen veri manasızdır.

Örnek Yoklaması

Nüfus sayımı, herkesi sayma çabasıdır; ve bir örnek değildir. Bir **örnek yoklaması**, bir parçasını örnek olarak ele almak yoluyla popülasyonun tümü hakkında enformasyon sağlamayı hedefler. Amaç, bir popülasyonun doğal durumunu bozmaksızın veya değiştirmeksizin o popülasyon hakkında enformasyon kazanmaktır. Örnekleme ile veri toplamak için çok çeşitli prosedürler kullanılır ve kitle medya araçlarının bize ulaştırdığı istatististiksel enformasyonun çoğu örnek yoklamalarının mahsülüdür. Çoğu zaman, örnek yoklamalarının ortaya koyduğu ilişkileri tanıtlamak için sonrasında kontrollü deneylere girişilir.

Mamafih, bir örneği anında geçersizleştirecek ve yararlı enformasyon elde etmeyi olanaksız kılacak şeylerden biri *yanlılık*tır. Eğer ciddi bir şekilde popülasyonu temsil kabiliyetinden yoksun ise, bir örnek *yanlı*dır. Yanlılığı izale etmenin başlıca tekniği, gözlem ünitelerinin seçiminde *rastgeleliğe*

uymaktır. Rastgeleleştirme, hem bilinen hem de bilinmeyen etkilerden ve tesirlerden bizi korur. Nihayetinde, örnek hacmi ne kadar büyürse, sonuçlar o kadar güzel olur, fakat önemli olan örneğin hacmidir, populasyona oranı değildir. Yani, hacmi 100 milyon birey olan bir popülasyondan çekilen 500 birey hacmindeki bir rastgele örnek, hacmi 100 bin birey olan bir popülasyondan çekilen 500 birey hacmindeki bir rastgele örnek kadar iyi temsil kabiliyetine sahip olabilir.

Değişkenlik örnek hacmi ile kontrol edilirken, *yanlılık* rastgeleleştirme ile kontrol edilir.

Deney

Kontrol-altına-alınmış bir öğrenme-çalışmasında, ki **deney** olarak adlandırılır, araştırmacı *gözlem ünitelerini* (insan, hayvan, nesne, durum, vb.) mütenasip gruplara rastgele atamak zorundadır. Gruplardan en az birine bir uygulama yapılır ve yanıtlar gözlenir. Mesela, içinde aspirin veya asetaminofen olan işaretsiz kapsüller hastalara rastgele verilebilir ve bu tedavinin etkileri ölçülür. Deneylerin pek çoğunda bir *muamele grubu* ve bir *kontrol grubu* vardır ve ideal koşullarda ne denekler ne de araştırmacı hangi grubun hangisi olduğunu bilir. Bu *çifte-kör* yaklaşımın en klasik misali, 1950'de Amerika birleşik Devletleri'nde yapılan çocuk felci aşısı deneyidir. Çocukların rastgele seçilen yarısına aşı yapılırken diğer yarısına plasebo yapıldı. Hem uygulayıcı doktorlar hem de çocuklar hangi çoçuğun aşı hangi çocuğun plasebo aldığından habersizdi. Neden-ve-etki ilişkileri ancak kontrol-altına-alınmış deneylerle belirlenebilir.

Bir öğrenme-çalışması, yanıtı gözlemek için, insanlara, hayvanlara, bitkilere veya nesnelere kasten herhangi bir muamele uyguladığımız zaman bir deneydir.

İyi deney dizaynının arkasındaki önemli prensipler *kontrol* (kimin hangi muameleye tabi tutulduğunun dışında, tüm gruplar için koşullar mümkün olduğunca benzer olmalıdır), *bloklama* (belli farklılıkları doğrudan hesaba katmak için, denekler kendi içinde benzerlik gösteren gruplara ayrılabilir), *rastgeleleştirme* (kimin hangi muameleye tabi tutulacağı rastgeleleştirilerek, bilinmeyen ve kontrol altına alınamayan farklılıkların üstesinden gelinir), *tekrarlama* (muamelelerin yeterli sayıda deney ünitesi ile tekrarlanması gerekir) ve *genelleştirilebilirliktir* (bir deneyin çeşitli kereler yinelenme, yani yeniden kurulabilme ve aynı sonuçları verebilme kabiliyeti).

Gözlemsel Öğrenme-çalışması

Örnek yoklamaları gözlemsel öğrenme-çalışmalarına bir misaldir. Gözlemsel öğrenme-çalışmalarında, gözlem ünitelerinin (yani deneklerin) muamele ve kontrol gruplarına rastgele ataması yapılmamıştır. Mesela bir araştırmacının 90 kişiye günde iki paket sigara içmelerini ve diğer 90 kişiye günde bir paket sigara içmelerini söylemesi etik olabilmez. Bunun yerine, yapabileceği sadece kendiliğinden bu miktarlarda sigara içen kişileri gözlemektir. Gözlemsel öğrenme-çalışmalarında bilimsel araştırmacı dikkate alınan yanıta hangi değişkenlerin etki ettiğini belirlemeye elbette kendince bir çaba gösterebilir. Fakat, sonuçlar ilişkileri ortaya koyuyor olsa bile, neden-ve-etki neticesine varmak hayli zordur.

Gözlemsel öğrenme-çalışmaları başlıca ve önemli veri kaynakları olmalarına rağmen, açıklayıcı değikende meydana gelen değişimin yanıt değişkeni üzerindeki etkisini ölçmenin zayıf bir motodudur. Açıklayıcı değişkende meydana gelen değişime gözlem ünitelerinin verdiği yanıtları değerlendirmek istiyorsak, bu değişimi kendimiz yaratmalıyız. Yani bir deney icra etmek zorundayız. Bu minvalde, bir değişkenin başka bir değişken üzerindeki etkisini inceleyen gözlemsel öğrenme-çalışmaları çoğu zaman başarısız olur. Çünkü açıklayıcı değişkenlerin etkileri başka değişkenlerle *karışmaktadır*.

Vaka Öğrenme-çalışması

Bir vaka öğrenme-çalışması bir bireyin veya az sayıdaki bireyin derinlemesine incelenmesidir. Araştırmacı ele aldığı bireyi ve ilgilendiği mevzu hakkında bilgisi olan başkalarını gözler ve onlarla

görüşme yapar. Mesela, ünlenmiş bir şifacı medyumu öğrenmeye-çalışmak için, bir araştırmacı onu iş başında gözleyebilir, onunla şifa teknikleri hakkında görüşebilir ve şifacının muamelesine maruz kalan müşterileri ile görüşebilir. Bu dersin müfredatına vaka öğrenme-çalışmalarını dahil etmiyoruz, çünkü vaka öğrenme-çalışmaları yalnızca betimseldir ve çıkarımsal istatistiksel metotları gerektirmezler. Şunu unutma ki bir vaka öğrenme-çalışmasının bulgularını üzerinde öğrenme-çalışması yapılandan başka kişi veya durumlara genelleştirebileceğini düşünmemelisin. Uygulamada, nadir ve temsilci-olmayan durumları hassasiyetle soruşturmak için vaka öğrenme-çalışmaları kullanılabilir.

Yoklamaların Planlaması ve Yürütülmesi

Çoğu veri toplama faaliyeti, kontrollü deneylerden ziyade, gözlemsel öğrenme-çalışmalarını ihtiva eder. Dahası, çoğu veri toplama faaliyeti bir maksada matuf olsa da pek çok öğrenme-çalışması veri toplandıktan ve incelendikten sonra akla gelir. Veri toplamanın yararlı olabilmesi için, elde edilecek örneğin mutlaka ilgilenilen popülasyonun bir temsilcisi olması gerekir. Örnek yoklamalarında en çok kullanılan veri toplama enstrümanı anketlerdir. Bir *anket*, alternatif cevapları ile beraber basılmış veya yazılmış sorulardır. Bu sorular anketçi tarafından yanıtlayıcılara sorulur ve yanıtlar kaydedilir. Kağıt veya dijital formu yanıtlayıcıların doldurmaları da istenebilir. Belirli bir araştırmaya mahsus olarak, o araştırmanın mahiyetine istinaden, bir anket, kapalı-uçlu soruların yanında açık-uçlu soruları da barındırabilir.

Anket bir veri toplama enstrümanıdır. Oysa, yoklama bir veri toplama ve analiz etme metodudur.

Basit Rastgele Örnekleme

İyi, yani temsilci bir örnek nasıl seçilebilir? En isabetli teknik popülasyonun her bir üyesinin adını bir kağıda yazmak, kağıtları bir kutu içerisinde karıştırmak ve belirli sayıda kağıdı bakmadan kutudan çekmektir. Bu metot popülasyondaki herkese örneğin bir parçası olarak eşit seçilme şansı verecektir. Ne yazık ki bu metot genellikle çok zaman alır ve maliyeti yüksektir. Bunların yanısıra, eğer yeterince karıştırılmadan çekim yapılırsa, yanlılık baş gösterebilmektedir.

Basit rastgele örnek öyle bir örnektir ki aynı hacimdeki her olanaklı örnek ile eşit seçilme şansına sahiptir. Bir basit rastgele örnek çekmenin en kolay yolu popülasyondaki her bireye bir numara vermek ve bir rastgele sayılar tablosu kullanmak suretiyle ya da bir kompüterde rastgele sayılar üretmek suretiyle seçilenleri belirlemektir.

Basit rastgele örnekleme, bir *örnekleme çerçevesi* gerektirir. Örnekleme çerçevesi, örneğin fiilen içinden çekileceği örnekleme ünitelerinin bir listesidir. Çerçevenin hedef popülasyonu kapsıyor olması (yani, popülasyondaki her hangi bazı grupları dışarda bırakmaması) önemlidir.

Misal 1 Bir İstatistik dersine 80 öğrenci kayıtlıdır. Öğretim üyesi uygulamalı bir sınav denemesi yapmak için 10 öğrenciden oluşan bir örneği rastgele seçmek istiyor. Öğretim üyesi önce öğrencilere 01, 02, 03, ..., 80 numaralarını atar. Rastgele sayılar tablosundan (Tablo 1) bir kerede iki hane okuyarak, 80'den büyük olanları ve tekrarlayanları gözardı ederek 10 farklı sayı belirler. Eğer tablonun başlangıç satırında 25638 70268 93605 87101 24685 98768 43202 03165 49815 02484 sayıları yer alıyor ise, öğretim üyesinin seçeceği öğrencilerin numaraları şunlar olacaktır: 25, 63, 02, 68, 60, 58, 71, 01, 24, 59. Dikkat edersen 87 ve 93 sayılarını görmezden geldik çünkü 80'den büyükler ve 68 sayısının ikinci tekrarını yine görmezden geldik.

İyi-dizayn-edilmiş ve İyi-yürütülmüş Bir Yoklamanın Özellikleri

İyi-dizayn-edilmiş bir yoklama daima şansı bünyesinde barındırır. Bu yukarıda anlatıldığı gibi bir rastgele sayılar tablosunun kullanılması ile veya bir kompüterin kullanılması ile sağlanabilir. Bununla beraber, olasılık tekniklerinin kullanılması bir temsilci örneği garanti etmekte yeterli değildir. Çoğu zaman, popülasyonun tam bir listesi elimizde olmaz ve böyle bir durumda "şansı" tam olarak nasıl uyguladığımız hakkında dikkatli olmamız gerekir. Hatta denekler şansa bağlı çekilse bile, yoklamaya yanıt vermeyi reddedenler veya yanıt vermek için müsait olmayanlar olabilir. Bu durumda sonuçta elde edilecek örneğin temsilciliğinin gerçekte nasıl olduğu tartışılır.

Soruların ifade edilmesi nötr olmalıdır. Yani soruya alınacak yanıta tesiri olmamalıdır. Tüm deneklerin aynı soruyu aynı manada ve aynı açıklıkla anlamasını temin etmek önemlidir. Yoksa denekler aynı yoklama sorusuna ifadesine bağlı olarak farklı yanıtlar verebilir.

Misal 2 Farzedelim ki küçük bir kasabada besleyici bir kahvaltı yapan yetişkinlerin yüzdesini belirlemekle ilgileniyoruz. Kasabanın telefon rehberinden 100 numarayı rastgele seçip, her birini

arayarak telefona çıkan yetişkine her sabah besleyici bir kahvaltı yapacak kadar zeki olup olmadığını soralım. Münasip mi?

Cevap: Rastgele seçim iyi, fakat cevaplanması gereken bir kaç soru var. Mesela, kasabada telefonu olmayan çok kişi var mı? Ya da numaralarını telefon rehberinde listeletmeyen? Aramaların yapıldığı günün saati seçilmiş olan kişinin ulaşılabilir olup olmadığını nasıl etkileyecek? Eğer kişiler ulaşılabilir değilse, yedekler yine aynı yoldan rastgele mi seçilecek yoksa bu durum belirli bir sınıf insanın daha az temsiline mi yol açacak? Nihayetinde, bu meselelere tatmin edici cevaplar verilse bile yoklama sorusunun ifadesi açıkça nötr değil. *Zeki olup olmama* ibaresi çıkarılsa bile yanıtlar yine de neredeyse manasız olacaktır. Çünkü *besleyici bir kahvaltı* ibaresi her denek için muhtemelen aynı manayı taşımıyor olacaktır.

Örnekleme Hatası: Bir Yoklamanın Özünde Olan Değişim

Bir yoklama ne kadar iyi-dizayn-edilmiş ve iyi-yürütülmüş olursa olsun, hala bir popülasyon *parametresi* için bir kestirim olarak bir örnek *istatistiği* verir. Farklı örnekler, hepsi aynı popülasyon parametresi için birer kestirim olarak farklı örnek istatistikleri verebilirler ve böylece *örnekleme hatası* olarak adlandırılan hata doğal olarak ortaya çıkar. (Yani, örnekleme hatası, yapılan bir yanlışlığı değil, aksine, aynı popülasyondan çekilen rastgele örnekler arasındaki doğal değişkenliği ifade eder.) Bu hata olasılık kullanılarak betimlenebilir. Bu da demektir ki belli büyüklükteki bir hatanın olabilirliğinin ne olduğunu hesaplayabiliriz. Genellikle, bu hatanın meydana gelme şansı örnek hacmi arttıkça azalır. Mamafih, verinin hangi yoldan toplandığı çok önemlidir. Geniş bir örnek hacimi ile, ne yazık ki kötü bir yoklama dizaynını veya yanlış bir veri toplama tekniğini kurtaramayız.

Çekilen örnek eğer popülasyonun temsilcisi değilse, örnekleme hatası mevcuttur.

Misal 3 Dört farklı kamuoyu araştırma şirketi yaptıkları örnek yoklamalarının sonucunda ülkenin görevdeki hükümetine seçmenlerin verdiği desteği %54.9, %55.6, %52.8 ve %49.7 olarak açıkladılar. Doğru yüzde kaçtır? Kamuoyu araştırma şirketlerinin en az üçü veya hepsi yanlışlık mı yaptı?

Cevap: Verilen enformasyona dayanarak, doğru popülasyon yüzdesini bilmenin yolu yok. Dört ayrı örnek yoklaması bu dört istatistiği ortaya koymuştur ki her biri popülasyon parametresinin ayrı bir kestirimidir. Çünkü, mesela, şansa bağlılığa riayet edilmediği, popülasyonun iyi bir temsilinin sağlanamadığı veya sorunun iyi ifade edilmediği gibi kötü bir örnek yoklaması yapmış olduğunu bilmeksizin hiç bir kamuoyu araştırma şirketinin yanlış yaptığını söyleyemeyiz.

Yoklamalarda Yanlılığın Kaynakları

Kötü dizayn edilmiş örnekleme teknikleri *yanlılığa* yol açar. Yanlılık, popülasyonun belli üyelerinin seçimi lehinde bir temayüldür. Eğer bir öğrenme-çalışması yanlı ise örnek hacminin bir yararı olmaz. Yani ne kadar geniş hacimde bir örnek alırsan al bununla o öğrenme-çalışmasının değersizliğini gideremezsin. Bir öğrenme-çalışmasına başlamadan önce yanlılık hakkında düşünmek gerekir. Çükü veriyi elde ettikten sonra, yanlı bir örneği düzeltmenin hiç bir yolu yoktur. Bazen daha dar bir örnekle bir pilot testi yapmak yanlılığın önceden görülmesini ve daha geniş bir örnek almadan önce giderilmesini sağlamakta bize yardımcı olabilir. Aşağıda yanlılığın muhtelif kaynakları tanımlanmaktadır. Bu kaynaklar her ne kadar ayrı ayrı tanımlanıyor olasalar da esasen biribirileriyle çakışmaktadırlar ve hepsinde olmasa bile yanlılığın ortaya çıktığı pek çok durumda bunların bir kaçı birarada mevcuttur.

***Hanehalkı yanlılığı*:** Eğer bir örneğe herhangi bir hanehalkının yalnızca bir üyesi dahil edilirse, geniş hanehalklarının temsil edilebilirliği azalır. Bu mahzura yol açmamak için anketçiler bazen daha geniş hanehalklarına örnek hacminde daha fazla ağırlık verirler.

***Yanıt-yok yanlılığı*:** Bunun en güzel misallerinden biri mektup veya e-mail yoluyla yapılan anketlerdir. Bu tip anketlerin yanıtlanma yüzdesi çok düşüktür ve çoğu zaman yanıtların popülasyonun hangi kısmından geldiği bilinemez. Bazen bir yoklama için seçilen kişiler yanıt vermeyi reddedebilirler sadece veya ulaşılabilir durumda olamayabilirler veya iletişim kurmak çok

zor olabilir. Yanıt oranlarını yükseltmek için başvurulan yollar arasında farklı iletişim kanallarını kullanarak ısrarcı olmak, para ödemek veya diğer atraksiyonlara başvurmak sayılabilir. Aynı zamanda, kısa kolay anlaşılabilir yoklamalar genellikle daha yüksek yanıt oranlarına sahip olmaktadır.

***Kota örneklemesi yanlılığı*:** Eğer anketçilere örneğe girecek kişileri kendince seçme yetkisi verilirse bu tip yanlılık ortaya çıkabilir. Mesela, örnekte belirli bir yüzdeyle öğrencilere, belirli bir yüzdeyle kadınlara ya da belirli bir yüzdeyle mültecilere yer vermeye çalışmak gibi.

***Yanıt yanlılığı*:** Yoklama sorusunun bizzatihi kendisi yanıltıcı sonuçlara yol açabilir. İnsanlar çoğu zaman kişisel durumlarının ve görüşlerinin açıkça bilinmesinden bir takım rahatsızlıklar duyabilirler. Böyle konular hakkındaki sorulara verilen yanıtlar da yüz yüze yapılan veya anonim olmayan yazılı formdaki anketlerde çoğu zaman doğru olmamaktadır. Hastalar doktorun talimatlarına uyma konusunda yalan söyleyebilirler, diyet yapanlar bir zayıflatma programını ne kadar sıkı uyguladıkları konusunda dürüst olmayabilirler, öğrenciler sınavlara kaç saat çalıştıkları konusunda gerçeği söylemeyebilirler ve televizyon seyircileri belli programları seyrettiklerini kabul etmek istemeyebilirler.

***Seçim yanlılığı*:** Bunun en meşhur misali Amerika Birleşik Devletleri'nde 1936 yılında yapılan başkanlık seçimlerinde yaşanmıştır. *Literary Digest* adlı dergi-gazete yaptığı örnek yoklaması sonucunda Cumhuriyetçi Partinin adayı Alfred London'ın seçimi Demokrat Partinin adayı Franklin D. Roosvelt'e karşı az bir farkla kazanacağını seçimden önce ilan etmiştir. Ve yanılmıştır. *Literary Digest* seçmenleri otomobil ve telefon kayıtları üzerinden yoklamıştır. Fakat 1936 yılında otomobil ve telefon sahipleri azınlıktaydı ki bu zenginlerin çoğunluğu Cumhuriyetçi Parti'ye oy vermekteydi. Mektupla yapılan yoklamaya iki milyondan fazla yanıt gelmesine rağmen *Literary Digest*'in yaptığı yoklama, seçim yanlılığından dolayı doğru sonuca ulaşamamıştır.

***Hacim yanlılığı*:** Eğer bir haritaya dart atarak ülkenin illerini örneklemeye kalkarsanız, bu örnekleme coğrafi olarak alanı daha geniş olan illerin lehine yanlı olur. Bir hastahanenin önünde durarak, insanlarla anket yapmak, örneklemenin hastalar lehine yanlı olmasına yol açar. Bir torbaya tedavüldeki madeni paraları eşit oranlarda doldurup sonra göstermeden tek tek çektirirseniz, büyük boyutlu madeni paraların daha yüksek oranlarda çekildiğini görürsünüz.

***Eksik-kapsama yanlılığı*:** Yeterli temsil olmadığı zaman bu tip yanlılık vuku bulur. Mesela, telefon yoklamaları basitçe telefonu olmayan olanaklı denekleri gözardı etmektedir. Bunun diğer bir misali kolay-erişimli örneklemedir ki alışveriş merkezlerinde yapılan yoklamalar bu tiptedir. En yakınınızdaki kişileri örneğe dahil edersiniz. Bu anketler tüm popülasyonu temsil etme kabiliyetinden oldukça yoksun veri üretir. Kapıyı çalarak yapılan hanehalkı yoklamaları üniversite öğrencilerini, mahkumları ve evsizleri ıskalar.

***Gönüllü yanıt yanlılığı*:** Katılımın bireylerin insiyatifine bırakıldığı yoklamalar tipik olarak soruya duyarlı kişileri öne çıkarır. Mesela, radyo programlarında silah kontrolü, kürtaj, ve sosyal ayrımcılık dibi duyarlı meselelere dair yapılan telefon çağrılı yoklamalarda üretilen veri popülasyonun söz konusu mevzularda hangi orantıda lehtar veya aleyhtar olduğuna delil olamaz. Yine web sitelerinde yayınlanan online yoklamlar da gönüllü yanıt yanlılığının modern bir kaynağıdır.

***İfade yanlılığı*:** Nötr olmayan veya kötü ifade edilmiş sorular popülasyonu temsil etmeyen yanıtlara yol açabilir. Böylesi bir yanlılıktan kaçınmak için yönlendirici soruları kullanmamak gerekir. Soruları açık ve nispeten kısa yazmak her zaman en iyisidir. Aynı zamanda, soruların sıralamasına da dikkat etmek gerekir ki denekleri belli yanıtlara sevketmesin.

Şunu tekrar ifade etmekte yarar var ki yukarıda sınıflandırılan yanlılıklar bazı durumlarda çakışabilecektir. Mesela, nötr-olmayan bir sorunun hem *yanıt* yanlılığına hem de *ifade* yanlılığına sahip olduğu söylenebilir. *Seçim* yanlılığı ve *eksik-kapsama* yanlılığı çoğu durumda beraber ortaya çıkacaktır. Hakeza, g*önüllü yanıt* yanlılığı *yanıt-yok* yanlılığı ile açıkça bağlantılıdır.

Diğer Örnekleme Metotları

Zamandan ve maliyetten tasarruf ettiren modifikasyonlar, basit rastgele örneklemeden başka örnekleme prosedürlerinin icra edilmesinde çoğunlukla kullanılır. Bunların başlıcaları: *Sistematik örnekleme, tabakalı örnekleme, öbek örnekleme* ve *aşamalı örnekleme*.

Sistematik örnekleme popülasyonun herhangi bir (mesela alfabetik) tertipte listelenmesini, rastgele bir başlangıç noktası belirlenmesini ve sonra her onuncu (veya yirminci, veya yüzüncü, veya bininci, vb.) kişinin listeden seçilmesini ihtiva eder. Listenin orjinal tertibi bir şekilde ele alınan değişkenler ile ilgili olmadığı sürece bu metot makul bir örnek üretir.

Tabakalı örneklemede popülasyon *tabaka* adı verilen kendi içerisinde homojen gruplara bölünür ve tüm tabakalardan bireylerin rastgele örnekleri seçilir. Mesela, bir insan popülasyonunu yaşa, cinsiyete, gelir düzeyine, veya mesleğe göre tabakaladıktan sonra her tabakadan insanların bir örneğini çekebiliriz. Dikkat etmelisin ki belirli bir tabakadaki tüm bireyler müşterek bir özelliğe sahiplerdir. Tabakalar içerisinde rastgele örnekleme *orantısaldır*. Yani, her tabakadan çekilen rastgele örneğin hacmi o tabakanın popülasyondaki orantısına bağlıdır.

Öbek örneklemede popülasyon *öbek* adı verilen kendi içerisinde heterojen gruplara bölünür ve tüm öbekler içerisinden öbeklerin bir rastgele örneği seçilir. Mesela, lise öğrencilerini yoklamak için belirli bir sayıda sınıfı rastgele seçip öğrenme-çalışmamıza dahil edebiliriz. Dikkat etmelisin ki her öbek tüm popülasyonun bir temsili olmalıdır.

Aşamalı örnekleme iki veya daha çok adımı ihtiva eden bir prosedürü tanımlar. Bu adımlardan her biri muhtelif örnekleme tekniklerinden herhangi birini ihtiva edebilir. Mesela Amerika Birleşik Devletlerinde gelenekselleşmiş olan Gallup yoklamaları çoğunlukla şu prosedürü takip eder: Önce ülke çapında bölgeler rastgele seçilir, sonra her bölge içerisinde semtler rastgele seçilir ve en sonunda her semt içerisinde haneler rastgele seçilir.

Misal 4 Ötenazi hakkındaki görüşlerini belirlemek ereği ile 5000 kayıtlı öğrencisi bulanan bir liseden 100 öğrencilik bir örnek çekilecektir. Bir metot, her öğrencinin adını ve soyadını birer kağıt parçasına yazıp, kağıt parçalarını bir kutuya doldurup bu kutudan 100 kağıt parçasını çekmek olurdu. Gel gör ki, kutuya doldurulan kağıt parçalarının ne kadar iyi karıştırılabileceği bir problem yaratacaktır. Mesela öğrencilerin adları ve soyadları kağıt parçalarına şubelerine göre bir sırada yazılacak ve kutuya atılacak ise, bunun sonucunda aynı şubenin öğrencilerinin ad ve soyadlarının yazılı olduğu kağıt parçaları kutu içerisinde birarada bulunacaktır. Bu durum çıkıma nasıl tesir edecektir? Diğer bir metot her bir öğrenciye 1 den 5000 e kadar birer numara atayıp, sonra bir rastgele sayılar tablosu kullanarak dört haneli sayılar çekmektir (tekrar edenleri ve 5000 den büyük olanları atmak suretiyle). Her iki durumda da basit rastgele örnekleme yapılmaya çalışılmaktadır. Acaba başka hangi alternatif prosedürler önerilebilir?

Cevap: Öğrencilerin bir listesinden, yoklamacı her ellinci öğrencinin adını ve soyadını kaydedebilir (sistematik örnekleme). Her sınıftaki öğrencilerin belli müşterek özelliklere sahip olabileceklerini düşünerek, yoklamacı birinci, ikinci, üçüncü ve dördüncü sınıfların ad ve soyad listelerinden 25 er öğrenciyi rastgele seçim metotlarından birini kullanarak seçebilir (tabakalı örnekleme). Yoklamacı, şubeleri sınıflarına göre ayırır; rastgele beş birinci sınıf şubesi, beş ikinci sınıf şubesi, beş üçüncü sınıf şubesi ve beş dördüncü sınıf şubesi seçer (öbek örnekleme); ve sonra her şubeden beş öğrenciyi rastgele seçer (aşamalı örnekleme). Yoklamacı oğlan ve kız öğrencilerin okulun tüm öğrencilerine orantılarına göre ayrı ayrı rastgele örneklerini seçebilir (orantısal tabakalı örnekleme).

Şuna dikkatini çekmek isterim ki bu misaldeki alternatif prosedürlerden hiç biri bir *basit rastgele örnek* değildir. Çünkü, 100 hacimlik her olanaklı örnek eşit seçilme şansına sahip değildir.

Olasılıksız örnekleme. *Elverişlilik örneklemesi, kota örneklemesi, uzman örneklemesi* ve *kartopu örneklemesi* popülasyonun bazı ünitelerinin seçilme şansının sıfır olduğu veya isabetlice belirlenemediği diğer örnekleme teknikleridir. Seçim rastgele yapılmadığı için, olasılıksız örneklemelerde örnekleme hatalarının kestirimi mümkün olmaz ve dolayısıyla örnekleme yanlılığı kaçınılmaz olabilir. Bu sebeple, böyle bir örneğin taşıdığı enformasyon popülasyona olasılıksal bir prosedürle genelleştirilemez.

Deneylerin Planlaması ve Yürütülmesi

Deneyleri münasipçe planlamaya ve yürütmeye dair başlıca bir kaç prensip vardır. Bunlardan birincisi olanaklı karışan değişkenleri, genellikle karşılaştırma metodunu kullanarak kontrol altına almaktır. İkincisi, denekleri muamele gruplarına şansa bağlı atamaktır. Üçüncüsü, çıkımların barındıracağı doğal değişimi azaltmak için mümkün olduğunca çok sayıda denek kullanmaktır.

Deneyler, Gözlemsel Öğrenme-çalışmaları ve Yoklamalar

Bir deneyde, bir değişikliği veya muameleyi uygularız ve sonucu veya yanıtı ölçeriz. Bir gözlemsel öğrenme-çalışmasında, kendiliğinden meydana gelen bir şeyi, mevcudiyetimiz vasıtasıyla herhangi bir değişikliğe neden olmaksızın sadece gözleriz ve ölçeriz. Bir örnek yoklaması bir gözlemsel öğrenme-çalışamsıdır ki münasipçe seçilmiş bir örneği nazar-ı dikkate almak suretiyle bir popülasyonun tümü hakkında neticeye varırız. Bir deney çoğu zaman bir nedensel ilişkiyi ortaya çıkarır, oysa bir gözlemsel öğrenme-çalışması ancak bir çağrışımın varlığını gösterebilir.

Nedensellik, bir değişkende meydana gelen değişmenin başka bir değişkende illa ki bir değişmeye yol açtığı ilişkileri haizdir. Nedensellik için üç koşulun sağlanması gerekir: kovaryasyon (karşılıklı-değişim), zamanda öncelik ve "üçüncü değişkenleri" kontrol altına almak. Sonki koşul gözlenen nedensel ilişkiler için alternatif açıklamaların mevcudiyetine delalet eder. Değişkenler arasındaki kovaryasyon (yani karşılıklı-değişim) bir nedensel etkiyi gösterdiği zaman, eğer bu kovaryasyon aslında altta-yatan bir müşterek nedenin (*gizli değişken*) sonucu ise, *suni ilişkiler* ortaya çıkar. Başka hallerde ise, yine kontrol altına alınmayan bir değişkenin yanıta etkisi açıklayıcı değişkeninkiyle *karışabilir*. Araştırma terminolojisinde karışan değikenler ve gizli değişkenler aynı zamanda *harici değişenlerdir*. Elbette her harici değişkenin yanıt değişkeni üzerinde bir karışan değişken veya bir gizli değişken etkisi olmak mecburiyeti yoktur.

Misal 5 Günlük kalsiyum takviyelerinin kadınlarda kemik kütlesini artırıp artırmadığını belirlemek için bir öğrenme-çalışması dizayn edilecektir. Bir gözlemsel öğrenme-çalışması nasıl icra edilebilir? Ya bir deney? Peki burada hangisi daha mütenasiptir?

Cevap: Bir gözlemsel öğrenme-çalışması eczanelerden veya kalsiyum takviyelerinin satıldığı dükkanlardan bu ürünleri satın aldığına şahit olunan kadınlarla mülakat yapıldıktan sonra bunların kemik kütlelerini ölçmek olabilir. Ya da belki de belirli bir periyotta hastahaneye yatan tüm hastalarla yapılacak mülakatlar sonucunda kalsiyum takviyesi alanlar tespit edilip bunların kemik kütlesi ölçülebilir. Kalsiyum takviyesi aldığı bilinen kadınların kemik kütlesi ölçümleri daha sonra kalsiyum takviyesi almadığı bilinen kadınların kemik kütlesi ölçümleri ile karşılaştırılabilir.

Bir deney, belirli sayıda denek üzerinde şöyle icra edilebilir: deneklerin şansa bağlı olarak seçilen yarısına kalsiyum desteği hapları verilirken diğer yarısına aynı görünümde plasebo hapları verilir ve her grubun kemik kütlesi muameleden önce ve sonra ölçülürek farkı kaydedilir.

Deneysel yaklaşım burada daha mütenasiptir. Gözlemsel öğrenme-çalışması ile kalsiyum takviyesi alan kadınlar ile kalsiyum takviyesi almayan kadınlar arasında kemik kütlesi bakımından ortaya çıkabilecek farklığın pek çok açıklaması olabilecektir. Mesela, kendiliğinden kalsiyum takviyesi almış kadınlar genel manada kendilerine daha iyi bakan kadınlar olabilir elbette ve böylece daha yüksek kemik kütlesine sahip olmalarının başka nedenleri olabilir. Deney deneklerin rastgele seçilen yarısına kalsiyum takviyesi vererek gizli değişkenleri kontrol altına almaya çalışır.

Misal 6 Uzun boylu kişilerin kısa boylu kişilere nazaran ömürlerini (yani boy uzunluğu ve ömür arasındaki ilişkiyi) incelemek için bir öğrenme-çalışması dizayn edilecektir. Hangisi daha mütenasiptir, bir gözlemsel-öğrenme çalışması mı yoksa bir deney mi?

Cevap: Bir gözlemsel öğrenme-çalışması eğer medikal kayıtlara dayanarak kişilerin ölüm anındaki boy uzunluklarını ve yaşlarını inceleyecek ise doğru görünüyor. Bir deney eğer rastgele seçilecek kişileri kısa veya uzun boylu yaptıktan sonra ölüm anlarında yaşlarını kaydedecek ise çığır açıcı olurdu (ve elbette manasız).

Açıklayıcı Değişken ve Yanıt Değişkeni

Deneyler *faktör* adı da verilen *açıklayıcı değişkenleri* ihtiva ederler ki bunlar *yanıt değişkeni* üzerinde bir etkisi olduğuna inanılan değişkenlerdir. Bir grup gözlem ünitesi, açıklayıcı değişkenin bir düzeyi ile muamele edilir ve yanıt değişkeninin çıkımı ölçülür. Buna göre bir faktörün her bir düzeyi bir muameledir. Aynı deney dizaynı içerisinde birden çok sayıda faktör yer aldığında, bu faktörlerin düzeylerinin her kombinasyonu yine bir muameledir.

Misal 7 Ders çalışma süresinin istatistik dersindeki başarıya etkisini belirlemek için, öğrenciler üç gruba ayrılabilir. Bir grup haftada 4 saat ders çalıştırılır, ikinci grup haftada 2 saat ders çalıştırılır ve üçüncü grup hiç ders çalıştırılımaz. Açıklayıcı değişken hangisidir, yanıt değişkeni hangisidir ve düzeyler nelerdir?

Cevap: Açıklayıcı değişken, haftalık ders çalışma süresidir ve üç düzeyi vardır: 4 saat, 2 saat ve 0 saat. Yanıt değişkeni belirtilmemiştir fakat yıl sonu başarı puanı, final sınavı puanı veya başka bir sınav puanı olabilir.

Misal 7'de üç muamele mevcut ki bu üç muamele bir faktörün üç düzeyine karşılık gelmektedir. Farzedelim ki öğrenciler aynı zamanda A ve B şubeleri olarak ta bölünmüş olsunlar. Şimdi birisi üç düzeyli ve birisi de iki düzeyli olmak üzere iki faktör ve toplamda ilaveten altı muamele (A şubesi ve 4 saat çalışma, A şubesi ve 2 saat çalışma, A şubesi ve 0 saat çalışma, B şubesi ve 4 saat çalışma, B şubesi ve 2 saat çalışma, B şubesi ve 0 saat çalışma) daha mevcut olacaktır.

Karışma

Bir etkiye hangi değişkenin neden olduğuna dair belirsizlik hasıl olduğu zaman, değişkenlerin *karışan* olduğunu söyleriz. Mesela, farzedelim ki iki çeşit gübre farklı sulama miktarları gerektiriyor olsun. Tek faktörlü bir deneyde, bitki gelişmesinde gözlenen farklılıkların gerçek nedeninin gübrelerdeki farklılıktan mı yoksa sulama miktarındaki farklıktan mı kaynaklandığını belirlemek zor olabilir. Bu karışmayı kontrol altına alabilmek için, her bir gübre çeşidini daha çok sayıda test parseline uygulayarak bu parsellerin bir yarısını bir miktarda diğer yarısını diğer miktarda sulayabiliriz, böylece gübre çeşidi ve sulama miktarının bitki büyümesine etkilerini biribirine karışmadan belirleyebiliriz.

Genel manada, karışmayı engellemek için açıklayıcı değişken dışında yanıt değişkenine etkisi olan tüm faktörleri sabit bir düzeyde tutmak esastır. Ancak bu pratikte büyük ölçüde olanaksız olduğundan, rastgeleleştirme prensibi deney dizaynında vazgeçilmezdir.

Gizli Değişken

Bir *gizli değişken* öyle bir değişkendir ki üzerlerinde ayrı ayrı etkili olduğu başka iki değişkenin gerçekte mevcut olmayan bir neden-ve-etki ilişkisine sahip oldukları izlenimini yaratır. Mesela, ayakkabı numarası büyük olan ilkokul öğrencilerinin okuma düzeylerinin daha iyi olduğu görülür. Mamafih, bir gizli değikenin mevcudiyeti söz konusudur ve bu gizli değişken yaştır ki diğer her iki değişken üzerinde de etkilidir. Yani yaşça daha büyük öğrenciler yaşça daha küçük öğrencilere göre daha büyük numaralı ayyakabı giyme eğilimindelerdir ve yaşça daha büyük öğrenciler aynı zamanda daha iyi okuma düzeylerine sahip olma eğilimindelerdir. Büyük numaralı ayakkabı giymek okuma düzeyini daha iyileştirmeyecektir. *Gizli değişken* ibaresini kullanmanın dışında, bu durumun bir *müşterek yanıt* olduğu da söylenebilir. Yani hem ayakkabı numarasındaki hem de okuma düzeyindeki değişikliklerin nedeni yaştaki değişikliktir.

Kontrol Grubu

Bir deneyde gözlem ünitelerinin bir grubu belirli bir muameleye maruz bırakılırken diğer bir grup doğal akışında bırakılır. Hiç bir muameleye maruz bırakılmayan bu ikinci gruba *kontrol grubu* denir. Deney muamele grubundaki yanıtları kontrol grubundaki yanıtlarla karşılaştırır. Denekleri muamele

ve kontrol gruplarına rastgele atamak karışmanın ve gizli değişkenin yol açtığı problemlerin giderilmesine yardımcı olur. Böylece, deney yaptığımız zaman bu tip problemlerin kontrol altına alınması, gözlemsel öğrenme-çalışması yaptığımız zamana göre daha kolaydır.

Plasebo etkisi ve Köreltme

Herhangi bir muamele algısına insanların psikolojik yanıt verdiği yadsınamaz bir gerçektir. Buna *plasebo etkisi* adı verilmektedir. Mesela, ameliyattan sonra sadece şeker hapı verilir, fakat hastaya bunun çok kuvvetli bir ağrı kesici olduğu söylenirse, pek çok hasta acılarında ani bir hafifleme hisseder. Pek çok öğrenme-çalışmasında, denekler bilinçli olarak veya bir bilinçaltı durumu olarak araştırmacının kanıtlamak istediği şeyi kanıtlamasına yardımcı olmak isterler. Böyle olunca, herhangi bir deneyde dikkat çekici yanıtlar ortaya çıktığı zaman, gerçek fiziksel yanıta psikolojik plasebo etkisinin neden olup olmadığından endişe edilir.

Denekler veya yanıtları değerlendiren kişi hangi deneğin muamelelerden hangisine (veya plaseboya) maruz kaldığını bilmediği zaman *köreltme* yapılmış olur. Gizli yanlılığı minimize etmeye yardım etmek için, en iyisi deneklerin hangi muameleye maruz bırakıldıklarını bilmemeleridir. Buna *tekli-köreltme* denir. Hangi deneğin hangi muameleye maruz bırakıldığını ne denkler ne de yanıtları değerlendiren kişiler bilmediği zaman buna da *çifte-köreltme* denir.

Plasebolar ve köreltme, deneylerde önemli olsalar da her zaman uygulanabilir olamayabilirler. Bunlarsız da pek çok deney yürütülebilmektedir.

Misal 8 Bir öğrenme-çalışmasında E vitamini ve beta karotenin kalp krizi oranlarına etkileri test edilmek isteniyor. Deney nasıl kurulmalıdır?

Cevap: Rastgeleleştirme kullanılarak, denekler dört gruba ayrılmalıdır: yalnızca E viatamini alanlar, yalnızca beta karoten alanlar, hem E viatmini hem de beta karoten alanlar ve ne E vitamini ne de beta karoten alanlar. Deneklerin muamele gruplarına ve kontrol grubuna rastgele atamaları için bir rastgele sayılar tablosu kullanılabilir. Eğer öğrenme-çalışması çifte-köreltilirse, yani ne denekler ne de onları muayene eden doktorlar hangi hastanın hangi muameleye maruz bırakıldığını bilmez ise daha manalı sonuçlar elde edilebilir.

Muameleler ve Deney Üniteleri

Bir deneyin üzerinde icra edildiği nesnelere *deney üniteleri* adı verilir, bunlar aynı zamanda *denekler* olarak ta adlandırılır. Deney üniteleri ve denekler genel olarak iki gruba ayrılır. Bir grup bir *muameleye* maruz bırakılır ki buna *muamele grubu* denir. Muamele grubunda kaydedilen yanıt herhangi bir muameleye maruz bırakılmayan *kontrol grubunda* kaydedilen yanıt ile karşılaştırılır.

Rastgeleleştirme

Gizli değişkenlerin etkisini ve karışmayı minimize etmek için *rastgeleleştirmeyi* kullanmak önemlidir. Yani hangi deneğin hangi gruba dahil edileceğine karar verirken şansa bağlı kalmak gerekir. İki grup arasında deney ünitelerinin sahip oldukları özellikler bakımından sistematik bir denge sağlamaya çalışmak yeterli değildir. Muamele grubunda ve kontrol grubunda mesela kadınların, müslümanların, bekarların, gençlerin, gözlüklülerin, öğrencilerin v.b. kendimizce eşit sayısını ayarlamaya çalışmak her ne kadar mantıklı görünüyor olsa bile, bu metot iyi iş görmez. Çünkü her zaman dikkate almayı aklımıza getiremeyeceğimiz başka değişkenler mevcuttur ki ancak deney sonuçları gelmeye başladığında bunların farkına varabiliriz. En iyi metot bir kompüterden, içine adların doldurulduğu bir kutudan, veya rastgele sayılar tablosundan yararlanmak suretiyle rastgeleleştirmeyi kullanmaktır.

Dikkat etmelisin ki *rastgeleleştirme* deney dizaynı bağlamında genellikle elimizdeki deneklerin muamelelere nasıl atanacağını kastetmektedir, popülasyonun tümü içerisinden bir denekler grubunun nasıl seçileceğini değil. Bir deneyin gayesi farklı muamelelerin farklı yanıtlara yol açıp açmadığını

görmektir. Öyle ise, denekleri muamelelere rastgele atarız ki değişkenliğin meçhul kaynaklarını dengeleyebilelim. Muamelelere rastgele atama, özellikle medikal deneylerde, ilaç deneylerinde olduğu gibi eğer denekler rastgele seçilmemiş ise çok daha önemlidir. Öğrenme-çalışmasının bulgularını genelleştirmek başka bir problemdir ki deneklerin başlangıçta nasıl seçildiğine bağlıdır.

Tamamen Rastgeleleştirilmiş Dizayn: İki Muamele

Rastgeleliği kullanarak iki muameleyi karşılaştırmak çoğu zaman favori dizayn seçimidir. Deney ünitelerini muamele gruplarına rastgele atarız. Böylece, her bir deney ünitesinin her bir muamele grubuna gitme şansı eşit olur. Her bir deney ünitesi diğer deney ünitelerinden bağımsız olarak atanır. Buna *tamamen rastgeleleştirilmiş dizayn* deriz. Bağımsız rastgeleleştirmelerin sayısının fazla olması muamele grupları veya kontrol grubu arasındaki karşılaştırmaların adil olmasını temin eder, çünkü gruplar arasında deney ünitelerinin değişimleri benzer olacaktır. Gizli değişkenlerin yüksek değerlerine ve düşük değerlerine sahip deney üniteleri tüm muamele geruplarında ve kontrol grubunda yakın orantılarda yer alacaklardır.

Misal 9 Bilekte bir basınç noktası vardır ki bazı doktorlar belli medikal operasyonları müteakiben yaşanan mide bulantısını kontrol altına almakta bu noktanın kullanılabileceğine inanmaktadırlar. Fikir, küçük bir mermer parçasının yerleştirilmiş olduğu bir bilekliği hastanın bileğine sıkıca bağlamaktır. Öyle ki bilekliğe yerleştirilmiş mermer tam olarak basınç noktasının üzerine gelmiş olsun. Operasyon geçirmiş 50 hasta üzerinde bir deneyin nasıl yürütülebileceğini betimle.

Cevap: Her hastaya 1'den 50'ye kadar olan numaralardan birini ata. Bir rastgele sayılar tablosundan bir defada iki hene olmak üzere, tekrarları, 00 sayısını ve 50'den büyük sayıları gözardı etmek suretiyle 25 sayıyı oku. Çektiğin sayılara tekabül eden hastalara, mermer parçası basınç noktasının tam üzerine gelecek şekilde bileklikleri bağla. Geriye kalan 25 hastayada üzzerinde mermer parçası olan bileklikleri bağla, fakat bunlarda mermer parçası basınç noktasının üzerinde değil de başka bir noktada olsun. Bir soruşturmacı bu 50 hastanın tümünü belirlenen zaman periyotlarında telefonla arayarak yaşadıkları mide bulantısının derecesini belirlesin. Ne hastalar ne de telefon aramalarını yapan soruşturmacı hangi hastanın bilekliğindeki mermer parçasının doğru basınç noktası üzerine başlanmış olduğunu bilmesin.

Misal 10 Bir kimyasal gübre şirketi yeni geliştirdikleri ürünün sebzeler üzerinde üstün bir etkisi olup olamdığını test etmek istiyor. Büyük bir tarlayı küçük parsellere böldükten sonra, deney nasıl ilerleyebilir?

Cevap: Eğer şirketin gübre için bir doz uygulaması önerisi varsa, parsellerin yarısı rastgele seçilebilir ve gübrenin reçetelendirilmiş dozajı uygulanabilir. Parsellerin rastgele seçiminde bir rastgele sayılar tablosu kullanılabilir. Parsellerin bu rastgele seçimi, ne gübre verilen bitkilerin ne de gübre verilmeyen bitkilerin toprak tipi, güneşlenme, sulanma, v.b. bakımından daha iyi parsellere kasten ekilmediğini garanti eder. Bitkilerin ekimini ve sulamasını yapacak işçilerden dolayı ortaya çıkabilecek yanlılıktan kaçınmak için, işçilerin hangi parsellere gübre uygulandığını bilmemeleri gerekir. Bunu sağlamak için her parselin yanına, yarısı aktif gübre dozunu kapsayan ve diğer yarısı kapsamayan benzer kutuların içerisinde benzer görünümlü ve benzer kokulu maddelerin bir karışımı dağıtılabilir. Nihayetinde, eğer sebzeler nicelik ve büyüklük olarak standart ölçekler ile değerlendirilecek ise ölçümler yanlılığı elbette daha az barındıracaktır. Mamafih, eğer sebzeler mesela tat veya koku gibi nitel bir değerlendirmeye tabii tutulacaksa, bu defa değerlendirmecilerin de hangi bitkilerin gübrelemeye maruz bırakıldığını ve hangilerinin bırakılmadığını bilmemesi gereklidir.

Eğer araştırmacılar aynı zamanda düzeyi de nazar-ı dikkate almak isterlerse, yani gübrenin değişik miktarlarını, bu durumda rastgeleleştirme daha fazla gruplandırmalar için yapılmalıdır. Mesela, gübrenin dört düzeyi 60 parsel üzerinde test edilecekse, bir rastgele sayılar tablosunun 01-60 aralığındaki ilk 12 adet biribirinden farklı sayıya tekabül eden parsellere gübrenin birinci düzeyi uygulanabilir, sonraki 12'sine ikinci düzeyi, sonraki 12'sine üçüncü düzeyi, sonraki 12'sine dördüncü düzeyi ve son 12'sine ise "plasebo" muamelesi uygulanabilir.

Rastgeleleştirilmiş Eşli Karşılaştırma Dizaynı

İki muamele eşli deneklerin yanıtları üzerinden karşılaştırılabilir. Eşli deneklerin birisi bir muameleye maruz bırakılırken öteki diğer bir muameleye (veya plaseboya) maruz bırakılır. Çoğu zaman eşli denekler aslında tek bir denektir ki her iki muamele bu deneğe farklı zamanlarda uygulanır.

Misal 11 Meşhur Pepsi ve Coca Cola testleri deneklere her bir içkinin örneklerinin tadını karşılaştırtır. Böyle bir eşli test nasıl dizayn edilir?

Cevap: Can alıcı nokta, böyle bir testte, testin köreltilmiş olması gerektiğidir. Yani, denekler içtikleri bardağın hangisinde Pepsi hangisinde Coca Cola olduğunu bilmemelidirler. Mesela üçgen işaretli bardaklar Pepsi örneği, kare işaretli bardaklar Coca Cola örneği olabilir. Dahası, gizli yanlılıktan kaçınmak için, deneklerin hangi bardağı önce içeceklerine de şansa bağlı olarak karar vermek gerekir. Mesela, sırası gelen deneğin hangi bardağı önce içeceğine bir madeni para atışıyla karar verilebilir. Yazı gelirse denek önce üçgen işaretli bardağı, tura gelirse önce kare işaretli bardağı içebilir. (**Not:** Her ne kadar deneklere bir içki içiriliyor ve bir rastgeleleştirme yapılıyor olsa bile, katı bir teorik istatistik-bilimi penceresinden bakıldığı zaman bu durumu bir deneyden ziyade, popülasyon orantısının kestirimini hedefleyen bir örnek yoklaması olarak mülahaza etmek daha doğru olabilir.)

Misal 12 Sarhoş sürücülerin neden olduğu trafik kazalarına ait fotoğrafları görmek, sarhoş sürücülerin cezalandırılması hakkında, kişilerin görüşlerine tesir eder mi? Bir karşılaştırma testi nasıl dizayn edilir?

Cevap: Fotoğraflar gösterilmeden önce ve gösterildikten sonra, deneklere sarhoş sürücülerin cezalandırılması hakkında sorular sorulur ve yanıtlardaki değişiklikler kaydedilir. Eğer bu kadarla kalırsa, bu zayıf bir dizayn olur. Çünkü bir kontrol grubu yok, rastgeleleştirme kullanılmıyor ve denekler fotoğrafları gördükten sonra, kendilerinden beklendiği düşüncesiyle yanıtlarını pekala değiştirebilirler.

Daha iyi bir dizayn, rastgeleştirme kullanarak denekleri iki gruba bölmek, bir yarısına yalnızca soruları sorarken diğer yarısına önce fotoğrafları gösterip sonrasında soruları sormak olurdu.

Bir diğer olanaklı dizayn birbirinin ikizi deneklerden oluşturulmuş grupları kullanmaktır. İkizlerden biri rastgele seçilerek fotoğrafları göstermeksizin soruları yanıtlaması istenirken diğerine önce fotoğraflar gösterilir ve sonra soruları yanıtlaması istenir. Her ikiz çiftinden birini rastgele seçmek için madeni para atışı yapılabilir veya bir rastgele sayılar tablosundaki tek haneli sayılar tek veya çift olarak kullanılabilir, ikinci durumda sıfırı göz ardı etmek gerekir elbette. İkizlerin her bir kümesinden gelen yanıtlar karşılaştırılır. Bu bir eşli karşılaştırma testidir ki kalıtım ve çevre gibi arkaplansal fenomenlere dair gizli değişkenleri minimize etmeye yardımcı olur.

Tekrarlama

Bir karşılaştırma testinde farklar gözlendiği zaman, araştırmacı bu farkların *istatistiksel olarak anlamlı* olup olmadığına veya doğal değişimle açıklanabilip açıklanabilmeyeceğine karar vermek durumundadır. Bu merhalede önem arzeden bir telakki örnek hacmidir. Örnek hacmi ne kadar geniş olursa, gözlem o kadar anlamlıdır. Buna *tekrarlama* prensibi denir. Yani, muamele deneklerin öyle yeterli bir sayısıyla tekrarlanmalıdır ki gerçek yanıt farkları daha aşikar olsun.

Bloklama

Örnekleme dizaynında tabakalama yaparken nasıl popülasyon tabaka adı verilen kendi içerisinde homojen gruplara bölünüyordu ise, deney dizaynında *bloklama* da denekleri önce blok adı verilen kendi içerisinde homojen alt gruplara böler. Bloklamayı her blokta ayrı bir deney yürütmek gibi düşünebilirsin. Bu teknik belli gizli değişkenleri, yanıta etkilerini ortaya çıkarmak suretiyle kontrol altına almaya ve neticelerin daha spesifik olmasına yardım eder. Eşli karşılaştırma dizaynı, bloklamanın özel bir durumudur ki her çift (veya eğer denekler kendi kontrolleri olarak yer alıyor ise,

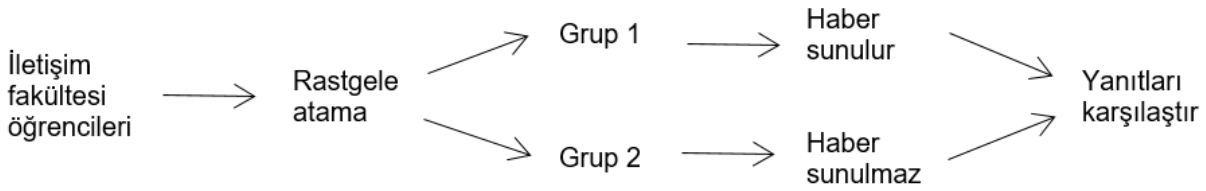
her denek) bir blok olarak mülahaza edilebilir.

Benzer şekilde, erkekleri ve kadınları, yüksek IQ'luları ve düşük IQ'luları, farklı meslek sahiplerini, farklı eğitim düzeyine sahip olanları, vb. ayrı ayrı analiz etmek için bloklama kullanılabilir.

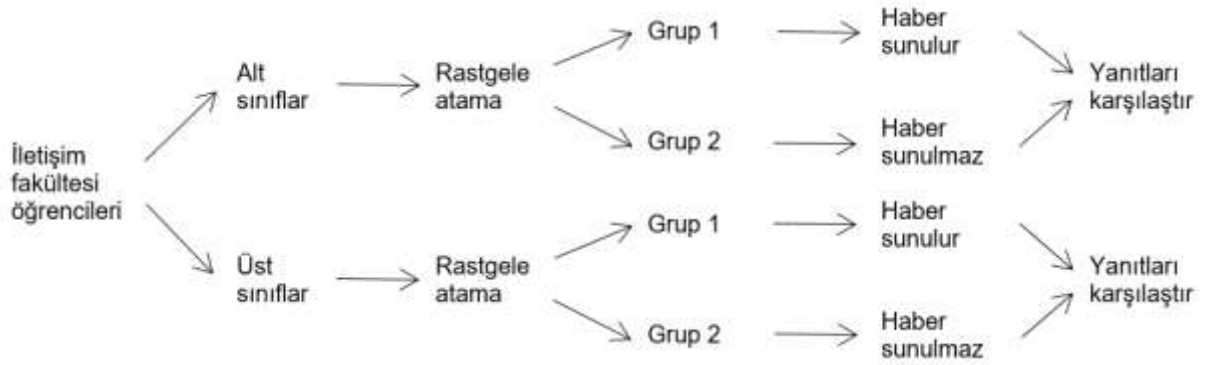
İstatistik-bilimi öğrenirken terminolojinin münasip kullanımı çok önemlidir. Mesela, deneylerin dili gözlemsel öğrenme-çalışmalarının dilinden farklıdır. *Bloklama* ile *tabakalamayı* sakın biribirine karıştırma.

Misal 13 Son yıllarda, iletişim fakültelerinde okuyan öğrencilerin henüz okullarını bitirmeden mesleğe atılmaları moda oldu. Meslekte çalışanlara ödenen maaşlar hakkında çıkan medya haberlerinin bu kararların alınmasına etkisi olup olmadığını değerlendirmek için bir öğrenme-çalışması icra ediliyor. Denekleri iki gruba ayırmak için rastgeleleştirme kullanılıyor ve gruplardan birindeki deneklere hazırlanan soruları yanıtlamalarını istemeden önce medya haberleri sunuluyor. Pekala, bu deneyin dizaynına bir blok dizaynı nasıl adapte edilebilir?

Cevap: Önce denekler alt sınıftakiler ve üst sınıftakiler olmak üzere iki bloğa ayrılabilirler. Sonra, her blok içerisinde, soruları yanıtlamaları istenmeden önce medya haberi sunulacak denekler rastgeleleştirme kullanılarak atanır. Bu dizayn ile, medya haberine maruz kalmanın alt sınıflar üzerindeki etkisi üst sınıflar üzerindeki etkisinden ayırdedilebilir.



Figür 1 Tamamen rastgeleleştirilmiş dizayn.



Figür 2 Blok dizaynı.

Sonuçların Genelleştirilebilirliği

Deneylerin bir başat amacı, sonuçları daha geniş popülasyonlara *genelleştirebilmektir*. Çoğu zaman bir deneyin çeşitli ortamlarda yinelenmesi gerekir. Mesela, bir televizyon reklamının Doğu Anadolu Bölgesin'de bir lisenin öğrencileri üzerindeki etkisini, aynı reklamın İstanbul'da yaşayan emekliler üzerindeki etkisine genelleştirmek zordur.

Genellikle, karşılaştırma ve rastgeleleştirme önemlidir, köreltme bazen çok önemlidir, mümkün olduğunca gizli yanlılıktan kaçınmaya dikkat etmek her zaman iyi-dizayn-edilmiş bir deneyin göstergesidir. Mamafih, şunu da unutmamak gerekir ki insan deneklerin veya diğer canlı deneklerin testlerin sonucuna bir takım öngörülemez etkileri vakidir. Özellikle insanlar üzerinde yapılacak testler ve deneyler onların doğal hallerini bozabilir ve bu da yapay yanıtlara yol açabilir.

İçsel Geçerlilik ve Dışsal Geçerlilik

Aralarındaki yegane manalı farkın uyguladığımız muameleden kaynaklandığını gösterecek şekilde, grupları oluşturmayı başarabildiğimiz ölçüde, deneyimizin *içsel geçerliliğe* sahip olduğunu söyleriz. Bu durumda, gruplar arasındaki içsel karşılaştırma geçerli bir karşılaştırmadır. İçsel geçerlilik ve rastgele atama arasındaki münasebete dikkatini çekmek isterim.

Eğer deney ünitelerimiz, genelleştirme yapmak istediğimiz popülasyonun bir temsilcisi ise, deneyimizin *dışsal geçerliliğe* sahip olduğunu söyleriz. Bu durumda, örnekten elde ettiğimiz bulguları popülasyona genelleştirmemiz geçerlidir. Dışsal geçerlilik ve rastgele seçim arasındaki münasebete dikkatini çekmek isterim.

Yarı-deneysel öğrenme-çalışmaları

Gözlemsel öğrenme-çalışmalarının bazılarına *yarı-deneysel öğrenme-çalışmaları* da denir. Bazı bilimsel araştırma öğrenme-çalışmaları vardır ki bilimsel araştırmacı açıklayıcı değişkeni gönlünce manipüle edemez (değiştiremez). Böyle bir manipülasyon bu tür değişkenlerin doğasına aykırıdır. Veya deney ünitelerinin muamele gruplarına rastgele ataması gerçekleştirilemez. Mesela projesinde yeni bir öğretim metodunun eskisinden daha etkin olup olmadığı sorusuna cevap arayan bir araştırmacı, deneyini gerçekleştireceği okuldaki mevcut iki sınıftan birinde yeni metodu diğerinde eski metodu deneyebilir. Bu durumda öğrencilerin (yani deneklerin veya deney ünitelerinin) öğretim metotlarına (yani muamele gruplarına) ataması rastgele yapılmamış olur.

Yarı-deneysel öğrenme-çalışmaları, neden-ve-etki ilişkilerini açığa çıkarmada arızalıdır. Çünkü içsel geçerlilik yeterince sağlanamaz. Yani açıklayıcı değişkenin, yanıt değişkeninde yarattığı etkinin miktarı bu türden bir öğrenme-çalışması ile tam belirlenemez. Ancak, koşulların zorlaması ile, yine de bu türden araştırmalara çok sık başvurulur.

Yine, örnek yoklamalarında da anket yapacağımız kişileri ne açıklayıcı değişken bakımından manipüle edebiliriz ve ne de yanıt değişkeni bakımından aynı koşullar altında ölçebiliriz. Bu durumda araştırmacının yapabileceği yegane kontrol, anket sorusunu her deneğe aynı hassasiyetle sorabilmekten ibarettir. Mesela projesinde bir partinin taraftarlığının cinsiyete göre değişip değişmediği sorusuna cevap arayan bir araştırmacı ne cinsiyet değişkenini manipüle edebilecek ve ne de yanıt değişkenini aynı koşullar altında ölçebilecektir.

Örnek yoklamaları içsel geçerlilik bakımından çok zayıf olmaları münasebeti ile değişkenler arasındaki neden-ve-etki ilişkilerini ortaya çıkarmaktan ziyade korelasyonları (yani çağrışımsal ilişkileri) bize bildirmekte yararlıdır.

Nihayetinde, bir deneyin içsel geçerliliği mükemmeldir, yarı-deneysel öğrenme-çalışmalarının içsel geçerliliği azalır ve örnek yoklamalarının (ki bunlara korelasyonel öğrenme-çalışmaları da denir) içsel geçerliliği çok zayıftır.

Kontrol-altına alınmış bir deneyde *rastgele seçim* ve *rastgele atama* prensiplerinin her ikisi de birlikte mevcuttur. Fakat bir yarı-deneysel öğrenme-çalışmasında bunlardan yalnızca biri mevcut olabilir. Ancak, ikisi birlikte asla mevcut değildir.

Mini Proje 07

Bir bilimsel dergide yayımlanmış, tercihen kendi bilim alanında, bir istatistiksel araştırma makalesi bul. Kopyasını al. Araştırma makalesini aşağıdaki hususiyetleri bakımından değerlendiren bir rapor hazırla.

1. Bir nüfus sayımı mı, bir örnek yoklaması mı, bir gözlemsel öğrenme-çalışması mı, yoksa bir dizayn-edilmiş deney mi? Niçin?
2. Araştırmacının amacı nedir (yani hangi yararlılığı güdüyor)?
3. Araştırmacının ereği/objektifi (yani araştırmanın sorusu veya hipotezi) nedir?
4. Bu bölümde öğrendiğin terminolojiyi kullanarak, istihdam edilen veri toplama metodunu detaylıca betimle.
5. Yanlış yapılan herhangi bir şey var mı? Varsa, doğrusu olarak ne önerirsin?
6. Araştırmacının vardığı neticeler nelerdir? Hemfikir misin?
7. Makalenin künyesini şu formatta belirt: Makalenin yazarı, yayımlandığı yıl, makalenin adı, yayımlandığı derginin adı, çıktığı sayı, sayfası.

BÖLÜM 08 | NİSPİ FREKANSLARDAN OLASILIKLARA

Giriş

Betimleyici istatistikler ve çıkarımsal istatistikler istatistik biliminin iki safhasını oluşturuyor. Çıkarımsal istatistikleri yapabilmek için, olasılığı çok iyi anlamış olmak gerekir. Esasen, bir popülasyonun bir rastgele örneğinden o popülasyon hakkında bir genelleştirmeye gitmek belirsizlikleri ihtiva eder. Mesela, bir örnek, içinden çekildiği popülasyonun iyi bir temsilcisi midir? Eğer aynı popülasyondan farklı bir örnek daha çekseydik, hala aynı betimleyici istatistikleri elde edebilecek miydik? Çok muhtemeldir ki aynı popülasyondan çekilecek farklı örnekler farklı betimleyici istatistikler vereceklerdir. Öyleyse, bir örnekten genelleştirme yapacağımız zaman, hata yapıyor olabilme şansımız da vardır. Peki, bu olası hatayı nasıl ölçebiliriz? Elbette bir *olasılık modeli* kullanarak. İşte, olasılığın bir istatistik dersine dahil olduğu yer burasıdır.

Bir olasılık modeli kullanarak, örnekten hesapladığın istatistiklerin mütekabil popülasyon parametrelerini ne kadar isabetlice kestirdiklerini değerlendirebilirsin.

Önce, orta öğretimdeki matematik derslerinde görmüş olduğun olasılık teorisinin temel kavramlarına ve olasılık kurallarına bir göz atacağız.

Rastgelelelik

Bir *rastgele* fenomenin tezahürünün (ya da daha hususi bir ifadeyle bir aktivitenin veya prosesin *rastgelece* gerçekleşmesinin veya gerçekleştirilmesinin) sonucu olarak, bir değerin gözlenmesine (ya da ölçülmesine) bir **istatistiksel deney** denir. Bir istatistiksel deney aynı koşullar altında her tekrarlanışında, önceden bilinen olanaklı **çıkımlardan** herhangi biri ile sonuçlanabilir. Deneyin mahiyeti gereği, bir çıkım sayısal veya kategorik bir değer olabilir.

Bir madeni parayı veya zarı attığını düşün. O anki çıkımını bilebilir misin (yani atmadan önce veya attıktan sonra ama bakmadan söyleyebilir misin)? Hayır! Ancak, yine de, paranın "yazı" veya "tura" gelmesinin; ve zarın 1, 2, 3, 4, 5, veya 6 gelmesinin ne kadar olabilirlikli olduğuna dair bir kanaatimiz vardır. Bu kanaat eğer sözkonusu deneyi defalarca tekrarlarsak, bilhassa doğrudur. Niçin? Çünkü pek çok sayıda ve minik etkiler sonucu belirlediği için, bir istatistiksel deneyin çıkımını önceden kesinlikle bilebilmek pratikte imkansızdır. Mamafih, spesifik bir çıkımı önceden bilebilmesek dahi, belirli bir çıkımın meydana gelmesinin olabilirliğine matuf bir *olasılığı* tayin edebiliriz.

Olasılığın Tanımı

Olasılık kavramı 400 yıldan daha fazla bir süredir yoğun olarak öğrenilmeye-çalışılan bir konudur. Fransız matematikçiler Pierre de Fermat'ın (1607-1665) ve Blaise Pascal'ın (1623-1662) kendi zamanlarında yazdıkları bazı yazılara kadar bir geçmişi vardır. Bu süre zarfında olasılığın çeşitli tanımları yapılmıştır. Burada yalnızca ikisinden bahsedeceğiz ve olasılıkların hesaplanmasında veya kestirilmesinde bunlardan birini en çok kullanacağız.

Ampirik Tanım: Büyük Sayılar Yasası

Bir çıkımın *nispi frekansı* bu çıkımın meydana geldiği kerelerin orantısıdır. Yani, çıkımın meydana geldiği deneme sayısının denemelerin toplam sayısına bölümüdür.

Misal 1 Eğer bir yıllık periyotta yağmurlu günlerin sayısı 73 ise, yağmurlu günlerin nispi frekansı $\frac{73}{365} = 0.2$ dir.

Nispi frekans bir istatistiksel deneyin her icra edilişinden sonra değişebilir. Mamafih, bir deney pek çok kere tekrar edildiği zaman, bir çıkımın nispi frekansı bu çıkımın *olasılığına* yaklaşma temayülü

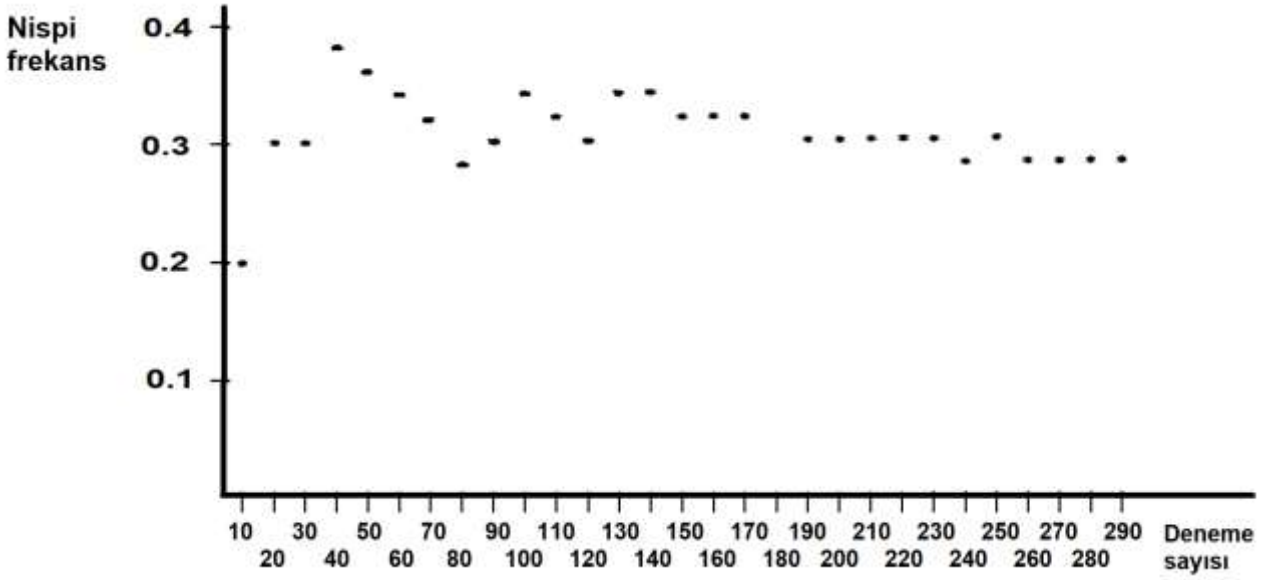
gösterir. Ki olasılığın ampirik tanımına göre, bir istatistiksel deneyin bir çıkımının **olasılığı** o çıkımın *uzun-vadeli nispi frekansıdır*.

Misal 2 Belli bir kavşaktan belirli bir günde geçen 64 aracın 16 sı kamyondur, bir başka gün geçen 50 aracın 19 u kamyondur. İki nispi frekans, $\frac{16}{64} = 0.25$ ve $\frac{19}{50} = 0.38$, biribirinden farklı. Eğer bu kavşaktan geçen araç sayılarını pek çok gün boyunca üst üste toplasaydık, kamyonların nispi frekansı ne olurdu? Her ne kadar farklı günler için kamyonların nispi frekansları kararsızlık gösteriyor olsa da, çok uzun bir vade sonunda toplanan sayılar için kamyonların nispi frekansı sabit bir değerde karar kılacaktır ki bu değere bu çıkımın olasılığı adını veriyoruz. Bu kavşaktan geçecek ilk aracın bir kamyon olması olasılığı her ne olursa olsun, farzedelim 0.32, bu idealize edilmiş sonuca, sonsuzca uzun bir vadede tekrarlanan deneylerin çıkımlarını toplamak suretiyle ulaşılacaktır.

Misal 3 (*Simülasyon*) Aşağıda verilen rastgele sayılar tablosundaki her bir rakamın büyük bir fabrikadaki mesailerine gelen işçilere tekabül ettiğini farzedelim. Sıfır, 1 ve 2 rakamları tütün içenler; 3 ten 9 a kadarki (9 da dahil olmak üzere) rakamlar da tütün içmeyenler olsun. Tabloda soldan sağa doğru giderek, her 10 kişiden sonra tütün içenlerin birikimli nispi frekanslarını hesapla ve sonuçları bir grafikte göster.

84177	06757	17613	15582	51506	81435	41050	92031	06449	05059
59884	31180	53115	84469	94868	57967	05811	84514	75011	13006
63395	55041	15866	06589	13119	71020	85940	91932	06488	74987
54355	52704	90359	02649	47496	71567	94268	08844	26294	64759
08989	57024	97284	00637	89283	03514	59195	07635	03309	72605
29357	23737	67881	03668	33876	35841	52869	23114	15864	38942

Cevap: Birikimli nispi frekanslar şöyledir: $\frac{2}{10} = 0.2, \frac{6}{20} = 0.3, \frac{9}{30} = 0.3, \frac{15}{40} = 0.375, \frac{18}{50} = 0.36, \frac{21}{60} = 0.35$, $\frac{23}{70} = 0.329$, $\frac{23}{80} = 0.288$, $\frac{27}{90} = 0.3$, $\frac{33}{100} = 0.33$ ve böyle gidiyor. Bu değerlerin buraya yazmadıklarımızla birlikte grafiği şudur:



Dikkat ediyorsan, birikimli nispi frekans grafiği önce hayli dalgalanıyor fakat daha sonrasında bir düzey tutturuyor. Bir istatistiksel deney pek çok kere tekrarlanmaya devam ettikçe, bir çıkımın nispi frekansının belli bir değere (yani bu çıkımın olasılığına ki $3/10$) gittikçe yaklaşması kavramı **büyük sayılar yasası** olarak adlandırılır.

Olasılığın ampirik tanımına göre, bir istatistiksel deneyin çıkımlarının olasılıkları birer kestirimdir. Yani birer "yaklaşımdır". Çünkü deneyi sonsuz kere tekrarlamak hiç bir zaman olanaklı değildir ve bu sebeple olasılığın değeri tam olarak hesaplanamaz.

Olayların Olasıklarını Değerlendirmek

Bir istatistiksel deneyin herhangi bir çıkımına veya birden fazla çıkımına bir arada bir **olay** adı verilir. Olasılık hesabı çıkımlar üzerinden yapılmakla birlikte, pratikte çoğu zaman olayların olasılıklarından konuşuruz. Mesela bir zar atışını bir deney olarak addedelim; zarın altı yüzünden her biri bu deneyin sonucunda olanaklı birer çıkımdır. Zarda 5 gelmesi bir olay olarak addedilebleceği gibi zarda tek sayı gelmesi başka bir olay olarak addedilebilir. İlk haldeki olay deneyin yalnızca bir çıkımından (ki 5) oluşurken ikinci haldeki olay deneyin üç ayrı çıkımından (ki 1, 3 ve 5) oluşmaktadır.

Bir *olayın* sonucu veya sonuçları olmaz, bir istatistiksel deneyin sonucu bir *olay* olur.

Bir istatistiksel deneyin sonucuda belirli bir olayın meydana gelmesi olasılığı, bu olayı oluşturan çıkımların meydana gelmesi olasılıklarına dair matematiksel bir ifadedir. Matematiksel notasyon gereği, olaylar alfabenin ilk sırasındaki büyük harflerle temsil edilir. Bir A olayının olasılığı $P(A)$ ile gösterilir.

Eğer bir istatistiksel deney (gözlem) n kere tekrarlanırsa ve n yeterince büyükse, n_A kere gözlenen bir A olayının olasılığı şöyle tanımlanır:

$$P(A) \cong \frac{n_A}{n} = \frac{A \text{ nın frekansı}}{\text{Deney sayısı}}$$

"≅" sembolü yaklaşık eşitlik manasına gelir, fakat pratikte onun yerine her zaman "=" sembolünü kullanırız. Mesela bu ders notunun yazarı asistanlık yıllarında eskimiş bir zarın hilesiz (veya adil) olup olmadığını test etmek istemişti. Zarı 60 kere attıktan sonra 1, 2, 3, 4, 5 ve 6 çıkımları için sırasıyla 12, 9, 9, 8, 12 ve 10 frekanslarını elde etti. Deneyin tekrar sayısı, sonsuz kere ile kıyaslandığında, elbette çok küçük kalıyor. Bunun bize isabetli olasılık kestirimleri temin edebileceği hayli kuşkulu olmakla beraber, mevzubahis eskimiş zarla yapılan bir atışta 5 gelmesi olasılığı,

$$P(5) = \frac{12}{60} = 0.2$$

ve tek sayı (1, 3 ve 5) gelmesi olasılığı,

$$P(\text{Tek sayı}) = \frac{12 + 9 + 12}{60} = 0.55$$

olarak hesaplanır.

Klasik Tanım: Eşit Olabilirlikli Çıkımlara Olasılıkların Atanması

Eğer bir istatistiksel deneyin m sayıdaki olanaklı çıkımlarının tümü eşit olabilirilikli ise, o zaman k sayıda çıkımı havi bir A olayının olasılığı şöyle tanımlanır:

$$P(A) = \frac{k}{m} = \frac{A \text{ daki çıkımların sayısı}}{\text{Deneyin tüm olanaklı çıkımlarının sayısı}}$$

Buna göre atılan hilesiz bir zarın üste gelen yüzündeki sayının 5 olması olasılığı,

$$P(5) = \frac{1}{6} = 0.167$$

ve tek sayı (1, 3 ve 5) olması olasılığı,

$$P(\text{Tek sayı}) = \frac{3}{6} = 0.5$$

olarak hesaplanır.

Eğer tüm olanaklı çıkımların eşit olabilirlikli olduğu varsayımı geçerli değilse, klasik olasılık tanımı yanlış ve yanıltıcı sonuçlar istihsal edecektir. Ampirik tanımın bu varsayıma bağımlı olmadığı açıktır,

öyle ki bu bölümde, bir çıkımın veya bir olayın olasılığını değerlendirmede veya kestirmede, olasılığın ampirik tanımını, olasılığın klasik tanımı bağlamı içerisinde, kullanmaya devam edeceğiz. Zaten bir gözlemin (deneyin) çıkımları eşit olabilirlikli olduğu zaman, olasılığın ampirik tanımı olasılığın klasik tanımının ampirik bir tatbiki olarak telakki edilebilir. Şöyle ki yürütülen denemelerin sonuçlarının kümesi tüm olanaklı çıkımların kümesi olarak mülahaza edilirken, ilgilenilen olayı meydana getiren sonuçlar o olaya ait çıkımlar olarak mülahaza edilir.

Olasılık Kuralları

Bir deneyin tüm olanaklı çıkımlarının bir kümesi **örnek uzayı** olarak adlandırılır ve S ile gösterilir. Bu örnek uzayının alt kümeleri olan muhtelif olayların olasılıkları daima 0 ve 1 arasında olmak zorundadır. Bir olasılığın 0 a yaklaşması, söz konusu olayın olabilirliğinin azaldığı manasına gelirken, bir olasılığın 1 e yaklaşması ise, söz konusu olayın olabilirliğinin arttığı manasına gelir. Bir istatistiksel deneyin tüm olanaklı çıkımlarının olasılıklarının toplamı daima 1 e eşit olmak zorundadır, yani $P(S) = 1$.

Misal 4 Bir markanın ürettiği bonbon şekerlerinden birini rasgele çektiğimizi ve rengini kaydettiğimizi farzedelim. Örnek uzayımız:

$$S = \{kahverengi, yeşil, turuncu, sarı, kırmızı\}$$

Bu örnek uzayında yer alan her bir çıkımın olasılığı, aynı renkteki tüm şekerlerin orantısıdır. (Genel manada bu idealize edilmiş bir durumu ifade eder ki teker teker şekerleri birbiri ardınca açıp renklerinin, sonsuza giden bir süreçte belirlenmesini gerektirir.) Üretici firmadan çıkımların hakiki olasılıklarını öğrendiğimizi yine farzedelim:

Renk	Kahverengi	Yeşil	Turuncu	Sarı	Kırmızı
Olasılık	0.3	0.2	0.2	0.1	0.2

Tüm olanaklı çıkımların olasılıklarının toplamı 1 dir:

$$0.3 + 0.2 + 0.2 + 0.1 + 0.2 = 1$$

Yani, $P(S) = 1$.

Ayrık Olaylar

Ortak çıkımı bulunmayan olaylar, *ayrıktır*. Ayrık olayların aynı anda (yani eş zamanlı olarak) meydana gelmesi olanaksız olduğu için, bu iki olay aynı zamanda *karşılıklı engelleyendir*.

Misal 5 Bir öğrenci yurdunda kalan öğrencilerin %40 ı ziraat fakültesine, %20 si tıp fakültesine ve %40 ı da edebiyat fakültesine kayıtlıdır. Bu yurtta kalan öğrencilerden herhangi birini rastgele çektiğimiz zaman, bu öğrencinin ziraat fakültesi öğrencisi olması bir olaydır, tıp fakültesi öğrencisi olması bir olaydır ve edebiyat fakültesi öğrencisi olması bir olaydır. Bir öğrenci yalnızca bir fakülteye kayıtlı olabileceği için, bunlar ayrık olaylardır.

Parayı bir kere attığımızda, paranın yazı gelmesi A olayı, tura gelmesi B olayı ise, bunlar ayrıktır. Çünkü, iki olay aynı zamanda gerçekleşemez. Parayı iki kere attığımızda, paranın birinci atışta tura gelmesi A olayı, ikinci atışta tura gelmesi B olayı ise, bunlar ayrık değildir. Çünkü iki olay aynı zamanda gerçekleşebilir (veya iki parayı aynı anda attığımızda).

Tamlayan Olaylar

Bir olayın meydana gelmemesi olasılığı, yani, tamlayanının olasılığı, o olayın olasılığının 1 den farkıdır:

$$P(A^C) = 1 - P(A)$$

Misal 6 Eğer bir müteahhitlik şirketinin bir ihaleyi kazanma olasılığı 0.3 ise, ihaleyi kazanamama olasılığı nedir?

Cevap: $1 - 0.3 = 0.7.$

Toplama Kuralı

Eğer iki olay *karşılıklı engelleyen* ise, yani, eş zamanlı olarak meydana gelmeleri olanaklı değilse, bu olaylardan en az birisinin meydana gelmesi olasılığı bu iki olayın tek başlarına meydana gelmeleri olasılıklarının toplamına eşittir:

$$\text{Eğer} \quad P(A \cap B) = 0 \qquad \text{ise} \qquad P(A \cup B) = P(A) + P(B)$$

ki burada $A \cap B$ ("A kesişim B" diye okunur) A nın ve B nin aynı anda meydana gelmesi manasındayken, $A \cup B$ ("A birleşim B" diye okunur) A nın veya B nin meydana gelmesi manasındadır.

Misal 7 Eğer bir komitenin başkanlığı için Müjgan'nın, Tahsin'in ve Jale'in seçilmesi olasılıkları sırasıyla 0.5, 0.3 ve 0.2 ise, komite başkanının Müjgan veya Jale olması olasılığı nedir?

Cevap: $0.5 + 0.2 = 0.7.$

Misal 8 Bir sosyal bilimci İngiltere'deki sosyal değişimi incelemek için babalardan ve oğullardan oluşan büyük bir örnek üzerinde öğrenme-çalışması yapmıştır. Bu örnekte yer alan her bir babanın ve oğulun sosyal statüsü, yani ait olduğu sosyal sınıf, öğrenimi ve mesleği gibi faktörlere bakılarak belirlenmiştir. Buna göre sosyal sınıf 1 (en düşük) ve 5 (en yüksek) arasında derecelendirilmiştir. Aşağıdaki tabloda, en düşük sınıftaki (1) bir babanın oğlunun mensup olacağı sosyal sınıf olasılıkları veriliyor.

Oğulun sosyal sınıfı:	1	2	3	4	5
Olasılık:	0.48	0.38	0.08	0.05	0.01

Buna göre şu olayları göz önüne alalım:

$$A: \text{oğulun 1. sınıfta kalması}$$
$$B: \text{oğulun en üst iki sınıftan birine ulaşması}$$

Tablodaki çıkımların olasılıklarını kullanarak, bu olayların olasılıkları:

$$P(A) = 0.48$$
$$P(B) = 0.05 + 0.01 = 0.06$$

Birinci sınıftan bir babanın oğlunun 1. sınıfta kalmaması olasılığı:

$$\begin{aligned} P(A^C) &= 1 - P(A) \\ &= 1 - 0.48 = 0.52 \end{aligned}$$

A ve B olayları ayrık olduğuna göre, 1. sınıftan bir babanın oğlunun yine 1. sınıfta kalması veya en üst 2 sınıftan birine ulaşması olasılığı:

$$\begin{aligned} P(A \; veya \; B) = P(A \cup B) &= P(A) + P(B) \\ &= 0.48 + 0.06 \\ &= 0.54 \end{aligned}$$

Olaylar karşılıklı engelleyen değilse, öyle basitçe olasılıklarını toplayamazsın.

İki olay karşılıklı engelleyen olmadıkları zaman, bu olayların olasılıklarının toplamı bu olayların ortak meydana gelişlerini iki kere sayar. Bu sebeple, iki olaydan en az birisinin meydana gelme olasılığını doğru hesaplayabilmek için, bu olayların tek başlarına meydana gelmeleri olasılıklarının toplamından, bu iki olayın beraber (veya aynı anda) meydana gelmesi olasılığını çıkarmak gerekir. Hal böyle olunca, müteakip kural geçerlidir.

Genel Toplama Kuralı

Herhangi iki A ve B olayı için, bu olaylardan en az birisinin meydana gelmesi olasılığı,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Misal 9 Geçmiş deneyimden biliniyor ki bir tıp fakültesinde bir öğrencinin anatomi dersinden kalması olasılığı 0.25, fizyoloji dersinden kalması olasılığı 0.21 ve her iki dersten de kalması olasılığı 0.13. Bir öğrencinin anatomi dersinden veya fizyoloji dersinden kalması olasılığı nedir?

Cevap:

$$0.25 + 0.21 - 0.13 = 0.33$$

Çarpım Kuralı

Bir olayın meydana gelme şansına bir başka olayın meydana gelmesinin veya gelmemesinin bir tesiri yoksa, yani iki olay *bağımsız* ise, bu iki olayın her ikisinin de meydana gelmesi olasılığı bu olayların tek başlarına meydana gelmeleri olasılıklarının çarpımıdır.

Misal 10 Farzet ki bir arkadaşın ardı ardına 4 kere atılan bir zarda gelen sayıları bilsin. Bu şaşırtıcı tahminin gerçekleşmesi olasılığı nedir?

Cevap: Zarı dört kere atmak aynı istatistiksel deneyi dört kere tekrar etmek demektir. Zarın önceki atılışındaki çıkımın sonraki atılışındaki çıkıma herhangi bir tesiri olmadığını biliyoruz. Buna "zarın belleği yoktur" denir. Öyle ise, bu ardışık bağımsız olayların olasılıklarını çarpabiliriz:

$$\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \left(\frac{1}{6}\right)^4 = \frac{1}{1296}$$

Olaylar bağımsız değilse, öyle basitçe olasılıklarını çarpamazsın.

Misal 11 Bir zarı ve bir parayı beraber attığımız zaman, zarın 2 ve paranın tura gelmesi olasılığı nedir?

Cevap: Zarın 2 gelmesi olasılığı $1/6$, paranın tura gelmesi olasılığı $1/2$ ve bu iki olay biribirinden bağımsız olduğuna, (yani, zarın 2 gelmesi veya gelmemesinin paranın tura gelmesine bir tesiri olmadığına ve paranın tura gelmesi veya gelmemesinin zarın 2 gelmesine bir tesiri olmadığına) göre, iki olayın olasılıklarını çarpabiliriz.

$$P(2 \text{ ve } tura) = \frac{1}{6}\left(\frac{1}{2}\right) = \frac{1}{12}$$

Çarpım kuralı ikiden çok olay için de genişletilebilir; yani, bir *bağımsız* olaylar dizisi verildiği zaman, bu olayların hepsinin meydana gelmesi olasılığı bunların tek başlarına meydana gelme olasılıklarının çarpımına eşit olacaktır.

Misal 12 Bir insanın kan grubu ebeveyninden kalıt olarak gelir. Genetik bilimi bu kalıtımın rasgele gerçekleştiğini söyler. Ebeveynlerin her biri, kan grubunu belirleyen iki ayrı gen taşır ve bu genlerden her birinin yavruya geçmesi olasılığı $1/2$ dir. Biri anadan diğeri babadan olmak üzere, yavruya geçen bu iki gen yavrunun kan grubunu belirler. Ebeveyn genleri birbirinden bağımsız olarak seçilir. Farzedelim ki ebeveynin her ikisi de O ve A genlerini taşıyor olsun. Eğer ebeveynin her ikisinden de yavruya O geni geçerse, yavrunun kan grubu O olacaktır, bunun dışında yavrunun kan grubu A olacaktır. Eğer A ile annenin yavruya O genini geçirmesi olayını ve B ile de babanın yavruya O genini geçirmesi olayını gösterirsek, yavrunun kan grubunun O olması olasılığı:

$$\begin{aligned} P(A \text{ ve } B) &= P(A)P(A) \\ &= (0.5)(0.5) \\ &= 0.25 \end{aligned}$$

Bu olasılık bize, uzun vadede, bu ebeveynden doğacak tüm çocukların $1/4$ ünün kan grubunun O

olacağını söylemektedir.

Bağımsızlığı ayrıklık ile karıştırma. Eğer A ve B ayrıksa, bu bize A meydana geldiğinde B nin meydana gelemeyeceğini söyler. Buna göre ayrık olaylar bağımsız değildir.

Paranın belleği olmadığından, eğer hilesiz bir parayı atıyorsak, ardışık atışların birbirinden bağımsız olduğunu kabul ederiz. Bunun manası, ilk atışın sonucunu görmüş bile olsak, mesela ikinci atışta tura gelme olasılığının hala $1/2$ olduğudur. Diğer taraftan, bir oyun kağıdı destesinden dağıtılan ardışık kartların renkleri bağımsız değildir. Bir standart 52 lik oyun kağıdı destesi 26 kırmızı ve 26 siyah karttan oluşur. Karılmış bir desteden ilk kartı dağıtırken kırmızı gelme olasılığı $26/52 = 0.50$. Dağıtılan kartın kırmızı olduğunu gördükten sonra, geri kalan kartlar içerisinde sadece 25 kırmızı olduğunu biliriz. İkinci çekilecek kartın kırmızı olması olasılığı, böylece, $25/51 = 0.49$. İlk dağıtılan kartın sonucunu bilmek, ikinci dağıtılacak kartın sonucuna ait olasılıkları değiştirdi.

Bir hemşire kan basıncını ardı ardına iki kere ölçtüğünde, bu istatistiksel deneyin iki sonucunun bağımsız olduğunu düşünmen makuldür. Çünkü ilk sonuç, ikinci okumayı yaparken ölçme aleti üzerinde bir tesir yaratmaz. Fakat, eğer bir IQ testi veya benzer bir zeka testine ardı ardına iki defa giriyorsan, test puanların bağımsız olmayacaktır. İlk girişinde bir öğrenme durumu yaşayacağından, bu ikinci girişinde alacağın puana tesir edecektir.

Misal 13 AIDS virüsünü tespit etmek maksadıyla yapılan bir tanılayıcı test 0.005 olasılıkla yanlış pozitif sonuç vermektedir. Yani, eğer AIDS virüsü taşımayan bir kişiye bu test uygulanırsa 0.005 olasılıkla bu virüsü taşıdığını gösteren bir yanlış sonuç ortaya çıkacaktır. Bir hastanede çalışan 140 kişiye bu test uygulanacak. Bu 140 kişiden hiçbirinin AIDS virüsünü taşımaması halinde, en az birinin yanlış pozitif çıkması olasılığı nedir?

Cevap: Farklı kişiler için test sonuçlarının bağımsız olduğu bir olasılık modeli düşünelim. AIDS virüsü taşımayan bir kişinin yanlış pozitif çıkması olasılığı 0.005 olduğundan, tamlama kuralına göre bu kişinin negatif çıkması olasılığı $1 - 0.005 = 0.995$ tir. Buna göre, 140 kişi içerisinde, en az bir kişinin yanlış pozitif çıkması olasılığı:

$$\begin{aligned} P(\text{en az bir yanlış pozitif}) &= 1 - P(\text{hiç biri pozitif değil}) \\ &= 1 - P(140 \text{ negatif}) \\ &= 1 - 0.496 \\ &= 0.504 \end{aligned}$$

Koşullu Olasılık ve Bağımsızlık

Yukarıda zikrettiğim bağımsızlık kavramını daha iyi anlayabilmek için, önce koşullu olasılık fikrini, yani, başka bir olayın meydana geldiği bilindiği halde, bir olayın olasılığını incelemek lazım gelir.

Misal 14 Aşağıdaki tablo 1200 üniveriste öğrencisinin Alkol ve tütün alışkanlıklarına ait bir yoklamanın sonuçlarını veriyor. Satır ve sütun toplamları da hesaplanmıştır.

	Alkol içmiyor	Alkol içiyor	Toplam
Tütün içmiyor	315	165	480
Tütün içiyor	585	135	720
Toplam	900	300	1200

Bu öğrencilerden herhangi birinin tütün içiyor olması olasılığı nedir?

Cevap: $P(\text{Tütün}) = \frac{720}{1200} = 0.6.$

Bir öğrencinin alkol içtiği bilindiği halde bu öğrencinin tütün içiyor olması olasılığı nedir?

Cevap: Şimdi ilgimizi alkol içen 300 öğrenciye yalnızca yöneltiyoruz, ve böylece,

$P(\text{Tütün}|\text{Alkol}) = \frac{135}{300} = 0.45$. (Kullanmış olduğum notasyonu şöyle okuruz: "alkol içtiği bilindiği halde bir öğrencinin tütün içmesi olasılığı").

Bir öğrencinin alkol içmesi olasılığı nedir?

Cevap: $P(\text{Alkol}) = \frac{300}{1200} = 0.25.$

Tütün içtiği bilindiği halde bir öğrencinin alkol içmesi olasılığı nedir?

Cevap: $P(\text{Alkol}|\text{Tütün}) = \frac{135}{720} = 0.1875.$

Koşullu olasılığın bir formülünü yazmamız gerekirse, şöyle yazabiliriz:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Misal 15 Bir bölgedeki tarım işletmelerinin %90'ı sığır yetiştiriyor ve %50 si hem sığır hem koyun yetiştiriyor. Bu bölgeden rastgele seçilen bir işletmenin sığır yetiştirdiği bilindiğine göre, aynı zamanda koyun yetiştirmesi olasılığı nedir?

Cevap:

$$P(\text{Koyun}|\text{Sığır}) = \frac{P(\text{Koyun} \cap \text{Sığır})}{P(\text{Sığır})} = \frac{0.5}{0.9} = \frac{5}{9}$$

Misal 16 Sağlaklığa karşı solaklığın IQ puanları üzerinde bir etkisi olup olmadığını araştıran bir psikolog 200 lise öğrencsini rastgele örneklemek suretiyle aşağıdaki veriyi toplamıştır.

	Sağlak	Solak	Toplam
Yüksek IQ	190	10	200
Normal IQ	1710	90	1800
Toplam	1900	100	2000

Bu gruptan rastgele seçilecek bir öğrencinin yüksek IQ lu olması olasılığı nedir?

Cevap:

$$P(\text{Yüksek IQ}) = \frac{200}{2000} = 0.1$$

Solak olduğu bilindiği halde, rasgele seçilen bir öğrencinin yüksek IQ lu olması olasılığı nedir?

Cevap: İlgilenilen popülasyonu 100 solak öğrenciye daraltıyoruz ve

$$P(\text{Yüksek IQ}|\text{Solak}) = \frac{10}{100} = 0.1$$

Dikkat etmelisin ki yüksek IQ luluk olasılığı öğrencinin solaklığından etkilenmiyor. Diyoruz ki iki olay, yüksek IQ ve solaklık, *bağımsızdır*.

Eğer $P(A|B) = P(A)$ ise, A ve B bağımsız iki olaydır. Daha önce belirtilmiş olduğu üzere, bu durumda aynı zamanda $P(A \cap B) = P(A)P(B)$ dir.

Olayların karşılıklı engelleyen olup olmadığı veya bağımsız olup olmadığı biribirinden çok farklı iki özelliktir. Birinci özellik olayların ayrık (yani kesişimsiz) olduğu manasına gelirken, ikinci özellik bir olayın meydana gelip gelmemesinin diğer olay üzerinde bir etkisi olmadığı manasına gelir. Dikkatini çekmek istiyorum ki karşılıklı engelleyen (yani ayrık) olaylar, bu olaylardan birinin olasılığı sıfır olmadığı sürece bağımsız değildir. Yani, karşılıklı engelleme durumunda $P(A \cap B) = 0$ iken, bağımsızlık durumunda $P(A \cap B) = P(A)P(B)$ dir. Ve bu iki durumun aynı anda doğru olabilmesinin tek yolu $P(A) = 0$ veya $P(B) = 0$ olmasıdır.

Sorulara sadece işlem sonuçları ile cevap vermek yerine, sonuçların nereden geldiğini

göstermelisin.

Sayma, Permütasyon ve Kombinasyon

Eğer bir olay *m* yoldan meydana geliyorsa ve bu yollardan her biri için bir ikinci olay *n* yoldan meydana geliyor ise, bu durumda iki olay birlikte *mn* yoldan meydana gelir. Buna saymanın *temel prensibi* denir. Aynı prensip üç veya daha fazla olay için de geçerlidir.

Misal 17 Bir büyük şirket, çalışanlarının kimlik kartları için sırasıyla iki harf ve üç rakamdan oluşan bir kod kullanmak istiyor. Biribirinden farklı kaç kod basılabilir?

Cevap:

$$29(29)(10)(10)(10) = 841000$$

Alıştırma 1 Misal 17 deki kod için eğer harfler ve rakamlar karışık kullanılabilseydi, kaç farklı kod basılabilirdi?

Misal 18 Bir sınıftaki öğrencilerin 16 sı oğlan, 12 si kız. Bu sınıftan bir oğlanı ve bir kızı kaç farklı yoldan seçebilirim?

Cevap: Oğlan için 16 farklı seçim yapılabilir ve kız için 12 farklı seçim yapılabilir. İki olayın birlikte meydana gelebileceği yolların sayısı:

$$16(12) = 192$$

Misal 19 Altı kişi arasından bir başkan ve bir başkan yardımcısı kaç farklı yoldan seçilir?

Cevap: Başkan için 6 farklı seçim yapılabilir ve geri kalan 5 kişi arasından da başkan yardımcısı seçilir:

$$6(5) = 30$$

Permütasyon

Mevcut *n* sayıda nesne arasından, *r* tanesini, sırasını gözeterek,

$$n(n-1)(n-2)(n-3)\cdots(n-r+1)$$

farklı yoldan seçebiliriz. Bu seçimi yapmanın her bir farklı yolu bir *permütasyon* olarak adlandırılır ve olanaklı permütasyonların toplam sayısı $P(n,r)$ ile temsil edilir:

$$P(n,r) = \frac{n!}{(n-r)!}$$

Misal 20 Altı kişi içerisinden bir başkan ve bir başkan yardımcısı seçimi kaç farklı yoldan yapılabilir?

Cevap:

$$P(6,2) = \frac{6!}{(6-2)!} = \frac{(6)(5)(4)(3)(2)(1)}{(4)(3)(2)(1)} = 6(5) = 30$$

Misal 21 On iki atletin katıldığı bir maratonun sonucunda ilk 3 sıra kaç farklı şekilde ortaya çıkabilir?

Cevap:

$$\begin{aligned} P(12,3) &= \frac{12!}{(12-3)!} \\ &= \frac{(12)(11)(10)(9)(8)(7)(6)(5)(4)(3)(2)(1)}{(9)(8)(7)(6)(5)(4)(3)(2)(1)} \\ &= 12(11)(10) \\ &= 1320 \end{aligned}$$

Kombinasyon

Mevcut *n* sayıda nesne arasından *r* tanesini, sırasını gözetmeksizin,

$$\frac{n(n-1)(n-2)(n-3)\cdots(n-r+1)}{r!}$$

yoldan seçebiliriz. Bu seçimi yapmanın her bir farklı yolu bir *kombinasyon* olarak adlandırılır ve olanaklı kombinasyonların toplam sayısı $C(n,r)$ veya $\binom{n}{r}$ ile temsil edilir:

$$C(n,r) = \frac{n!}{(n-r)!\,r!}$$

Misal 22 Altı kişi içinden, iki kişilik bir komite kaç farklı yoldan seçilebilir?

Cevap:

$$C(6,2) = \frac{6!}{(6-2)!\,2!} = \frac{6(5)(4)(3)(2)(1)}{4(3)(2)(1)(2)(1)} = 15$$

Misal 23 A, B, C ve D harflerinden 3 ü ile kaç permütasyon ve kaç kombinasyon elde edilir?

Cevap:

a,b,c	a,c,b	b,a,c	b,c,a	c,a,b	c,b,a
a,b,d	a,d,b	b,a,d	b,d,a	d,a,b	d,b,a
a,c,d	a,d,c	c,a,d	c,d,a	d,a,c	d,c,a
b,c,d	b,d,c	c,b,d	c,d,b	d,b,c	d,c,b

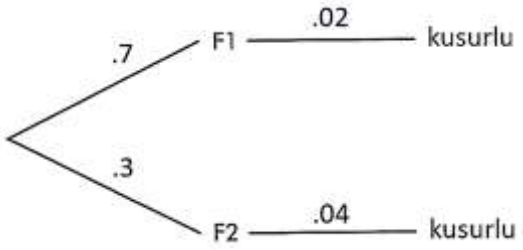
$$P(6,3) = \frac{4!}{(4-3)!} = \frac{4(3)(2)(1)}{1} = 24$$

$$C(6,3) = \frac{4!}{(4-3)!\,3!} = \frac{4(3)(2)(1)}{1(3)(2)(1)} = 4$$

Aşamalı Olasılık Hesapları

Misal 24 Bir cep telafonu üreticisi kullandığı baskı devrelerin %70 ini F1 fabrikasında ve geri kalanını da F2 fabrikasında yaptırmaktadır. Geçmiş deneyime göre, F1 fabrikasında yaptırdığı baskı devrelerin %2 si kusurluyken, F2 fabrikasında yaptırdıklarının %4 ü kusurludur. Üreticiye ulaşan herhangi bir parçanın kusurlu olması olasılığı nedir?

Cevap: Bu gibi problemlerde aşağıdaki gibi bir ağaç diyagram çizmek yararlıdır.

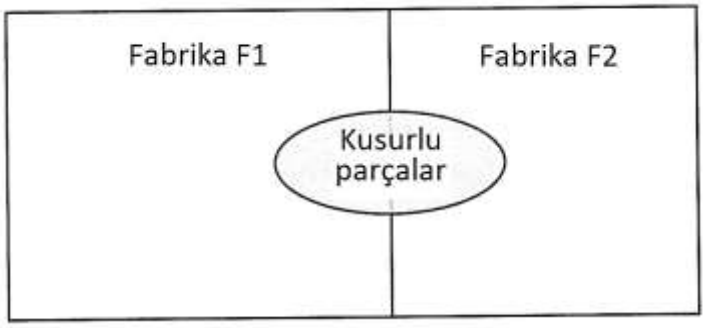


Ağaç diyagramı takip ederek,

$$P(\text{F1} \cap \text{kusurlu}) = P(\text{F1})\,P(\text{kusurlu}|\text{F1}) = 0.7(0.02) = 0.014$$

$$P(\text{F2} \cap \text{kusurlu}) = P(\text{F2})\,P(\text{kusurlu}|\text{F2}) = 0.3(0.04) = 0.012$$

Bu aşamada problemi bitirmek için bir Venn diyagramı yararlı olur:



$$P(\text{kusurlu}) = P(\text{F1} \cap \text{kusurlu}) + P(\text{F2} \cap \text{kusurlu}) = 0.014 + 0.012 = 0.026$$

Yukarıdaki analizi bir aşama daha ileri götürebiliriz ve şu gibi soruları cevaplayabiliriz: Eğer fabrikaya ulaşan baskı devreler içerisinden rastgele seçilen bir parça kusurlu ise, F1 fabrikasından gelmesi olasılığı nedir? F2 fabrikasından gelmesi olasılığı nedir?

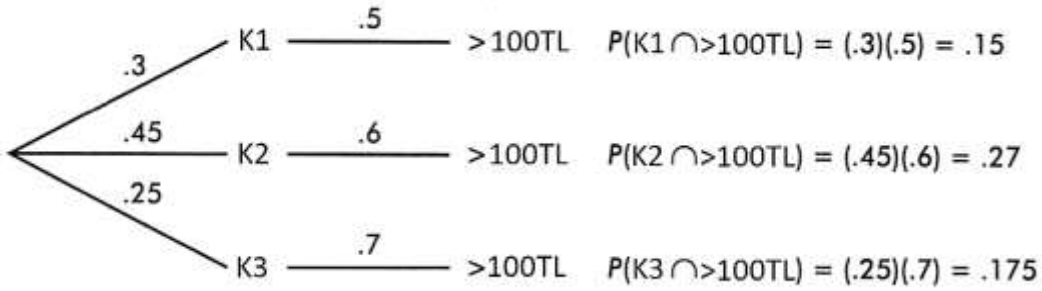
$$P(\text{F1}|\text{kusurlu}) = \frac{P(\text{F1} \cap \text{kusurlu})}{P(\text{kusurlu})} = \frac{0.014}{0.026} = 0.5385$$

$$P(\text{F2}|\text{kusurlu}) = \frac{P(\text{F2} \cap \text{kusurlu})}{P(\text{kusurlu})} = \frac{0.012}{0.026} = 0.4615$$

Koşullu olasılıkları öğrenmeye-çalışırken, ağaç diyagramlar çok yararlı olabilir.

Misal 25 Bir marketin üç farklı yöne açılan kapılarında birer kasası var ve bu kasalardan geçen günlük ortalama müşteri sayıları, sırasıyla 120, 180 ve 100. Geçmiş deneyimden biliniyor ki herhangi bir müşterinin bir geçişte 100TL den fazla harcaması olasılığı, bu kasalar için, sırasıyla 0.5, 0.6 ve 0.7. Markette bulunan müşterilerden biri rastgele seçilirse, 100TL den fazla harcaması olasılığı nedir? Müşterinin 100TL den fazla harcadığı bilindiği halde, birinci kasadan geçmiş olması olasılığı nedir?

Cevap: $P(\text{K1}) = \frac{120}{400} = 0.3$, $P(\text{K2}) = \frac{180}{400} = 0.45$, $P(\text{K3}) = \frac{100}{400} = 0.25$ olduğuna göre,



Şimdi, $P(> 100\text{TL}) = 0.15 + 0.27 + 0.175 = 0.595$ ve $P(\text{K1}| > 100\text{TL}) = \frac{0.15}{0.595} = 0.252$.

Mini Proje 08

Bir olay seç, öyle ki bu olay herkes için aynı çıkımla sonuçlanıyor olsun. Mesela, gelecek Pazar gününün yağmurlu olması gibi. Ve öyle bir olay olsun ki bu, olasılığı açıkça 0 veya 1 olmasın. On kişiden bu olayın olasılığını değerlendirmelerini iste. Biribirlerinin cevaplarını işitmelerine izin verme. Aynı kişilere, öte taraftan, hilesiz bir oyun kağıdı destesinden rastgele seçilecek bir kartın kupa olması olasılığını da sor. Her iki soruya verilen yanıtların değişkenliklerini belirle ve biribiriyle karşılaştır. Niçin farklı olduğunu açıkla. Yani, bu olasılıklar eşit olabilirlikli çıkımların bir kümesinin olasılığı olarak mı hesaplanıyor, gözlenen çıkımların uzun-vadeli bir kümesine mi dayanıyor veya bir subjektif değerlendirme olarak mı yapıldı?

BÖLÜM 09 | RASTGELE DEĞİŞKENLER VE KESİKLİ OLASILIK DAĞILIMLARI

Giriş

İstatistiğin toplumda iki rolü var. Birincisi, İstatistik steryotipler yaratma işidir. Haydi "kadınlar adamlara göre daha güleryüzlüdür" hipotezini mülahaza edelim. Steryotipi tesis etmek için, veriyi toplamaya ve onu kullanarak bir netice çıkarmaya ihtiyacımız var. Çoğu zaman, netice verinin bir farkı teklif etmesi veya etmemesidir. Netice (neredeyse) hiç bir zaman bir farkın hakikaten var olup olmadığı değildir. Yalnızca verinin bir farkı teklif edip etmediği ve ne kadar ısrarla teklif ettiğidir. İstatistik toplumdaki bu rolünü en az 100 yıldır oynamaktadır.

İstatistiğin toplumdaki daha az bilinen ikinci rolü steryotipleri madara etmektir. İstatistik değişkenliği takdir etmek (kıymetlendirmek) hakkındadır. Değişkenliği anlamak, onu açıklamak ve onu kontrol altına almak hakkındadır. Şunu beklerim ki yeterli veri ile kadınların adamlara göre ortalamada gerçekten daha güleryüzlü olduğunu gösterebiliriz. Çok fazla veri toplamak münasebetli değişkenleri kontrol altına almaya yardım eder ve bir netice çıkarmamızı temin eder. Fakat yine şunu beklerim ki rastgele bir kadının rastgele bir adama göre daha güleryüzlü olup olmadığını asla *isabetlice* tahmin edebilmeyiz. İnsanlar arasında çok fazla değişkenlik var. Gruplar arasında farklar olduğu zaman bile, grupların içindeki bireylerin gösterdiği değişkenlik ile karşılatırıldığında, bu farklar çoğunlukla bariz değildir. Bireylere dair beyanlar için, çok fazla veri toplamak yalnızca yüz yüze geldiğimiz belirsizliğin sınırlarını daha isabetlice beyan etmemize yardım eder.

Nihayetinde, İstatistik değişkenliğe tabi verinin analizinden bir netice çıkarabilmek ve eşitçe çıkarabilmemek hakkındadır. Tüm veri değişkenliğe tabidir. Değişkenliği nicelemek için, istatistikçiler olasılık kavramlarını kullanır. Veri için olasılık modelleri yaratmak suretiyle, veriyi analiz ederler.

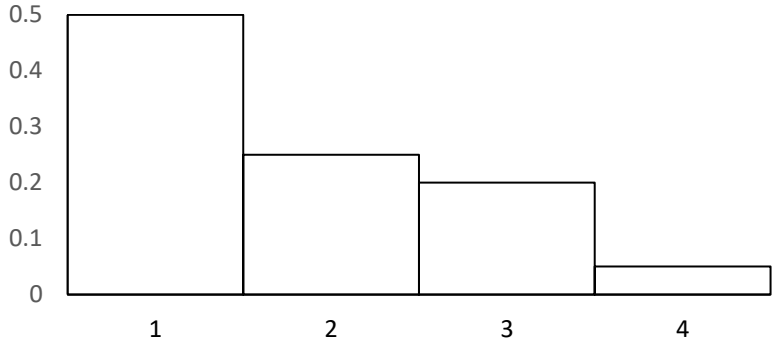
Rastgele Değişken

Bir istatistiksel deneyin çıkımlarının sayısal bir betimlemesi bir **rastgele değişken** olarak adlandırılır. Rastgele değişkenler genellikle İngiliz alfabesinin sonundaki büyük harflerle temsil edilir, mesela W, X, Y, veya Z gibi. Eğer Y bizim rastgele değişkenimizi temsil ediyorsa, olanaklı çıkımlarını $Y = y_1, y_2, y_3 \cdots$ olarak listeleriz. Mesela bir madeni parayı dört kere atarsak, çıkımı yazıların ve turaların bir dizilimi olarak kaydedebiliriz, mesela TYYT gibi. Bu çıkımın bir sayısal betimlemesi, mesela dört atışta gelen yazıların sayısı olabilir. İstatistiksel notasyonu takibederek, yazıların sayısını Y ile temsil edelim. Eğer bizim çıkımımız TYYT ise, o zaman $Y = 2$ dir. Bir sonraki çıkım YYYT ise, $Y = 3$ olur. Bu istatistiksel deneyin olanaklı çıkımları için, Y nin alacağı değerler 0, 1, 2, 3 ve 4 tür. Yani, $Y = 0, 1, 2, 3, 4$ olur.

Bir rastgele değişkenin bir çıkımının olasılığı, meydana gelme şansının bir ifadesi olarak çoğu zaman 0 ve 1 arasında bir reel sayıdır. Mesela, bir parti malda beş kusurlu parça çıkması olasılığı $1/20$ olabilir, bir şirketin bir yıl içinde 7 ihale alması olasılığı 0.01 olabilir ya da bir hastalığa yakalanan üç kişinin de iyileşmeleri olasılığı 0.95 olabilir.

Standart matematiksel notasyona göre, bir Y rastgele değişkeninin y değerini alması olasılığı $P(Y = y) = P_Y(y) = p(y)$ ile gösterilebilir. Böylece, $P(Y = 0) = P_Y(0) = p(0)$, $P(Y = 1) = P_Y(1) = p(1)$, ... yazılabilir.

Misal 1 Hastahane kayıtları belirli bir hastalık teşhisi ile yatışı yapılan hastaların $1/2$ sinin 1 gün, $1/4$ ünün 2 gün, $1/5$ inin 3 gün ve $1/20$ sinin 4 gün içinde taburcu edildiğini gösteriyor. Eğer Y ile yatış günlerinin sayısını temsil edersek, Y bir rastgele değişkendir ve $1/2$ olasılıkla 1 değerini, $1/4$ olasılıkla 2 değerini, $1/5$ olasılıkla 3 değerini ve $1/20$ olasılıkla 4 değerini alır.



Misal 1 deki rastgele değişkenin **kesikli** oduğunu söyleriz, çünkü yalnızca sayma sayılarını (yani, tam sayıları) değer olarak alabilmektedir. Öte taraftan, Misal 2 deki rastgele değişkenin **sürekli** olduğunu söyleriz, çünkü tam sayıların yanısıra sayı doğrusu üzerindeki bir aralıkta bulunan herhangi bir başka değeri de (yani reel sayıları) alabilmektedir.

Misal 2 Bir sürat teknesinin hızını Y rastgele değişkeni ile temsil edelim. Sürat teknesi belki saatte 423.275 km hız yapacak; belki 423.275078 km, belki 423.2750781000256 km, belki de belli bir aralıktaki herhangi bir başka hızda gidecek. Bu durumda, sürat teknesinin hızının saatte 423 km ve 424 km arasında olması olasılığını sorabiliriz.

Bir Y kesikli rastgele değişkenine ait **olasılık dağılımı** (veya *olasılık modeli*), bu rastgele değişkenin alabileceği her bir y değerini ve $P(Y = y)$ olasılığını veren bir liste veya formüldür.

Mesela para atmak gibi pek çok uygulamada, yalnızca *iki olanaklı çıkım* mevcuttur. Başka benzeri misaller, şunlardır: yarın yağmur yağacak veya yağmayacak, aldığın ürün sağlam çıkacak veya çıkmayacak, işçiler greve gidecek veya gitmeyecek, ameliyat başarılı olacak veya olmayacak, rastgele seçilen bir kişinin aylık kazancı 5000TL den daha yüksektir veya değildir.

Yalnızca iki olanaklı çıkımı olan bir deneye **Bernoulli deneyi** adı verilir. Bir Bernoulli denemesi olasılık ve istatistikte yürütebileceğin en basit deneydir ki yalnızca iki olanaklı çıkımdan birisi ile sonuçlanır. Bu çıkımlardan birisi *başarılı (veya arzu edilen) sonuç* olarak nitelenir ve sayısal olarak 1 ile betimlenirken, diğeri *başarısız (veya arzu edilmeyen) sonuç* olarak nitelenir ve sayısal olarak 0 ile betimlenir. Başarılı sonucun olasılığı π ile simgelenir. Buna göre başarısız sonucun olasılığı daima $1 - \pi$ dir.

Misal 3 Önceki deneyim gösteriyor ki belirli bir fabrikanın ürettiği ampullerin 10 da 1 i arızalıdır. Bu fabrikanın ürettiği ampullerden birini satın alırsan, arızalı çıkması ve arızalı çıkmaması olasılıkları nedir?

Cevap: Rastgele seçilen bir ampulün arızalı çıkması olasılığı,

$$P(Y = 1) = \pi = \frac{1}{10} = 0.1.$$

Arızalı çıkmaması olasılığı,

$$P(Y = 0) = 1 - \pi = 1 - \frac{1}{10} = 1 - 0.1 = 0.9.$$

Bazı durumlarda, iki-çıkımlı proseslerin (yani istatistiksel deneylerin) sonuçları birden çok kere gözlenir. Mesela, bir üretim bandından rastgele seçilen dört parça üründen en çok birinin kusurlu çıkması şansını öğrenmek isteyebiliriz, dikilen 9 fidandan en az 7 sinin tutması şansını ya da yapılan 4 atıştan tam iksinin hedefe isabet etmesi şansını öğrenmek isteyebiliriz. Bunlar gibi iki-çıkımlı bir prosesin (yani bir Bernoulli deneyinin) belli sayıda kere bağımsızca tekrarlandığı ve olanaklı iki çıkımdan her birinin olasılığının her tekrarlamada aynı kaldığı bir durumda, başırılı sonuçların sayısı için **binomiyal (iki-terimli) olasılıklar** hesaplanabilir.

Misal 4 Bir ampulün bozuk olması olasılığı 0.1 dir. Rastgele seçilecek dört ampulün tümünün de bozuk olması olasılığı nedir?

Cevap: Bağımsızlıktan dolayı (yani, bir ampulün bozuk olmasına başka herhangi bir ampulün bozuk olup olmamasının bir tesiri olmamasından dolayı), tüm ampullerin bozuk olması olasılığını bulmak için, ampullerin teker teker bozuk olmaları olasılıklarını biribiriyle çarpabiliriz:

$$P(Y = 4) = 0.1(0.1)(0.1)(0.1) = 0.1^4 = 0.0001$$

Misal 5 Yine, bir ampulün bozuk olması olasılığı 0.1 dir. Rastgele seçilecek üç ampulden tam olarak ikisinin bozuk olması olasılığı nedir?

Cevap: Problemi şu şekilde parçalayabiliriz: İlk iki ampulün bozuk ve üçüncünün sağlam olması olasılığı $0.1(0.1)(0.9) = 0.009$ dur. (Dikkatini çekmek istiyorum ki ampulün bozuk olması olasılığı 0.1 olduğuna göre, sağlam olması olasılığı 0.9 dur). İlk ampulün sağlam ve diğer iki ampulün bozuk olması olasılığı $0.9(0.1)(0.1) = 0.009$ dur. Nihayetinde, ikinci ampulün bozuk ve diğer ikisinin sağlam olmaları olasılığı $0.1(0.1)(0.9) = 0.009$ dur. Her üç kombinasyonun olasılıklarını topladığımızda, üç ampulden tam ikisinin bozuk olması olaslığını $0.009 + 0.009 + 0.009 = 0.027$ olarak buluruz.

Misal 6 Misal 5 ten devamla, rastgele seçilen sekiz ampulden tam üçünün bozuk olması olasılığı nedir?

Cevap: Yine problemi parçalayalım. Mesela, ikinci, beşinci ve sekizinci ampullerin bozuk olaması olasılığı,

$$0.9(0.1)(0.9)(0.9)(0.1)(0.9)(0.9)(0.1) = 0.1^3 0.9^5 = 0.00059049.$$

Üçüncü, dördüncü ve yedinci ampullerin bozuk geri kalanların sağlam olması olasılığı,

$$0.9(0.9)(0.1)(0.1)(0.9)(0.9)(0.1)(0.9) = 0.1^3 0.9^5 = 0.00059049.$$

Anlaşılabileceği üzere, üç bozuk ve beş sağlam ampulün herhangi bir belirli diziliminin olasılığı $0.1^3(0.9^5) = 0.00059049$ dur. Böyle kaç dizilim mevcuttur? Başka bir deyişle, bozuk ampuller için, sekiz pozisyondan üçünü kaç farklı yoldan çekebiliriz (geriye kalan beş pozisyon sağlam ampuller içindir). Cevabı bize *kombinasyonlar* verir:

$$\binom{8}{3} = \frac{8!}{5!\,3!} = \frac{8(7)(6)}{3(2)} = 56$$

Bu 56 dizilimden her birinin olasılığı 0.00059049 dur. Buna göre, sekiz ampulden tam üçünün bozuk olması olasılı ğı $56 \times 0.00059049 = 0.03306744$ tür. Ya da:

$$P(Y = 3) = \binom{8}{3} 0.1^3(0.9^5) = 56(0.1^3)0.9^5 = 0.03306744$$

Misal 7 Bir şirkette çalışanların %30 u tütün içiyor. Beş kişilik bir grubu rastgele seçersen, bunlardan tam ikisinin tütün içicisi olması olasılığı nedir?

Cevap: Şöyle akıl yürütüyoruz. Bir kişinin tütün içmesi olasılığı $\%30 = 0.3$ olduğuna göre, bir kişinin tütün içmemesi olasılığı $1 - 0.3 = 0.7$. Tütün içen iki kişinin ve tütün içmeyen üç kişinin belirli bir diziliminin olasılığı $0.3^2(0.7^3) = 0.03087$. Böyle dizilimlerin sayısı,

$$\binom{5}{2} = \frac{5!}{2!\,3!} = 10.$$

Bu 10 dizilimin her birinin olasılığı 0.03087 olduğuna göre, nihai cevabımız $10 \times 0.03087 = 0.3087$. Ya da:

$$P(Y = 2) = \binom{5}{2} 0.3^2(0.7^3) = \frac{5!}{2!\,3!} 0.3^2 0.7^3 = 0.3087$$

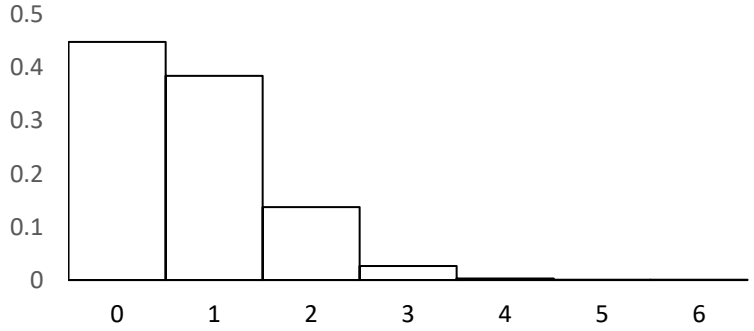
Binomiyal (İki-terimli) Olasılık Dağılımı

Yukarıda rastgele fenomenlerin iki-çıkımlı (binomiyal) sonuçlarına dair yaptığımız olasılık

hesaplarını şimdi bir genel prensiple ifade edeceğiz. Farzedelim ki bir istatistiksel deneyin *başarı* ve *başarısızlık* olarak adlandırılan iki olanaklı çıkımı var; başarı olasılığı π ye ve başarısızlık olasılığı $1-\pi$ ye eşit. Elbette ki $\pi+(1-\pi)=1$. Ve yine farzedelim ki deney n defa tekrarlanıyor ve herhangi bir belirli defaki çıkımın herhangi bir başka defaki çıkıma hiç bir tesiri yok. Buna göre, Y rastgele kesikli değişkeni için yapılan n tane gözlem içerisinde tam y tane başarı olasılığı (ve böylece $n-y$ tane başarısızlık olasılığı) şöyle hesaplanır:

$$P(Y=y)=\binom{n}{y}\pi^y(1-\pi)^{n-y}=\frac{n!}{y!\,(n-y)!}\pi^y(1-\pi)^{n-y}$$

Misal 8 Bir sığır ırkında bir buzağının albino doğması olasılığı 0.125 tir. Bu ırktan altı tane gebe ineği bulunan bir işletmecinin tam olarak bir tane albino buzağıya sahip olması olasılığı nedir?

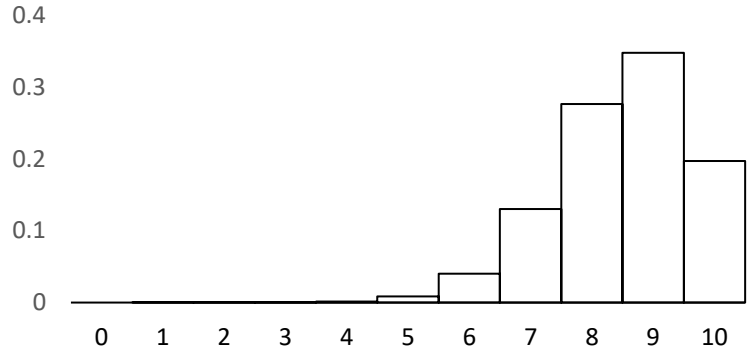


Cevap: Eğer bir buzağının albino doğması olasılığı 0.125 ise, albino doğmaması olasılığı $1-0.125=0.875$ tir. Eğer altı buzağıdan birisi albino ise, $6-1=5$ buzağı albino değildir. Buna göre, bilmek istediğimiz olasılık,

$$P(Y=1)=\binom{6}{1}0.125^1 0.875^5=\frac{6!}{1!\,5!}0.125^1 0.875^5=0.385.$$

Olasılık uygulamaları pek çok zaman, "en az", "en çok", "-den az" ve "-den çok" gibi ibareleri havidir. Böylesi durumlarda, problemin çözümü, iki veya daha fazla hal için sonuçların toplanmasını gerektirir.

Misal 9 Bir fabrikada bir parti malın üretimi tamamlandıktan sonra, şu kalite kontrol prosedürü takip edilmektedir: Eğer rastgele seçilen 10 parçadan en az sekizi tüm spesifikasyonları sağlıyor ise, o partinin tümünün sevkiyatı onaylanıyor. Eğer üretilen bir parti malın, gerçekte, %85 i tüm spesifikasyonları sağlıyor ise, kalite kontrol prosedürü uygulandıktan sonra bu partinin sevkiyatının onaylanması olasılığı nedir?



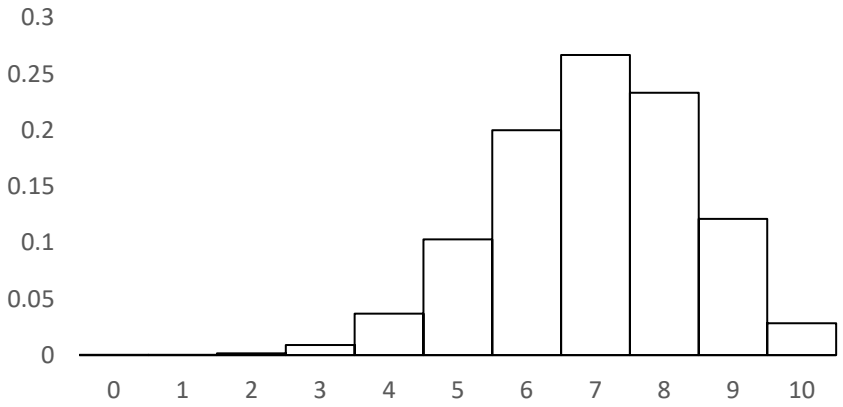
Cevap: Rastgele seçilecek bir parçanın spesifikasyonları sağlaması olasılığı 0.85 ve böylece spesifikasyonları sağlamaması olasılığı $1-0.85=0.15$ olmalıdır. On parçadan en az sekizinin spesifikasyonları sağlaması olasılığını bulmak istiyoruz. Yani, tam sekiz veya tam dokuz veya tam 10 parçanın spesifikasyonları sağlaması olasılığı. Bu üç halin binomiyal olasılıklarını toplayacağız:

10 parçanın tam 8 tanesinin spesifikasyonları sağlaması		10 parçanın tam 9 tanesinin spesifikasyonları sağlaması		10 parçanın tam 10 tanesinin spesifikasyonları sağlaması
$\binom{10}{8}0.85^8 0.15^2$	+	$\binom{10}{9}0.85^9 0.15^1$	+	$\binom{10}{10}0.85^{10} 0.15^0$

$$P(Y \geq 8) = \frac{10!}{8!\,2!}0.85^8 0.15^2 + \frac{10!}{9!\,1!}0.85^9 0.15^1 + \frac{10!}{10!\,0!}0.85^{10} 0.15^0$$
$$= 0.82$$

$Y \sim B(n, \pi)$ notasyonu Y değişkeninin örnek hacmi n ve başarı olasılığı π olan bir binomiyal dağılıma sahip olduğunu ifade eder.

Misal 10 Misal 9 deki probleme devamla, üretilen bir parti malın, gerçekte, %70 i tüm spesifikasyonları sağlıyor ise, kalite kontrol prosedürü uygulandıktan sonra sevkiyatının onaylanması olasılığı nedir?

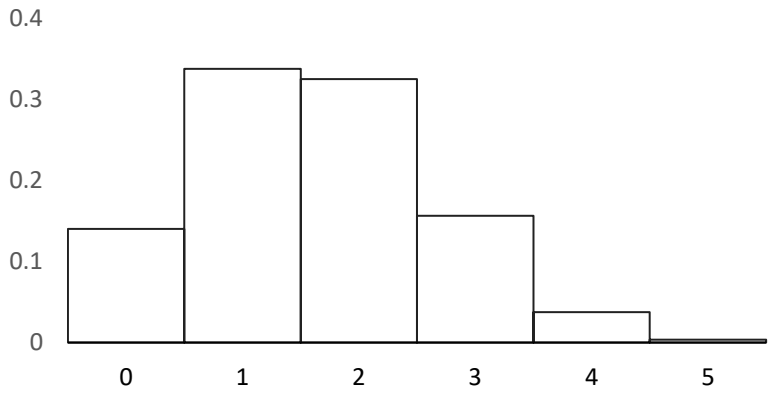


Cevap:

$$P(Y \geq 8) = \binom{10}{8}0.7^8 0.3^2 + \binom{10}{9}0.7^9 0.3^1 + \binom{10}{10}0.7^{10} 0.3^0$$
$$= 450.7^8 0.3^2 + 100.7^9 0.3^1 + 10.7^{10} 0.3^0$$
$$= 0.383$$

Bazı durumlarda, tamlayan olayın olasılığını hesaplamak ve 1 den çıkarmak daha kolaydır.

Misal 11 Atılhan Döver adlı boksörün kariyerindeki tüm maçları gözlendiğinde, nakavt ile kazandıklarının orantısı 0.325. Bu boksörün çıkacağı beş maçtan en az birini nakavtla kazanma olasılığı nedir?



Cevap: Tam bir nakavt, tam iki nakavt, tam üç nakavt, tam dört nakavt ve tam beş nakavt olasılıklarını toplayabiliriz. Mamafih, "en az bir nakavt"ın tamlayanı "sıfır nakavt"tır. Tam sıfır

nakavt olasılığı,

$$P(Y=0)=\binom{5}{0}0.325^0 0.675^5=0.675^5=0.14$$

ve buna göre, beş maçta en az bir nakavt olasılığı,

$$P(Y\geq 1)=1-P(0)=1-0.14=0.86.$$

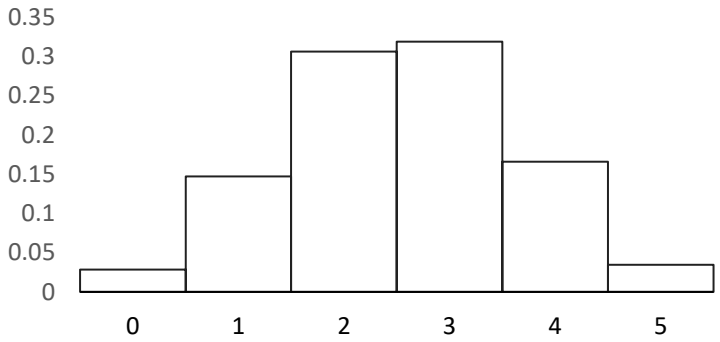
Misal 12 Bir süpermarket müşterilerinin %35 inin ödemelerini nakit parayla yaptığını belirlemiştir. Bir kasada bekleyen yedi müşteri varsa, ödemelerin dörtten azının nakit parayla yapılması olasılığı nedir?

Cevap: Bu durumda, "dörtten az"ın manası "sıfır veya bir veya iki veya üç"tür.

$$\begin{aligned}P(Y<4)&=\binom{7}{0}0.35^0 0.65^7+\binom{7}{1}0.35^1 0.65^6+\binom{7}{2}0.35^2 0.65^5+\binom{7}{3}0.35^3\times 0.65^4\\&=0.65^7+(7)0.35^1 0.65^6+(21)0.35^2 0.65^5+(35)0.35^3 0.65^4\\&=0.8\end{aligned}$$

Bazen, olanaklı tüm çıkımların herbirinin olasılığını hesaplaman istenebilir, senden. Böyle olunca, sonuçların toplamda 1 i vermeleri gerektiğini unutmamalısın.

Misal 13 Bir bebeğin kız doğması olasılığı 0.51 ise, beş çocuklu bir ailenin tüm çocuklarının kız olması olasılığı nedir? Tam dört kız? Tam üç kız? Tam iki kız? Tam bir kız? Tümü oğlan?



Cevap:

$$P(5\text{ kız})=\binom{5}{5}0.51^5 0.49^0=0.51^5=0.0345$$

$$P(4\text{ kız})=\binom{5}{4}0.51^4 0.49^1=(5)0.51^4 0.49=0.1657$$

$$P(3\text{ kız})=\binom{5}{3}0.51^3 0.49^2=(10)0.51^3 0.49^2=0.3185$$

$$P(2\text{ kız})=\binom{5}{2}0.51^2 0.49^3=(10)0.51^2 0.49^3=0.3060$$

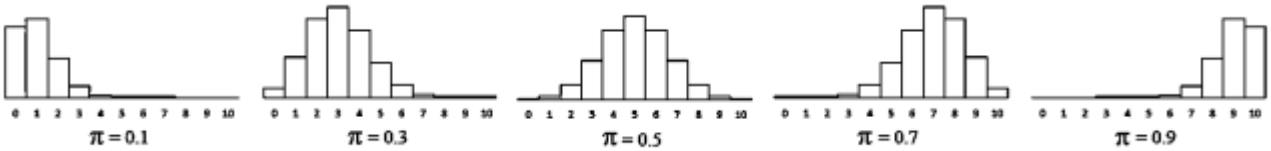
$$P(1\text{ kız})=\binom{5}{1}0.51^1 0.49^4=(5)(0.51)0.49^4=0.1470$$

$$P(0\text{ kız})=\binom{5}{0}0.51^0 0.49^5=0.49^5=0.0283$$

$$\sum P(y)=0.0345+0.1657+0.3185+0.3060+0.1470+0.0283=1.0000$$

Misal 13 teki gibi bir liste bir *olasılık dağılımını* gösterir ki tüm çıkımların ve bu çıkımların olasıklarının bir listesidir. Bir olasılık dağılımı aynı zamanda bir *olasılık modelidir*. İstatistik biliminde, her popülasyon mütekabil bir olasılık dağılımı ile özdeşleştirilir.

Binomiyal olasılık modeli söz konusu olduğunda, verilen misallerden açıkça anlaşılmış olacağı üzere, π nin aldığı değer dağılımın şeklini belirlemektedir. Mesela $n = 10$ olduğu zaman, $\pi = 0.1$, $\pi = 0.3$, $\pi = 0.5$, $\pi = 0.7$ ve $\pi = 0.9$ için çizilen histogramlar şöyledir.

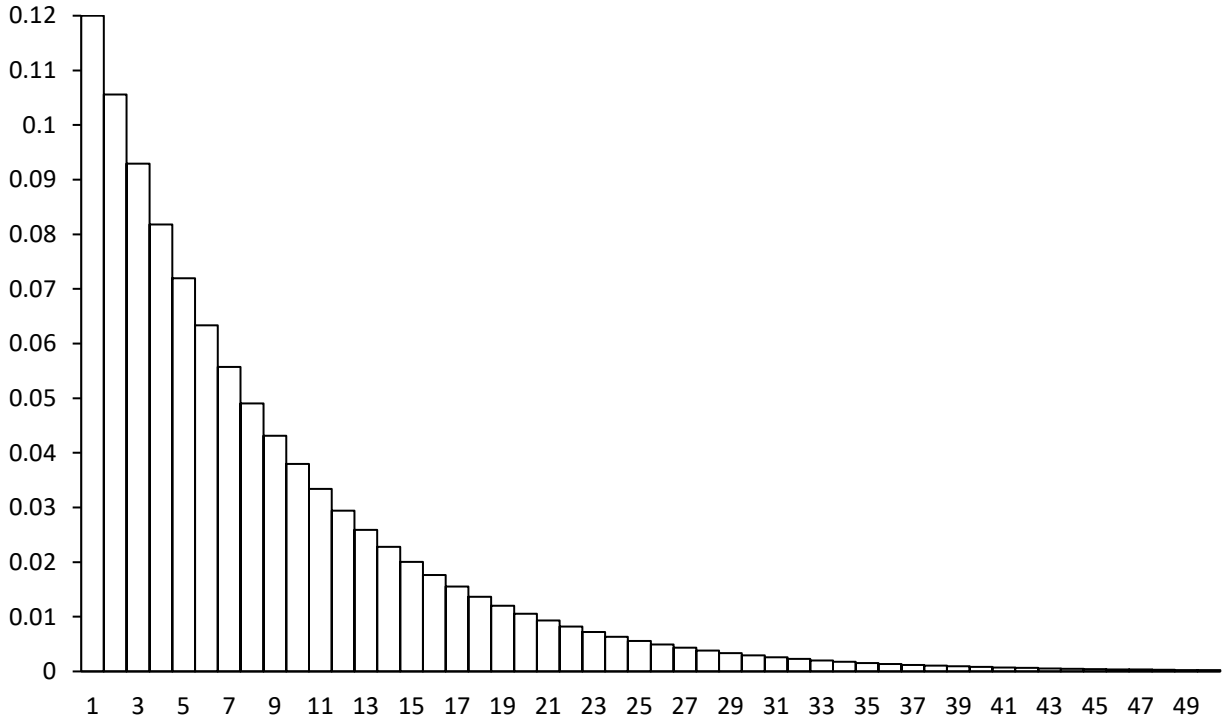


Geometrik Olasılık Dağılımı

Eğer bir istatistiksel deneyin *başarı* ve *başarısızlık* olarak adlandırılan iki olanaklı çıkımı var, başarı olasılığı π ye ve başarısızlık olasılığı $1 - \pi$ ye eşit ve deneyler biribirinden bağımsız ise, Y rastgele kesikli değişkeni için yapılan gözlemlerde ilk başarının y inci deneyde elde edilmesi olasılığı şöyle hesaplanır:

$$P(Y = y) = (1 - \pi)^{y-1}\pi$$

Misal 14 Bir patates cipsi markasının ürünlerinin %12 sinde hediye çeki vardır. Bir müşterinin ilk hediye çekini üçüncü cips paketinde bulması olasılığı nedir?



Cevap:

$$P(Y = 3) = 0.88^2 0.12 = 0.092928$$

İlk hediye çekini en geç dördüncü cips paketinde bulması olasılığı nedir?

Cevap:

$$P(Y \leq 4) = 0.12 + 0.88(0.12) + 0.88^2 0.12 + 0.88^3 0.12 = 0.40030464$$

Olasılık Dağılımlarının Simülasyonu

Cebirsel hesaplamalar yapmak yerine, olasılık problemlerini çözmek üzere bazen simülasyon kullanırız.

Misal 15 Futbol maçlarına dair yapılan bir öğrenme-çalışması, bir önceki maçını kazanmış olan takımların, çıkacağı ilk maçı kazanma şansının %60 olduğunu gösteriyor. Buna göre, bir önceki maçlarını kazanmış beş takımdan en az dördünün çıkacakları ilk maçları kazanacaklarına inanmak akla uygun mudur? Bu soruyu simülasyon kullanarak cevapla.

Cevap: Rastgele değişkenin gözlenen değerlerinin bir simülasyonunu yapacağız. Bu maksatla rastgele sayılar tablosununun (Tablo A) ilk altı satırını kullanacağız. Bir önceki maçını kazanmış olan takımların, çıkacağı ilk maçı kazanma şansı %60 olduğu için, tablodaki iki basamaklı sayılardan 01 ve 60 arasında olanları bu çıkımı temsil etsin. Tabloda bu koşulu sağlayan çiftlerin altını çiziyoruz. Her beş çiftlik grup içerisinde, kaç tanesinin 01 ve 60 arasında olduğunu kaydediyoruz. sayıların kendini tekrar etmesinin mahsuru yok.

8417706757	1761315582	5150681435	4105092031	0644905059
5988431180	5311584469	9486857967	0581184514	7501113006
6339555041	1586606589	1311971020	8594091932	0648874987
5435552704	9035902649	4749671567	9426808844	2629464759
0898957024	9728400637	8928303514	5919507635	0330972605
2935723737	6788103668	3387635841	5286923114	1586438942

Mesela, beş çiftlik ilk grupta, beş çiftten ikisi 01 ve 60 arasındadır (17 ve 57). İkinci grupta, beş çiftten üçü istenen aralıktadır (17, 31 ve 55). Nihayetinde, tüm çıkımların frekanslarını bir tabloya dökersek,

Galibiyet sayısı	Frekans
0	1
1	0
2	4
3	8
4	12
5	5

Bu simülasyonda görüyoruz ki 30 kerenin 17 sinde, bir önceki maçını kazanmış olan beş takımdan en az dördü çıkacakları ilk maçları kazanacaklardır. Böylece, asıl sorunun cevabına gelirsek, bir önceki maçını kazanmış olan beş takımdan en az dördünün çıkacakları ilk maçları kazanacaklarına inanmak akıl dışı değil. (Peki, rastgele sayılar tablosunun tamamını kullanmak suretiyle, eğer 30 kereden sonra simülasyonu devam ettirseydik, sonuç ne olurdu?)

Şunu unutmaman gerekiyor ki *örneklerin sayısı* ile *örnek hacmi* aynı şey değildir. Dikkat ettiysen, yukarıda her biri beş maçtan oluşan 30 örneği simüle ettik (yani hacmi 5 olan bir örnek çekme işlemini 30 kere yineledik).

Misal 15 te örnekleme yapmak için kullandığımız simülasyon, binomiyal dağılıma ilişkindi. Müteakip misalde ise, geometrik dağılıma ilişkin olacak. Biliyoruz ki bir geometrik dağılım ilk başarılı çıkım elde edilinciye kadar icra edilmesi gereken deney sayısına ait olasılıkları gösterir.

Bir simülasyon yaptığın zaman, öyle detaylıca betimlemelisin ki başkaları da senin prosedürünü tekrarlayabilsin.

Misal 16 Pazara yeni giren bir otomobil markasının tanınırlık oranı 0.3 tür. Simülasyon kullanarak, rastgele seçilecek birinci kişinin, ikinci kişinin, üçüncü kişinin ve sonraki kişilerin markayı tanıyacak ilk kişi olması olasılıklarını hesapla.

Cevap: Rastgele sayılar tablosunu, örneklerin simülasyonu için kullanıyoruz. Tabloda yeralan 0, 1 ve 2 sayıları markanın tanınması çıkımını temsil ederken, kendileri de dahil olmak üzere 3 ve 9 arasındaki rakamlar markanın tanınmaması çıkımını temsil etsin. Soldan sağa doğru, tablodaki rakamları birer basamak olarak okuyoruz ve 0, 1 veya 2 rakamlarına gelinceye kadar kaçar tane basamak geçtiğimizi kaydediyoruz. Misal 15 te yaptığımız gibi, Tablo A nın baştan altı satırını kullanacağız.

84177 06757 17613 15582 51506 81435 41050 92031 06449 05059
59884 31180 53115 84469 94868 57967 05811 84514 75011 13006
63395 55041 15866 06589 13119 71020 85940 91932 06488 74987
54355 52704 90359 02649 47496 71567 94268 08844 26294 64759
08989 57024 97284 00637 89283 03514 59195 07635 03309 72605
29357 23737 67881 03668 33876 35841 52869 23114 15864 38942

Rastgele sayılar tablosundan şu örnekleri çekiyoruz:

841: üçüncü deneyde ilk başarı,
770: üçüncü deneyde ilk başarı,
67571: beşinci deneyde ilk başarı,
761: üçüncü deneyde ilk başarı,
31: ikinci deneyde ilk başarı
⋮

ve böylece devam ettikten sonra, tüm çıkımların frekanslarını bir tabloya dökersek,

İlk başarının kaçıncı deneyde elde edildiği	Frekans
1	21
2	19
3	16
4	7
5	9
>5	12

Buna göre, olasılıkları şöyle kestirebiliriz:

İlk başarının kaçıncı deneyde elde edildiği	Olasılık kestirimi (yani nispi frekans)
1	$21/84 = 0.25$
2	$19/84 = 0.23$
3	$16/84 = 0.19$
4	$7/84 = 0.08$
5	$9/84 = 0.11$
>5	$12/84 = 0.14$

Öte taraftan, bu çıkımların matematiksel (yani a priori) olasılıkları sırasıyla şunlardır: 0.3, $0.7 \times 0.3 = 0.21$, $0.7^2 \times 0.3 = 0.147$, $0.7^3 \times 0.3 = 0.1029$, $0.7^4 \times 0.3 = 0.07203$ ve $1 - (0.3 + 0.21 + 0.147 + 0.1029 + 0.07203) = 0.16807$.

Misal 17 Binomiyal ve geometrik dağılımların haricinde bir simülasyon yapacağız, bu misalimizde. Şöyle ki bir şirket müdürü, üç adam ve altı kadın elemanı arasından iki kişilik bir komiteyi rastgele seçecektir. Bu seçimin tam bir adam ve bir kadından oluşması olasılığı nedir? Cevabı kestirmek için simülasyon kullan.

Cevap: Simülasyon için rastgele sayılar tablosunu (Tablo A) kullanıyoruz. Her bir basamaktaki 1, 2 ve 3 rakamları adamları temsil ederken, 4, 5, 6, 7, 8 ve 9 rakamları kadınları temsil etsin. Aynı kişi komitede yalnızca bir kere yer alabileceği için, yapacağımız örneklemenin *yerine koymadan* bir örnekleme olması gerkiyor. Prosedürümüz şöyle ki rastgele sayılar tablosunun istediğimiz herhangi bir yerinden başlayarak, biribirinden farklı iki rakam seçinceye kadar, her defasında bir sonraki basamaktaki rakamı çekeceğiz. Yine, şirketin toplamda 9 elemanı olduğu için, sıfırları da gözardı edeceğiz.

8417706757 1761315582 5150681435 4105092031 0644905059
5988431180 5311584469 9486857967 0581184514 7501113006
6339555041 1586606589 1311971020 8594091932 0648874987
5435552704 9035902649 4749671567 9426808844 2629464759

0898957024	9728400637	8928303514	5919507635	0330972605
2935723737	6788103668	3387635841	5286923114	1586438942

Tablodan rastgele örnekleri çekmeye başlayınca,

(8, 4): bir adam, bir kadın
(1, 7): bir adam, bir kadın
(7, 6): iki kadın
(7, 5): iki kadın
(7, 1): bir adam, bir kadın
(7, 6): iki kadın
(1, 3): iki adam
⋮

ve böylece devam ettikten sonra, çekilen örnekler bize nihai bir sonuç olarak şunu verir:

İki adam	:	11
Bir adam ve bir kadın	:	53
İki kadın	:	66

Buna göre, komitenin tam bir adam ve bir kadından oluşması olasılığının kestirimi,

$$\frac{53}{11+53+66} = \frac{53}{130} = 0.41.$$

Öte taraftan, bu olayın matematiksel (yani a priori) olasılığını şöyle hesaplarız: tam bir adamın ve bir kadının çekilmesi iki farklı yoldan meydana gelebileceği ve çekilişler yerine koymadan yapılacağına göre,

$$\frac{3}{6}\left(\frac{6}{8}\right) + \frac{6}{9}\left(\frac{3}{8}\right) = 0.5.$$

N hacminde ve K sayıda belli bir özelliğe sahip nesneyi haiz sınırlı bir popülasyondan, yerine koymadan yapılacak n sayıdaki çekilişte y sayıda başarı (yani çekilen nesnenin belli özellikte olması) olasılığı *hipergeometrik olasılık dağılımı* kullanılarak ta hesaplanabilir.

$$P(Y=y) = \frac{\binom{K}{y}\binom{N-K}{n-y}}{\binom{N}{n}} = \frac{\binom{3}{1}\binom{6}{1}}{\binom{9}{2}} = \frac{18}{36} = 0.5.$$

Hipergeometrik dağılımın aksine, Binomiyal dağılım *yerine koyarak* yapılacak n sayıdaki çekilişte y sayıda başarı olasılığını betimler.

İstatistiksel Anlamlılık

Farzedelim ki basketbol topuyla 10 serbest atış yaptın ve 10 unu da çemberden geçirdin. Bu beklediğinden daha fazla olurdu herhalde. Fakat, gerçekte, bu ne kadar beklenmediktir? Geçmiş deneyimine göre, serbest atışlarda isabet yüzdenin %60 olduğuna inanıyorsun. Bu durumda, 10 atışın hepsini isabet ettirme olasılığın, yalnızca $\binom{10}{10}(0.6)^{10}(0.4)^{0} = 0.006$ dır. Yani binde altı. Eğer senin isabet yüzden gerçekten %60 ise, bu çok bu çok nadir görülebilir. Bunun manası şudur ki bu sonucun şans eseri meydana gelmesi aşırı olabilirliksizdir. Bu da şunun kanıtıdır ki senin serbest atışlarda isabet yüzden gerçekte %60 tan daha yüksek olabilir. Bu sonuç **istatistiksel anlamlılık** olarak tanımlanır. Çünkü, yalnızca şans eseri meydana geldiğini farzetmek makul değildir. Makul olmamasının sebebi, eğer serbest atışlarda isabet yüzden gerçekten %60 ise, bu sonucun hayli nadirliğidir.

Alıştırma 1 Eskimiş ve yıpranmış bir zar 60 kere atılıyor ve 1, 2, 3, 4, 5 ve 6 çıkımları için sırasıyla 12, 9, 9, 8, 12 ve 10 frekansları elde ediliyor. Bu zar hilesizlik (veya adillik) vasfını kaybetmiş olabilir mi? Bir anlamlılık testi icra ederek, sonuucun interpretasyonunu yapabilir misin?

Mini Proje 09

Hilesiz bir madeni parayı dört kere at ve gelen turaların sayısını kaydet. Bu istatistiksel deneyi 30 kere tekrar et. Üç atıştaki tura sayılarının nispi frekanslarını hesapla ve bir tabloya yaz. Bu rastgele değişkenin dağılımının binomiyal olasılık dağılımına uyup uymadığını tartış. Eğer deneyi 60 kere veya 120 kere tekrarlasaydın, bu uyum hakkında ne söyleyebilirdin?

Deney No.	1. Atış	2. Atış	3. Atış	4. Atış	Tura sayısı
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					

BÖLÜM 10 | BEKLENEN DEĞERLER

Giriş

Rulette şansını deneyen birini düşünelim. Elindeki çipi çarktaki bir rakama yerleştirdikten sonra, çok para kazanmak için az şansı ve az para kazanmak için çok şansı olacaktır. Bir bahsin "adil" olup olmadığını belirlemek istediğimiz zaman, bahisçi için oyunun *beklenen değerini* hesaplamamız gerekir. Bir sigorta şirketi, belirli bir yılda nispeten az sayıda kişiye büyük miktarlarda para öderken, başka çok sayıda kişilerden küçük miktarlarda para tahsil eder. Mesela yaşam sigortası poliçesi primlerini belirlemek için, bir aktüer, poliçe sahiplerinin çeşitli yaş gruplarındaki ölümlerinin beklenen değerlerini hesaplamak zorundadır. Bir önceki bölümde, rastgele değişken ve olasılık dağılımı kavramlarını geliştirdik. Bu bölümde, bir rastgle değişkenin beklenen değerlerini veya aritmetik-ortlamasını ve varyansını nasıl belirleyeceğimizi tartışacağız.

Bir Rastgele Değişkenin Beklenen Değeri: Aritmetik ortalama

Misal 1 Üniversiteden mezun olan ve fakat kendi mesleğini icra edebileceği bir iş bulamayan Oğuz adlı genç, kentin işlek bir caddesinde su satmaya başlar. Geçmiş deneyimine göre, güneşli günlerde 60TL, yağmurlu günlerde 45TL ve karlı günlerde 15TL kazanmaktadır. Kışa girerken paltosunu yenilemek isteyen Oğuz, bir sonraki ay günlük kazancının ne kadar olacağını bilmek istemektedir. Okumuş bir genç olması hasebiyle, Meteoroloji Genel Müdürlüğü'ne başvurarak, şu enformasyonu temin eder. Kasım ayında herhangi bir günün güneşli, yağmurlu veya karlı olması olasılıkları sırasıyla $1/2$, $1/3$ ve $1/6$ dır.

Hülasa, elimizde 60, 45 ve 15 değerlerini alan bir Y rastgele değişkeni var. Şöyle ki

Çıkım:	Güneşli	Yağmurlu	Karlı
Olasılık:	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$
Rastgele değişken (Y):	60	45	15

Kasım ayında Oğuz'un beklenen günlük kazancı nedir?

Cevap: Şöyle akıl yürütüyoruz. Kasım ayı 30 gün çektiğine göre, günlerin $1/2$ sinde hava güneşli (=15 gün), $1/3$ ünde yağmurlu (=10 gün) ve $1/6$ sında karlı (=5 gün) olacaktır. Öyleyse, Oğuzun bir günlük beklenen kazancı,

$$\frac{(15)60 + (10)45 + (5)15}{30} = 47.5\text{TL}.$$

Dikkat edersen, payda yer alan üç terimin her birini önce 30 a böldükten sonra toplayarak ta aynı sonuca ulaşabiliriz,

$$\frac{15}{30}60 + \frac{10}{30}45 + \frac{5}{30}15$$

ve kesirleri daha sadeleştirerek yazarsak,

$$\frac{1}{2}60 + \frac{1}{3}45 + \frac{1}{6}15 = 47.5\text{TL}$$

Esasen, görüldüğü üzere, 30 günü telakki etmeye ihtiyaç yoktu. Basitçe olasılıkları tekabül ettikleri değerlerle çarpıp, çarpımları toplamak kafiydi.

Misal 1 deki son hesaplama bizi beklenen değerin tanımına sevk edecek. Bir Y rastgele değişkeninin *beklenen değeri* (veya başka bir deyişle *aritmetik ortalaması*), olanaklı y_i değerlerinin mütekabil $P(y_i)$ olasılıklarıyla çarpımlarının toplamıdır:

$$E(Y) = \mu_Y = \sum y_i \, P(y_i)$$

Başka kitaplarda *beklenen değer* yerine *matematik ümit* teriminin de kullanıldığına rastlayabilirsin.

Alıştırma 1 Kasım ayında Oğuz'un beklenen toplam kazancı nedir?

Misal 2 Bir piyongo çekilişi için 10 bin adet bilet satıldı. Her bir biletin bedeli 1 TL ve kazanana ödenecek ikramiye 7500 TL dir. Herhangi bir bahisçinin beklenen kazancı nedir?

Cevap: Gerçek kazanç 7499TL, çünkü çekilişi kazanabilmek için zaten bileti satın alırken 1TL ödemiş olmak gerekiyor. Buna göre,

Çıkım:	Kazanç	Kayıp
Olasılık:	$\frac{1}{10000}$	$\frac{9999}{10000}$
Rastgele değişken (Y):	7499	–1

$$E(Y) = 7499\frac{1}{10000} + (-1)\frac{9999}{10000} = -0.25$$

Böylece, piyangoyla bahse giren (yani piyango bileti satın alan) herhangi bir kişi için aritmetik ortalama kazanç – 0.25TL dir. Başka bir deyişle, piyango organizatörünün sattığı her bir bilet başına beklenen kazancı 0.25TL dir.

Misal 3 Bir şirket müdürü üç opsiyon arasında seçim yapmak durumunda kalır. A opsiyonunu seçerse 250 bin TL kazanma şansı %10, aksi taktirde 10 bin TL kaybedecek. B opsiyonunu seçerse 40 bin TL kazanma şansı %50, aksi taktirde 2 bin TL kaybedecek. C opsiyonunu seçerse 800 bin TL kazanma şansı %5, aksi taktirde 20 bin TL kaybedecek. Müdür hangi opsiyonu seçmeli?

Cevap: Her üç opsiyonun beklenen değerlerini hesaplayalım:

	A opsiyonu		B opsiyonu		C opsiyonu	
Çıkım:	Kazanç	Kayıp	Kazanç	Kayıp	Kazanç	Kayıp
Olasılık:	0.10	0.90	0.50	0.50	0.05	0.95
Rastgele değişken:	250000	–10000	–40000	–2000	–800000	–20000

$$E(A) = 0.10(25000) + 0.90(-10000) = 16000\text{TL}$$

$$E(B) = 0.50(40000) + 0.50(-2000) = 19000\text{TL}$$

$$E(C) = 0.05(800000) + 0.95(-20000) = 21000\text{TL}$$

Mevzubahis müdürün ilk elde seçmesi gereken opsiyon, C opsiyonu olmalı! Mamafih, her ne kadar C opsiyonunun aritmetik ortalaması (veya beklenen değeri) en yüksek olsa dahi, müdür çeşitli opsiyonları nispi risklilikleri bakımından mülahaza etmek isteyebilir. Eğer, mesela 5 bin TL nin üstünde bir kayıp şirket için tahripkar olacak ise, müdür B opsiyonunu seçmek isteyebilir, çünkü B opsiyonunun maksimum olanaklı kaybı 2 bin TL dir. (Bu noktada sezgisel olarak açıkça görülebilir ki riski ölçerken, varyans kavramı burada işimize yarardı.)

Binomiyal Rastgele Değişken

Aldığı değerler bir binomiyal olasılık dağılımından gelen "başarı" sayıları olan bir rastgele değişkenin beklenen değerini daha spesifik bir yoldan hesaplayabiliriz.

Misal 4 Belirli bir fabrikada üretilen otomobillerin %40 ının kusurlu olduğu biliniyor. Bir şirket bu

fabrikada üretilmiş otomobillerden 5 adet satın alıyor. Kusurlu otomobillerin beklenen değeri nedir?

Cevap: Aritmetik ortalamanın veya beklenen değerin 5 in %40 ı olduğunu $[5(0.4) = 2]$ söyleyebiliriz. Fakat, formal tanımı kullanmak suretiyle bu sonuca yeniden varalım. Kusurlu otomobillerin sayısını Y ile temsil edersek,

$$P(0) = \binom{5}{0} 0.4^0 0.6^5 = 0.6^5 = 0.07776$$

$$P(1) = \binom{5}{1} 0.4^1 0.6^4 = 5(0.4)0.6^4 = 0.25920$$

$$P(2) = \binom{5}{2} 0.4^2 0.6^3 = (10)0.4^2 0.6^3 = 0.34560$$

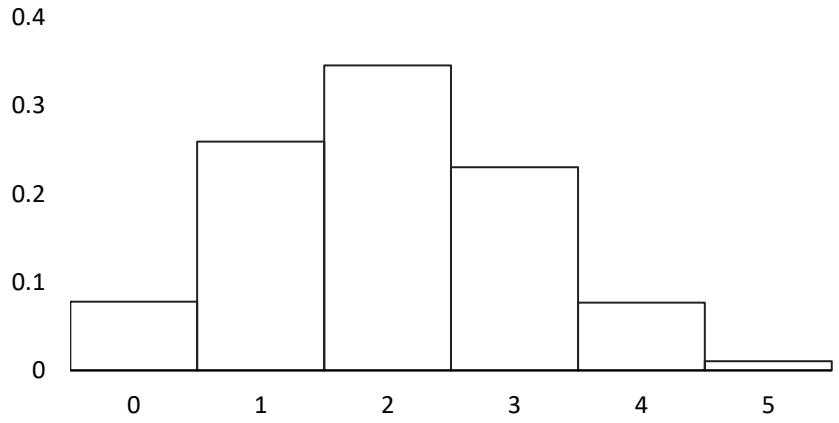
$$P(3) = \binom{5}{3} 0.4^3 0.6^2 = (10)0.4^3 0.6^2 = 0.23040$$

$$P(4) = \binom{5}{4} 0.4^4 0.6^1 = (5)0.4^4 0.6^1 = 0.07680$$

$$P(5) = \binom{5}{5} 0.4^5 0.6^0 = 0.4^5 = 0.01024$$

Rastgele değişkenin olanaklı çıkımlarını ve mütekabil olasılıklarını bir tabloya dökersek,

Çıkım:	0	1	2	3	4	5
Olasılık:	0.07776	0.25920	0.34560	0.23040	0.07680	0.01024
Rastgele değişken (Y):	0	1	2	3	4	5



$$E(Y) = 0(0.07776) + 1(0.2592) + 2(0.3456) + 3(0.2304) + 4(0.0768) + 5(0.01024) = 2$$

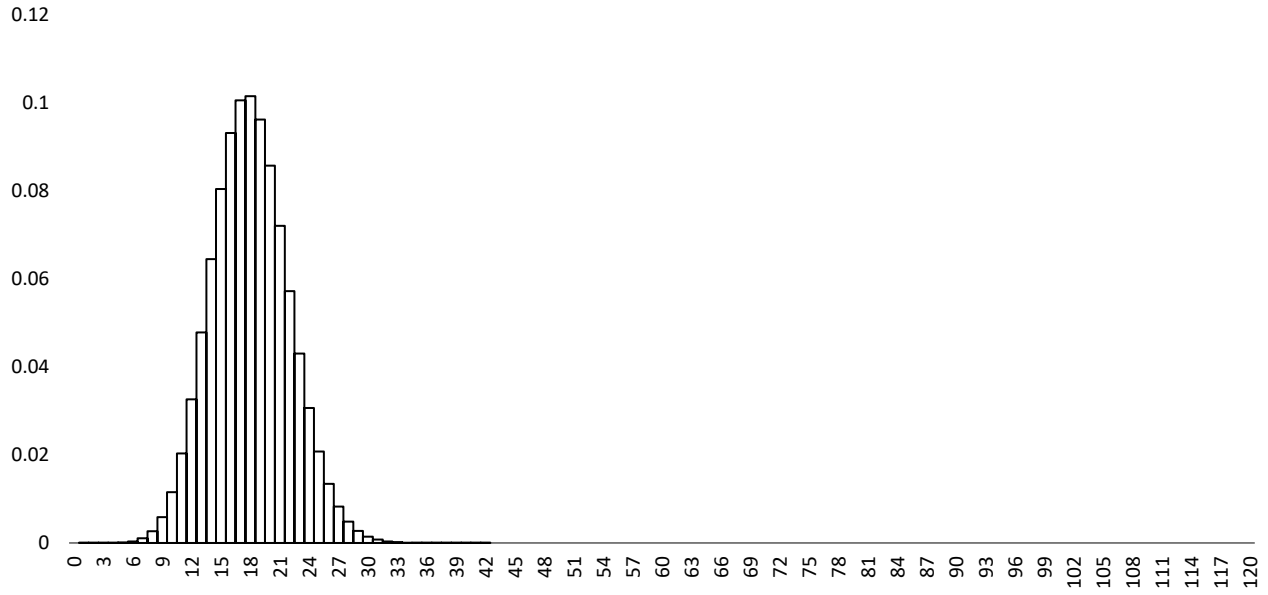
Böylece, cevap, basitçe "başarı" olasılığını binomiyal deney sayısıyla çarparak elde ettiğimiz sonuçla aynı çıkmış oldu.

Bir binomiyal olasılık durumunda, başarı olasılığı π ye ve deney sayısı n ye eşit ise, n deneydeki başarı sayısının beklenen değeri $n\pi$ dir.

Misal 5 Bir sigorta satış temsilcisi, görüştüğü kişilerin %15 ine poliçe satabilmektedir. Önümüzdeki bir ay içinde, bu temsilci 120 kişi ile görüşecek olduğuna göre, satacağı poliçelerin beklenen değeri nedir?

Cevap: Bu bir binomiyal deney, başarı olasılığı 0.15 ve deneme sayısı 120. Buna göre, başarıların beklenen değeri (veya aritmetik ortalaması),

$$n\pi = 120(0.15) = 18.$$



Bir Rastgele Değişkenin Beklenen Değerinden Kareli Sapmasının Beklenen Değeri: Varyans

Bir rastgele değişkenin beklenen değeri, yani aritmetik ortalaması, $\sum y_i\, P(y_i)$, önemlidir, fakat aynı zamanda bir rastgele değişkenin standart sapmasını da bilmek isteriz. Bunu bilmek bize rastgele değişkenin değişkenliğini ölçme becerisi kazandırır. Anımasayacağın üzere, önceki bölümlerde, bir veri kümesinin değişkenliğini ölçmek için *varyans* fikrini (*standart sapma* ile beraber) zaten kullanmıştık: Aritmetik ortalamadan sapmaların karelerinin aritmetik ortalaması. Aynı fikri takip ederek, burada $(y_i - \mu_Y)^2$ terimlerini yine bir rastgele değişkenin değerleri olarak telakki edersek, bu değerlerin olasılıkları y_i olasılıkları ile aynı olduğundan dolayı, bu yeni rastgele değişkenin beklenen değeri $\sum(y_i - \mu_Y)^2\, P(y_i)$ olur. Ki bu ifade bir rastgele değişkenin varyansı ile tam olarak aynı tanıma sahiptir:

$$E[(Y - \mu_Y)^2] = \sigma_Y^2 = \sum (y_i - \mu_Y)^2 P(y_i)$$

Bir rastgele değişkenin varyansı $var(Y)$ ile de temsil edilir. Hali hazırda ve yine biliyoruz ki standart sapma, σ, varyansın, σ^2, kare köküdür.

Misal 6 Misal 1 de üniversite mezunu Oğuz'un Kasım ayında beklenen (veya aritmetik ortalama) günlük kazancını hesaplamıştık.

$$E(Y) = \mu_Y = \frac{1}{2}60 + \frac{1}{3}45 + \frac{1}{6}15 = 47.5\text{TL}$$

Bu rastgele değişkenin beklenen değerinden kareli sapmasının beklenen değeri (veya varyansı) nedir?

Cevap:

$$\begin{aligned} E[(Y - \mu_Y)^2] = \sigma_Y^2 &= \sum (y_i - \mu_Y)^2 P(y_i) \\ &= \sum (60 - 47.5)^2 \frac{1}{2} + (45 - 47.5)^2 \frac{1}{3} + (15 - 47.5)^2 \frac{1}{6} \\ &= 256.25\text{TL}^2 \end{aligned}$$

Binomiyal Rastgele Değişken

Aldığı değerler bir binomiyal olasılık dağılımından gelen "başarı" sayıları olan bir rastgele değişkenin

beklenen değerinden kareli sapmasının beklenen değerini daha spesifik bir yoldan hesaplayabiliriz.

Misal 7 Misal 4 te bir şirketin belirli bir otomobil fabrikasından satın aldığı beş otomobil içerisindeki kusurlu otomobillerin sayısının beklenen değerini (veya aritmetik ortalamasını) hesaplamıştık.

$$E(Y) = \mu_Y = 0(0.07776) + 1(0.2592) + 2(0.3456) + 3(0.2304) + 4(0.0768) + 5(0.01024) = 2$$

Bu rastgele değişkenin beklenen değerinden kareli sapmasının beklenen değeri (veya varyansı) nedir?

Cevap:

$$\begin{aligned} E[(Y-\mu_Y)^2] &= \sigma^2 \\ &= \sum (0-2)^2 0.07776 + (1-2)^2 0.2592 + (2-2)^2 0.3456 + (3-2)^2 0.2304 \\ &\quad + (4-2)^2 0.0768 + (5-2)^2 0.01024 \\ &= 1.2 \end{aligned}$$

Bu sonucu daha spesifik bir yoldan hesaplayabilir miydik? Bu bir binomiyal rastgele değişken olduğu için, evet, beklenen değerinden kareli sapmasının beklenen değerini (veya varyansını) şöyle hesaplayabiliriz,

$$n\pi(1-\pi).$$

Yani,

$$E[(Y-\mu_Y)^2] = \sigma_Y^2 = n\pi(1-\pi) = 5(0.4)0.6 = 1.2.$$

Bu bölümde şimdiye kadar öğrendiklerimizi genel manada şöyle özetleyebiliriz. Bir Y rastgele değişkeni için,

Beklenen değer	:	*Aritmetik ortalama*	:	$\mu = \sum y_i P(y_i)$
Beklenen değerden kareli sapmanın beklenen değeri	:	*Varyans*	:	$\sigma^2 = \sum (y_i - \mu)^2 P(y_i)$
		Standart sapma	:	$\sigma = \sqrt{\sum (y_i - \mu)^2 P(y_i)}$

Başarı olasılığının π ye ve deney sayısının n ye eşit olduğu, bir **binomiyal olasılık dağılımı** durumunda, eğer n deneydeki başarı sayısını Y ile temsil edersek,

Beklenen değer	:	*Aritmetik ortalama*	:	$\mu = n\pi$
Beklenen değerden kareli sapmanın beklenen değeri	:	*Varyans*	:	$\sigma^2 = n\pi(1-\pi)$
		Standart sapma	:	$\sigma = \sqrt{n\pi(1-\pi)}$

Başarı olasılığının π ye eşit olduğu, bir **geometrik olasılık dağılımı** durumunda, eğer ilk başarının elde edildiği deney sayısını Y ile temsil edersek,

Beklenen değer	:	*Aritmetik ortalama*	:	$\mu = \frac{1}{\pi}$
Beklenen değerden kareli sapmanın beklenen değeri	:	*Varyans*	:	$\sigma^2 = \frac{1-\pi}{\pi^2}$

		Standart sapma	:	$\sigma = \sqrt{\frac{1-\pi}{\pi^2}}$

Misal 8 Tropikal bir bölgede yaşayan örümceklerin %30 unun zehirli olduğu biliniyor. Bu bölgede beş kere örümcek ısırığına maruz kalınması durumunda zehirlenme sayısının aritmetik ortalaması ve standart sapması nedir?

Cevap: Bu rastgele değişken bir binomiyal olasılık dağılımına sahip olduğu için,

$$\mu = n\pi = 5(0.3) = 1.5$$

$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{5(0.3)(0.7)} = 1.05$$

Misal 9 Basketbola yeni başlayan bir genç atışlarının %20 sini çemberden geçirebiliyor? Bu genç ilk sayısını ortalama kaçıncı atışında yapar? Standart sapması nedir?

Cevap: Bu rastgele değişken bir geometrik olasılık dağılımına sahip olduğu için,

$$\mu = \frac{1}{\pi} = \frac{1}{0.2} = 5$$

$$\sigma = \sqrt{\frac{1-\pi}{\pi^2}} = \sqrt{\frac{1-0.2}{0.2^2}} = 4.4721$$

Bir Sabitle Çarpmak ve Toplamak

Bazen bir rastgele değişkeni bir sabitle çarpmak veya toplamak icap edebilir. Mesela, derece Fahrenheit cinsinden ölçülmüş sıcaklık gözlemlerini derece Celsius cinsinden ifade etmek istersek, 32 çıkarıp $5/9$ ile çarparız. Böylesi doğrusal dönüştürme hallerinde şu kurallar geçerlidir:

$$\mu_{Y\pm a} = \mu_Y \pm a$$

$$\sigma^2_{Y\pm a} = \sigma^2_Y$$

$$\mu_{bY} = b\mu_Y$$

$$\sigma^2_{bY} = b^2\sigma^2_Y$$

Misal 16 İstatistik sınavında sorduğu soruların hayli zor olduğunu ve sınıf ortalamasının düşüklüğünü gören Profesör Korkut, puanların aritmetik ortalamasını 10 puan artırmaya karar verir. Bunu başarabilmek için, her öğrencinin puanına kaç puan daha eklemesi gerekir?

Cevap: Yukarıda verilen kural gereğince,

$$\mu_{Y\pm a} = \mu_Y \pm a$$

olduğuna göre, sınıf ortalamasını 10 puan daha artırmak için, öğrencilerin puanlarına 10 ar puan eklemesi gerekir.

Bu işlem sonrasında, puanların yayılımı nasıl farklılaşır?

Cevap: Mevcut puanlara 10 ar puan daha eklemek, verinin yayılımını farklılaştırmaz.

Misal 17 Ataman Bey'in fabrikasında çalışan işçilerin aldıkları aylık maaşların aritmetik ortalaması 2347.6 TL ve standart sapması 184.2 TL dir. Ataman Bey, gelecek ay, her işçinin maaşını 1.3 katına çıkarmayı düşünüyor. İşçi başına kaç TL lik artış olacaktır?

Cevap: Mevcut maaşı X, zamlı maaşı Y ve maaş artışını da W ile temsil edelim. Yukarıda ifade

ettiğimiz kurallar gereğince,

$$\mu_Y = b\mu_X \qquad \text{ve} \qquad \mu_W = \mu_Y - \mu_X$$

olduğuna göre, zamlı maaşların aritmetik ortalaması

$$(1.3)2347.6 = 3051{,}88 \text{ TL}$$

ve işçi başına (yani, aritmetik ortalama) artış,

$$3051.88 - 2347.6 = 704{,}28 \text{ TL}$$

olacaktır.

İşçi başına artışın standart sapması kaç olacaktır?

Cevap:

$$Y = 1.3X \qquad \text{ve} \qquad W = Y - X$$

olduğuna göre,

$$W = 1.3X - X = 0.3X.$$

Böylece, yine yukarıdaki kural gereğince,

$$\sqrt{\sigma_W^2} = \sqrt{\sigma_{0.3X}^2} = \sqrt{0.3^2\sigma_X^2} = \sqrt{0.3^2 184.2^2} = 55.26.$$

Bağımsız Rastgele Değişkenlerin Kombinasyonu

Rastgele değişkenleri şans olaylarına tekabül eden sayısal değerler olarak tanımladık. Aynı gözlem üniteleri üzerinde biribirinden farklı rastgele değişkenlerin aldıkları değerleri gözleyebilmekteyiz. Pek çok istatistiksel deneyde her bir çıkıma bir sayı çifti tekabül edebilmektedir. Mesela iki zarı birlikte attığımızı farzedelim, her zarda gelen sayı ayrı birer rastgele değişken olduğu gibi iki zarda gelen sayıların toplamı da ayrı bir rastgele değişkendir. İki rastgele değişkenin toplamından veya farkından türeyen yeni bir rastgele değişkenin dağılımının merkezini ve yayılımını orijinal değişkenler cinsinden nasıl ifade ederiz? Bu bölümde bunu göreceğiz.

Rastgele değişkenler biribirinden bağımsız olduğu zaman, iki rastgele değişkenin toplamının veya farkının aritmetik ortalamasını ve standart sapmasını hesaplamak kolaydır.

Bağımsızlık

Eğer X ve Y rastgele değişkenlerinin bağımsız olduğunu söylersek, her x ve y değeri için $P(y|x) = P(y)$ dir. Aynı manada, eğer X ve Y bağımsız ise, her x ve y değeri için $P(x \cap y) = P(x)P(y)$ dir.

Yukarıdaki ifadeleri, bilindiği üzere, $P(Y = y|X = x) = P(Y = y)$ ve $P(X = x \cap Y = y) = P(X = x)P(Y = y)$ formunda da yazabiliriz.

Misal 10 Motorlu araçlar üreten bir şirkete ait bir bayide üç farklı model ATV satılmaktadır. Bu üç modelin perakende fiyatları sırasıyla 20 bin TL, 25 bin TL ve 30 bin TL dir. Her satışta, satış elemanı 500 TL veya 750 TL prim almaktadır. Çeşitli çıkımlara ait olasılıklar şunlardır:

		Prim	
		500 TL	750 TL
Satış tutarı	20000 TL	0.30	0.05
	25000 TL	0.20	0.20
	30000 TL	0.05	0.20

Eğer "satış tutarı" rastgele değişkenini X ile gösterirsek, X in olasılık dağılımı nedir?

Cevap: Her satırı topladıktan sonra,

x:	20000	25000	30000
$P(X = x)$:	0.35	0.40	0.25

Eğer "prim" rastgele değişkenini Y ile gösterirsek, Y nin olasılık dağılımı nedir?

Cevap: Her sütunu topladıktan sonra,

y:	500	750
$P(Y = y)$:	0.55	0.45

Satış tutarının 30 bin TL olması halinde, satış elemanının 500 TL prim alması olasılığı nedir?

Cevap:

$$P(500\text{ TL prim}|30000\text{ TL satış}) = \frac{P(500\text{ TL prim} \cap 30000\text{ TL satış})}{P(30000\text{ TL satış})}$$

$$= \frac{0.05}{0.25} = 0.2$$

Mamafih, $P(500\text{ TL prim}) = 0.55$ ve böylece $P(500\text{ TL prim}|30000\text{ TL satış}) \neq P(500\text{ TL prim})$. Buna göre, diyoruz ki satış tutarı rastgele değişkeni ve prim rastgele değişkeni bağımsız değildir.

Misal 11 C vitamini ve soğuk algınlığı üzerine yürütülen bir öğrenme-çalışmasında, denekler rastgele iki gruba ayrıldıktan sonra, her gün bir gruba hiç C vimini içermeyen, bir gruba ise 500 mg C vitamini

içeren tabletler içirildi ve soğuk algınlığına yaklanıp yakalanmadıkları kaydedildi. "C vitamini" rastgele değişkenini X ile gösterirsek, 0 ve 500 değerlerini almaktadır. "Soğuk algınlığı" rastgele değişkenini Y ile gösterirsek, 0 (soğuk algınlığı yok) ve 1 (en az bir soğuk algınlığı var) değerlerini almaktadır. Müteakip tablo bize gözlenen olasılıkları, yani bu iki değişkenin ortak dağılımına ait olasılık kestirimlerini (veya gözlenen nispi frekansları) veriyor:

		Soğuk algınlığı (Y)	
		0	1
C vitamini (X)	0	0.10	0.30
	500	0.15	0.45

C vitamini rastgele değişkeni ve soğuk algınlığı rastgele değişkeni bağımsız mıdır?

Cevap: Satırları ve sütunları topladıktan sonra,

$$P(X = 0) = 0.10 + 0.30 = 0.40$$

$$P(X = 500) = 0.15 + 0.45 = 0.60$$

$$P(Y = 0) = 0.10 + 0.15 = 0.25$$

$$P(Y = 1) = 0.30 + 0.45 = 0.75$$

Şimdi koşullu olasılıkları kontrol edelim,

$$P(Y = 0|X = 0) = \frac{0.10}{0.40} = 0.25 = P(Y = 0)$$

$$P(Y = 0|X = 500) = \frac{0.15}{0.60} = 0.25 = P(Y = 0)$$

$$P(Y = 1|X = 0) = \frac{0.30}{0.40} = 0.75 = P(Y = 1)$$

$$P(Y = 1|X = 500) = \frac{0.45}{0.60} = 0.75 = P(Y = 1)$$

Böylece, şu neticeye varıyoruz ki X ve Y bağımsız rastgele değişkenlerdir. Alternatif olarak, bağımsızlığı şu yoldan da kontrol edebiliriz: $P(X = 0, Y = 0) = 0.10$ ve $P(X = 0)P(Y = 0) = 0.40(0.25) = 0.10$. Benzerce, $P(X = 0, Y = 1) = 0.30$ ve $P(X = 0)P(Y = 1) = 0.40(0.75) = 0.30$, $P(X = 500, Y = 0) = 0.15$ ve $P(X = 500)P(Y = 0) = 0.60(0.25) = 0.15$, $P(X = 500, Y = 1) = 0.45$ ve $P(X = 500)P(Y = 1) = 0.60(0.75) = 0.45$.

Bağımsız İki Rastgele Değişkenin Farkının ve Toplamının Aritmetik ortalamaları

Ugulamada, birden çok popülasyonu karşılaştırmak icap ettiği zaman, sıkça başvurduğumuz yöntem bu popülasyonlar arasındaki farkları analiz etmektir.

Misal 12 $X = \{2,9,11,22\}$ ve $Y = \{5,7,15\}$ kümelerine sahibiz. X kümesinin aritmetik ortalaması,

$$\mu_x = \frac{2 + 9 + 11 + 22}{4} = 11$$

ve Y kümesinin aritmetik ortalaması,

$$\mu_y = \frac{5 + 7 + 15}{3} = 9$$

Farkların W kümesi için, Y nin her elemanını X in her elemanından çıkarıyoruz,

$$W = \{2-5, 2-7, 2-15, 9-5, 9-7, 9-15, 11-5, 11-7, 11-15, 22-5, 22-7, 22-15\} = \{-3, -5, -13, 4, 2, -6, 6, 4, -4, 17, 15, 7\}$$

W nun aritmetik ortalaması nedir?

Cevap:

$$\mu_w = \frac{-3-5-13+4+2-6+6+4-4+17+15+7}{12} = \frac{24}{12} = 2$$

Dikkatini şuna celbetmek istiyorum ki

$$\mu_w = \mu_x - \mu_y = 11 - 9 = 2.$$

Genel olarak, bir farklar kümesinin aritmetik ortalaması, orijinal kümelerin aritmetik ortalamalarının farkına eşittir. Yine, bir toplamlar kümesinin aritmetik ortalaması, orijinal kümelerin aritmetik ortalamalarının toplamına eşittir. Hatta daha genel olarak, çeşitli kümelerden birer eleman almak suretiyle yeni bir toplamlar kümesi elde edilir ise, böylesi toplamların aritmetik ortalaması basitçe orijinal kümelerin aritmetik ortalamalarının toplamıdır.

Misal 13 İki rastgele değişken, *X* ve *Y*, müteakip ortak olasılık dağılımı tablosuna sahiptir:

		Y		
		5	8	9
X	3	0.10	0.05	0.15
	4	0.20	0.10	0.00
	6	0.05	0.20	0.15

X, *Y* ve *X*+*Y* rastgele değişkenlerinin beklenen değerleri (yani aritmetik ortalamaları) nedir?

Cevap:

$$\begin{aligned} E(X) &= \sum x_i P(x_i) \\ &= 3(0.10+0.05+0.15) + 4(0.20+0.10+0.00) + 6(0.05+0.20+0.15) \\ &= 4.50 \end{aligned}$$

$$\begin{aligned} E(Y) &= \sum y_j P(y_j) \\ &= 5(0.10+0.20+0.05) + 8(0.05+0.10+0.20) + 9(0.15+0.00+0.15) \\ &= 7.25 \end{aligned}$$

$$\begin{aligned} E(X+Y) &= \sum (x+y)_{ij} P(x,y)_{ij} \\ &= (3+5)0.10 + (3+8)0.05 + (3+9)0.15 + (4+5)0.20 + (4+8)0.10 \\ &\quad + (4+9)0.00 + (6+5)0.05 + (6+8)0.20 + (6+9)0.15 \\ &= 11.75 \end{aligned}$$

Görüldüğü üzere,

$$E(X+Y) = E(X) + E(Y).$$

Genel olarak, bu ifade, rastgele değişkenlerin toplamlarının beklenen değeri için geçerlidir. Aynı zamanda, şu da geçerlidir ki

$$E(X-Y) = E(X) - E(Y).$$

Bağımsız İki Rastgele Değişkenin Farkının ve Toplamının Varyansları

Misal 14 Misal 12 de, X ve Y kümelerinin mütekabil elemanlarının farklarından bir W kümesi oluşturmuştuk. W nun varyansı orijinal X ve Y kümelerinin varyanslarıyla nasıl ilişkilidir?

Cevap:

$$\sigma_x^2 = \frac{(2-11)^2 + (9-11)^2 + (11-11)^2 + (22-11)^2}{4} = \frac{206}{4} = 51.5$$

$$\sigma_y^2 = \frac{(5-9)^2 + (7-9)^2 + (15-9)^2}{3} = \frac{56}{3} = 18.67$$

$$\sigma_w^2 = \frac{(-3-2)^2 + (-5-2)^2 + (-13-2)^2 + \cdots + (7-2)^2}{12} = \frac{842}{12} = 70.17$$

Burada dikkat etmelisin ki

$$\sigma_w^2 = \sigma_x^2 + \sigma_y^2 = 51.5 + 18.67 = 70.17.$$

Genel olarak, bir farklar kümesinin varyansı, orijinal kümelerin varyanslarının toplamına eşittir.

$$\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2$$

Yine, bir toplamlar kümesinin varyansı, orijinal kümelerin varyanslarının toplamına eşittir.

$$\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$$

Daha genel olarak, çeşitli bağımsız kümelerden birer eleman almak suretiyle yeni bir toplamlar veya farklar kümesi elde edilir ise, böylesi toplamların veya farkların varyansı basitçe orijinal kümelerin varyanslarının toplamıdır.

Toplamın varyansını hesaplamak için varyansları toplayabildiğimiz gibi, bu işlemi tersine döndürüp, toplam varyansın, kendisini oluşturan çeşitli kaynaklarca nasıl paylaşıldığını da belirleyebiliriz. Buna *varyans dekompozisyonu* veya *varyansın parçalarına ayrılması* adı verilir. Mesela, bir şirketin satış elemanlarının yaptıkları satış sayılarının varyansının ne kadarının satış elemanları arasındaki bireysel farklılıktan kaynaklandığını, ne kadarının bölgeler arasındaki farklılıktan kaynaklandığını ve ne kadarının satılan ürün çeşitleri arasındaki farklılıktan kaynaklandığını bulabiliriz. Bu prensipler, daha geniş manada, rastgele değişkenler için geçerlidir.

Misal 15 Misal 13 te verilen, X ve Y rastgele değişkelerine ait ortak olasılık dağılımı tablosuna göre, X, Y ve $X+Y$ rastgele değişkenlerinin beklenen değerlerinden kareli sapmalarının beklenen değerleri, yani varyansları nedir?

Cevap:

$$\begin{aligned} E[(X-\mu_X)^2] &= \sum (x_i - \mu_X)^2 P(x_i) \\ &= (3-4.5)^2 0.3 + (4-4.5)^2 0.3 + (6-4.5)^2 0.4 \\ &= 1.65 \end{aligned}$$

$$\begin{aligned} E[(Y-\mu_Y)^2] &= \sum \left(y_j - \mu_Y\right)^2 P\left(y_j\right) \\ &= (5-7.25)^2 0.35 + (8-7.25)^2 0.35 + (9-7.25)^2 0.3 \\ &= 2.8875 \end{aligned}$$

$$E\left[\left((X+Y) - \mu_{X+Y}\right)^2\right] = \sum \left((x+y)_{ij} - \mu_{X+Y}\right)^2 P(x,y)_{ij}$$

$$= (8-11.75)^2 0.10 + (11-11.75)^2 0.05 + (12-11.75)^2 0.15$$
$$+ (9-11.75)^2 0.20 + (12-11.75)^2 0.10 + (13-11.75)^2 0.00$$
$$+ (11-11.75)^2 0.05 + (14-11.75)^2 0.20 + (15-11.75)^2 0.15$$
$$= 5.5875$$

Görüldüğü üzere,

$$E\left[\left((X+Y)-\mu_{X+Y}\right)^2\right] \neq E[(X-\mu_X)^2] + E[(Y-\mu_Y)^2].$$

Yani bu iki rastgele değişkenin toplamının beklenen değerinden kareli sapmasının beklenen değeri, orijinal değişkenlerin beklenen değerlerinden kareli sapmalarının beklenen değerlerinin toplamına eşit değil. Bir başka notasyonla, bu halde,

$$\sigma_{X+Y}^2 \neq \sigma_X^2 + \sigma_Y^2.$$

Yani, bu iki değişkenin toplamının varyansı, varyanslarının toplamına eşit değil.

Binnetice, varyansları aritmetik ortalamalar gibi her halde toplayamıyoruz. Müteakip misali, bağımsız iki değişken bağlamında mülahaza edelim.

Misal 16 İki rastgele değişken, *X* ve *Y*, şu ortak olasılık dağılımı tablosuna sahiptir:

		Y			
		2	5	10	
X	5	0.07	0.05	0.08	0.20
	7	0.28	0.20	0.32	0.80
		0.35	0.25	0.40	

X, *Y* ve *X*+*Y* rastgele değişkenlerinin beklenen değerlerinden kareli sapmalarının beklenen değerleri (yani varyansları) nedir?

Cevap: Önce, beklenen değerleri (yani aritmetik ortalamaları) hesaplayalım,

$$E(X) = \sum x_i P(x_i)$$
$$= 5(0.20) + 7(0.80) = 6.6$$

$$E(Y) = \sum y_j P(y_j)$$
$$= 2(0.35) + 5(0.25) + 10(0.40) = 5.95$$

$$E(X+Y) = E(X) + E(Y) = 12.55$$

Şimdi, beklenen değerlerden kareli sapmaların beklenen değerlerini (yani varyansları) hesaplayalım,

$$E[(X-\mu_X)^2] = \sum (x_i - \mu_X)^2 P(x_i)$$
$$= (5-6.6)^2 0.2 + (7-6.6)^2 0.8$$
$$= 0.64$$

$$E[(Y-\mu_Y)^2] = \sum \left(y_j - \mu_Y\right)^2 P(y_j)$$
$$= (2-5.95)^2 0.35 + (5-5.95)^2 0.25 + (10-5.95)^2 0.4$$
$$= 12.2475$$

$$E\left[\left((X+Y)-\mu_{X+Y}\right)^2\right]=\sum\left((x+y)_{ij}-\mu_{X+Y}\right)^2 P(x,y)_{ij}$$
$$=(7-12.55)^2 0.07+(10-12.55)^2 0.05+(15-12.55)^2 0.08$$
$$+(9-12.55)^2 0.28+(12-12.55)^2 0.20+(17-12.55)^2 0.32$$
$$=12.8875$$

Görüldüğü üzere,

$$E\left[\left((X+Y)-\mu_{X+Y}\right)^2\right]=E[(X-\mu_X)^2]+E[(Y-\mu_Y)^2].$$

Yani bu iki rastgele değişkenin toplamının beklenen değerinden kareli sapmasının beklenen değeri, orijinal değişkenlerin beklenen değerlerinden kareli sapmalarının beklenen değerlerinin toplamına eşit. Bir başka notasyonla, bu halde,

$$\sigma_{X+Y}^2=\sigma_X^2+\sigma_Y^2.$$

Yani, bu iki rastgele değişkenin toplamının varyansı, varyanslarının toplamına eşit. Eğer ortak olasılık dağılımı tablosunda tüm koşullu olasılıkları hesaplasaydın, görürdün ki her x ve y değeri için,

$$P(y|x)=P(y)$$

ve dolayısıyla X ve Y bağımsızdır.

Genel olarak, eğer iki rastgele değişken bağımsız ise, bu iki rastgele değişkenin toplamının varyansı, ayrı ayrı varyanslarının toplamına eşittir. Aynı zamanda, eğer iki rastgele değişken bağımsız ise, bu iki rastgele değişkenin farkının varyansı, ayrı ayrı varyanslarının toplamına eşittir. Yani,

$$E\left[\left((X+Y)-\mu_{X+Y}\right)^2\right]=E[(X-\mu_X)^2]+E[(Y-\mu_Y)^2]$$

ve

$$E\left[\left((X-Y)-\mu_{X-Y}\right)^2\right]=E[(X-\mu_X)^2]+E[(Y-\mu_Y)^2]$$

Bağımsız rastgele değişkenlerin varyanslarını toplayabilirsin, fakat standart sapmalarını toplayamazsın!

Misal 17 İstatistik dersinin birinci vize sınavından öğrencilerin aldıkları puanların aritmetik ortalaması 42.5 ve standart sapması 10 dur. İkinci vize sınavında aritmetik ortalama 47.5 ve standart sapma 10 dur. İki vize sınavında öğrencilerin aldıkları puanların toplamının aritmetik ortalaması ve standart sapması hakkında ne söylenebilir?

Cevap: Birinci ve ikinci vize puanlarının toplamının aritmetik ortalaması,

$$42.5+47.5=90.$$

Birinci ve ikinci vize puanlarının ayrı ayrı varyanslarını toplasak,

$$10^2+10^2=200$$

ve sonra karekökünü alsak,

$$\sqrt{200}=14.142,$$

birinci ve ikinci vize puanlarının toplamının standart sapmasına eşit olur mu? Olmaz, çünkü bağımsızlık mevcut değil. Şu sebeple ki birinci vizeden daha yüksek puan alan öğrenciler, genel olarak, ikinci vizeden de daha yüksek puan alma eğilimindelerdir ve daha düşük puan alanlar, genel olarak, daha düşük puan alma eğiliminde. Burada, toplam puanların standart sapması, verilen enformasyondan hesaplanabilmez.

Varyansların ve Standart Sapmaların Spesiyal Halleri

Gördük ki iki rastgele değişkenin farklarının aritmetik ortalaması her zaman aritmetik ortalamalarının farkına ve toplamlarının aritmetik ortalaması her zaman aritmetik ortalamalarının toplamına eşittir. Fakat bu eklenirlik kuralı, varyanslar için, yalnızca değişkenlerin bağımsızlığı halinde geçerlidir. Bu bölümün başında rastgele değişkenlerin bağımsızlığının ne manaya geldiğini tartıştık.

Eğer rastgele değişkenler bağımsız ise, bu değişkenlerin değerleri arasında bir çağrışım yoktur ve varyansları toplanabilir. Eğer X ve Y rastgele değişkenleri bağımsız değilse, varyanslarının toplanmasına mani olan şey bu iki değişken arasındaki çağrışımdır. Bu durumda, iki değişkenin toplamlarının varyansı, kendi bireysel varyanslarına bağlı olduğu kadar, ikisi arasındaki **korelasyona** da bağlıdır. Bölüm 6 da, aynı örnekteki bireyler üzerinde ölçülen iki değişken arasındaki r korelasyon katsayısını, standartlaştırılmış x ve y gözlemlerinin çarpımlarının aritmetik ortalaması olarak tanımlamıştık. İki rastgele değişken arasındaki koralasyon da aynı şekilde yine ağırlıklandırılmış bir ortalama olarak tanımlanır, ancak bu defa olanaklı değerler olasılıklar ile ağırlıklandırılır. Elbette burada daha fazla detayına girmeyeceğiz. Sadece, iki rastgele değişken arasındaki korelasyonun, örnek veriden hesaplanan r korelasyonu ile aynı temel özelliklere sahip olduğunu bilmen kafi. İki rastgele değişken arasındaki korelasyonu göstermek için ρ simgesini kullanırız. Yine, ρ korelasyonu, iki rastgele değişken arasındaki ilişkinin yönünü ve gücünü ölçen, -1 ve $+1$ arasında bir sayıdır. **Bağımsız iki rastgele değişken arasındaki korelasyon sıfırdır.**

Mini Proje 10

Bu bölümde, bir rastgele değişkenin alabileceği sayısal değerleri ve onların olasılıklarını bildiğin zaman, bu rastgele değişkenin beklenen değerini nasıl hesaplayacağını öğrendin. Bu enformasyonu kullanarak, insanların beklenen parasal kazançlarını maksimize etmelerini sağlayacak kararları verebileceklerini düşünür müydün? Fakat çoğu insan bu yönde hareket etmez. Eğer etselerdi, piyango biletlerini veya sigorta poliçelerini satınalmazlardı.

Rastgele seçeceğin 20 kişi ile, bir örnek yoklaması yürüt. Deneklere şunu sor: "Eğer aşağıdaki alternatif çiftleriyle karşılaşmış olsaydın, hangilerini tercih ederdin?"

Alternatif çiftleri	Tercihin?
(A) Garanti 240 TL kazanacaksın. (B) 1000 TL kazanma şansın %25 ve hiç bir şey kazanamama şansın %75.	A mı, B mi?
(C) Garanti 740 TL kaybedeceksin. (D) 1000 TL kaybetme şansın %25 ve hiç bir şey kaybetmeme şansın %75.	C mi, D mi?

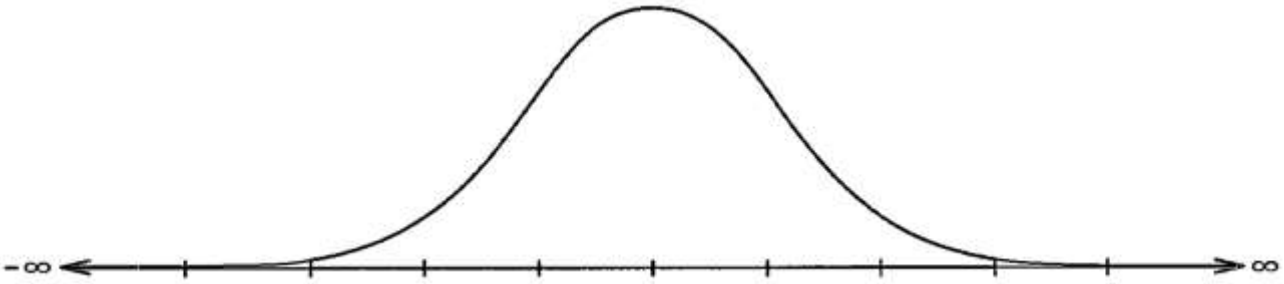
a. A yı B ye ve C yi D ye tercih edenlerin yüzdelerini hesapla.
b. Bulgularının interpretasyonunu yap. Öğrenme-çalışmanın sorusu bağlamında, ne manaya geliyor bu yüzdeler? Rakamların dilini herkesin anlayabileceği bir dile tercüme et.
c. Örnek yoklamanı nasıl yürüttüğünü açıkla. Veri toplamaya dair Bölüm 7 de tartışılan potansiyel zorluklarla nasıl başa çıktığını anlat.

BÖLÜM 11 | **NORMAL DAĞILIM**

Giriş

Daha evvel rastgele değişkenlerin kesikli ve sürekli olmak üzere iki tipinin olduğunu görmüştük ve kesikli olasılık dağılımlarından bazılarını öğrenmeye-çalışmıştık. Anımsayalım, eğer örnek uzayındaki çıkımları sürekli bir spektrum boyunca gerçekleşiyorsa, bir rastgele değişkeni sürekli olarak tanımlamıştık. Bir Y sürekli rastgele değişkenin olanaklı değerlerinin sayısı sonsuz olacağı için, Y nin olasılık dağılımı kesikli rastgele değişkenlerinkinden daha farklıdır. Temel fark Y nin bir y değerini alması olasılığının sıfır olmasıdır. İlaveten, kesikli olasılık dağılımlarına benzemezce, sürekli olasılık dağılımları düzgün (yani pürüzsüz) bir eğridir. Y nin örnek uzayındaki belirli bir aralıktaki herhangi bir değeri alması olasılığı, o aralıkta olasılık dağılımı eğrisinin altında kalan alana eşittir.

En kullanışlı sürekli olasılık dağılımlarının bazıları simetrik ve çan eğrisi şeklindedir. Çan eğrisi uçlardan basık ortası şişkin bir şekildir ve merkez etrafında yığılmış çok sayıda gözlemlerden ve dağılımın her iki ucuna yayılmış daha az saydaki gözlemlerden oluşur. Bir çan eğrisi dikey olarak tam ortasından (yani merkezinden) ikiye bölündüğü zaman biribirinin yansıması iki parça oluşturur.

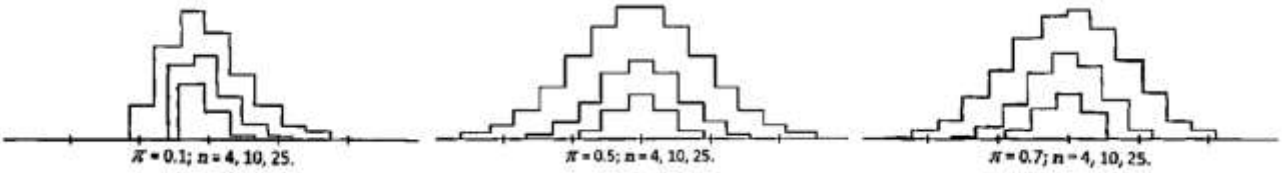


Bu bölümde böylesi olasılık dağılımlardan biri olan normal dağılımı konu edeceğiz. Simetrik ve tek-tepeli dağılımların tümünün normal olmadığını en baştan söylemeliyim. Çan eğrisi şeklindeki normal dağılım eğrisi iki parametreli (μ ve σ) bir matematiksel fonksiyonla tanımlanır:

$$f(y|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Normal dağılım eğrisi sürekli sayısal değerler ile gözlenen çeşitli doğal fenomenleri betimlemek için paha biçilmezdir. Mesela ağırlık, uzunluk ve sıcaklık gibi bir çok değişkenin frekans dağlımları, örnek hacmi arttıkça mükemmel birer çan eğrisi şekline yaklaşır. (Bununla beraber, istatistik biliminde normal dağılımın gerçek önemi, örnekleme prosedürlerinin sonuçlarını betimlemek için kullanılabiliyor olmasındandır.)

Normal dağılımı binomiyal dağılımın bir limiti olarak düşünebiliriz. Daha spesifik olarak, herhangi bir sabit π başarı olasılığı ile başlayıp, n keyfi olarak artırıldığı zaman, binomiyal olasılık dağılımının şekli bakımından ne olacağını mülahaza edelim. Bu limitin kısmi görsel bir temsilini temin etmek niyetiyle, aşağıda her bir sabit π değeri (0.1, 0.5 ve 0.7) için artan n değerleriyle (4, 10 ve 25) sırasıyla içeriden dışarıya doğru binomiyal histogramlar çizilmiştir.



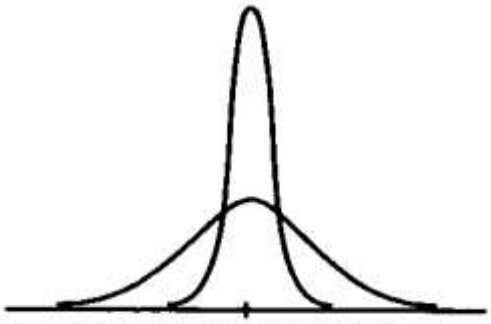
Diyagramın anlaşılabilirliğini artırmak maksadıyla, n in farklı değerleri için çizilen binomiyal histogramlar farklı boyutlarda ölçeklendirilmiştir ki aslında bunların gerçekte aynı büyüklükte olmaları gerektiği açıktır. Yukarıdaki diyagramlara baktığımızda, $\pi = 0.5$ olduğu zaman, n hangi

değeri alırsa alsın, binomiyal dağılımın şekli ile normal dağılımın şeklinin biribirine çok benzediğini görebiliyoruz. Yine, π nin farklı değerleri söz konusu olduğu zaman, mantığın gereği olarak kabul edilmelidir ki n yeterince büyüdükçe ortaya çıkan binomiyal histogramların şekli düzgün bir çan eğrisine, yani normal dağılım eğrisine yaklaşma temayülü göstermektedir.

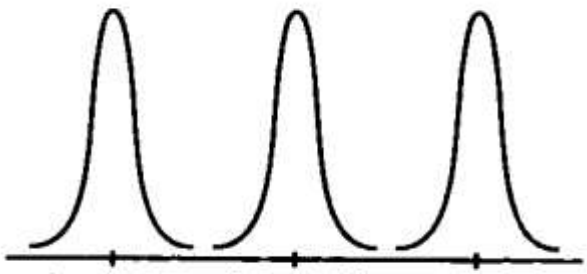
Normal Dağılımın Özellikleri

Normal dağılım eğrisi çan şeklinde ve simetriktir ve sonsuz bir tabana oturur. Uzun, düz görünümlü iki kuyruğu çok sayıda değeri kapsar, fakat eğri altındaki toplam alanın yalnızca çok küçük bir orantısıdır. Kuyrukların düz görüntüsü yanıltıcıdır. Gerçekte, kuyruklar uzadıkça, eğri artan bir orantısallıkla aşağı doğru gider. Başka bir deyişle, eşit genişlikte iki aralığı karşılaştırdığımız zaman, merkeze yakın olanı daha büyük bir sayısal düşüşe sahip olurken, kuyruğa yakın olanı sayısal olarak daha küçük bir düşüşe ve fakat aralığın başlangıç noktasının yüksekliğine orantısal olarak daha büyük bir düşüşe sahip olur.

Normal dağılım eğrisi verinin aritmetik ortalamasının değerine göre yatay eksen üzerinde yer değiştirirken, standart sapmasının farklı değerlerine göre de dikleşir veya basıklaşır. Bu manada biribirinden farklı normal dağılımlar söz konusudur. Ama, hangisi olursa olsun, bir normal dağılımın aritmetik ortalaması ortancasına eşittir ve eğrinin merkezinde yer alır.



Aritmetik-ortalamaları aynı fakat standart sapmaları farklı normal eğriler



Standart sapmaları aynı fakat aritmetik-ortalamaları farklı normal eğriler

$Y \sim N(\mu, \sigma)$ notasyonu Y değişkeninin aritmetik ortalaması μ ve standart sapması σ olan bir normal dağılıma sahip olduğunu ifade eder.

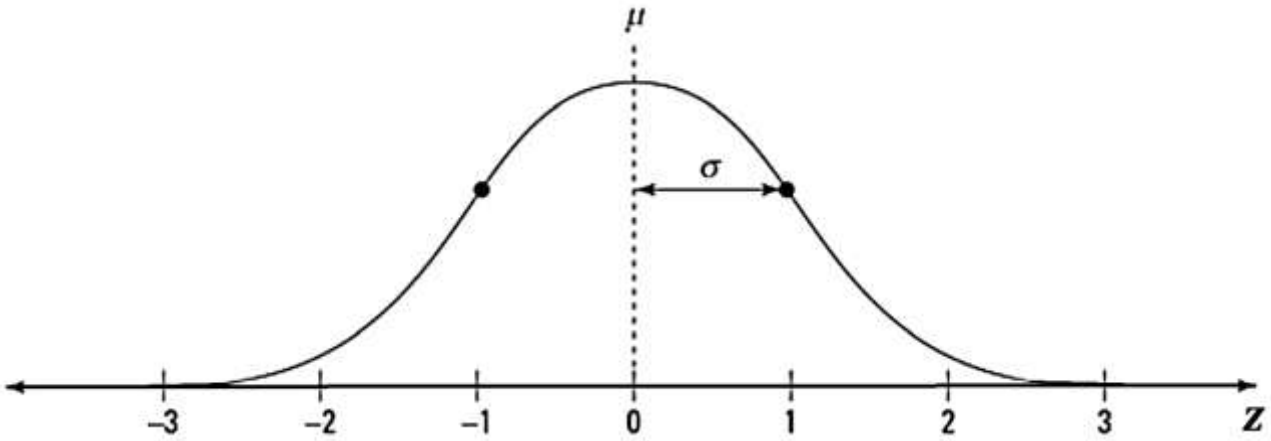
Bir nispi frekans grafiğini göz önüne alalım, dikey eksen ölçeğinde her bir sınıfa düşen verinin orantısı kaydedilmiştir. Yani bir frekans dağılımı grafiğinde dikdörtgenlerin alanı nispi frekanslarla orantılıdır. Bir ferkans dağılımını normal eğri ile betimlediğimizde de eğrinin altında kalan alan nispi frekanslara eşittir. Nispi frekans grafiğinde nispi frekansların alanı her zaman 1 e (yani %100 e) eşit olduğu gibi, normal eğrinin altında kalan toplam alan 1 dir.

Normal dağılım fonksiyonunun orijini Fransız matematikçi Abraham de Moivre (1667-1754) ye kadara izlenebilir. Bu adamın kumara karşı bilimsel bir alakası vardı ve pek çok zaman kumarbazlara olasılıkların belirlenmesine dair danışmanlık yapardı. Yine para atışlarının olasılık dağılımını öğrenmeye-çalışıyordu. Mesela 100 atışta 60 veya daha çok sayıda yazı gelmesi gibi olasılıkları hesaplamaya yarayacak bir matematiksel ifade bulmakla uğraşıyordu. Bu soruya bir cevap olarak çan şekilli dağılımı (binomiyal dağılımın bir limiti olarak) türetti ki bu dağılım müştereken normal eğri olarak adlandırılmıştır. Bunun daha öncesinde, İtalyan astronom Galileo di Vincenzo Bonaiuti de Galilei (1564-1642) astronomik gözlemlerde enstrüman ve gözlemci kusurlarından dolayı meydana gelen hataların dağılımının simetrik olduğunu ve küçük hataların büyük hatalara göre daha yüksek frekanslarla ortaya çıkma eğiliminde olduğunu göstermişti zaten. Mamafih, bu hataların normal dağılıma uyduğunu göstermek ancak biribirlerinden bağımsızca bu fonksiyonu yeniden keşfeden matematikçiler İrlandalı-Amerikalı Robert Adrain (1775-1843) ve Alman Carl Friedrich Gauss (1777-1855) a nasip oldu. Fransız Pierre-Simon Laplace (1749-1827) çok ünlü ve tesirli merkezi limit teoremini geliştirmeye çalışırken, yine bağımsızca aynı dağılımı keşfetti ve kullanım alanının genişlemesine katkıda bulundu. Belçikalı bilimci Lambert Adolphe Jacques Quetelet (1796-1874)

normal dağılımı insan karakteristiklerine ilk defa uyguladı ve mesela boy, ağırlık ve güç gibi karakteristiklerin dağılımlarının normal dağılıma uyduğuu gösterdi.

Standart Normal Dağılım

Normal dağılım eğrisinin bir matematiksel fonksiyon tarafından tanımlandığını söyledik. Fakat bu fonksiyonu, istatistiksel problemlerin çözümü için μ ve σ parametrelerinin alacağı farklı değerlere göre her defasında yeniden hesaplamak biraz külfetlidir. Bu sebeple, öyle bir ölçüm birimi isteriz ki herhangi bir normal dağılıma en güzel şekilde uygulanabilsin. Bu ölçüm birimi doğal olarak eğrinin şeklinden neşet eder. Eğrinin her iki yanında birer nokta vardır ki bu noktalarda eğim en diktir. Bu iki nokta *büküm noktaları* olarak adlandırılır ve aritmetik ortalamanın bu noktaların her birine olan uzaklığı tam olarak 1 standart sapmaya eşittir. Böylece, normal eğri altındaki mesafeleri *z*-puanları cinsinden ölçmek elverişlidir (Bölüm 2 den anımsayacağın üzere *z*-puanları gözlem değerlerinin aritmetik ortalamadan uzaklıklarının standart sapma cinsinden miktarlarıdır; yani, bir gözlem değerinin içinde yer aldığı gözlem değerleri kümesinin aritmetik ortalamasından kaç standart sapma uzakta olduğunu söyler bize).



Ham değerler aritmetik ortalamadan standart sapmalarına dönüştürüldükten, yani *z*-puanlarına dönüştürüldükten sonra, matematiksel olarak, normal eğri yatay eksen üzerinde sonsuz sayıdaki *z*-puanlarının bir fonksiyonu olarak, dikey eksende mütekabil $f(z)$ değerlerine dönüşür. Bu fonksiyonu

$$f(z) = e^{-z^2/2}$$

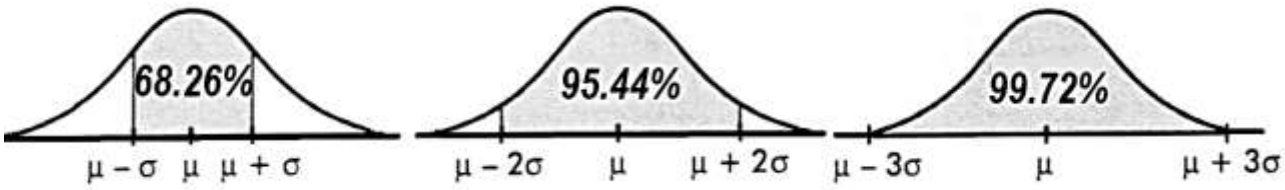
formunda ifade edebiliriz ki burada $f(z)$ bir *z*-puanı üzerindeki nispi yükseklikti*r*. *Nispi yükseklik* derken, aritmetik ortalamanın üzerindeki yüksekliğin orantısını kastediyoruz. Ancak, bizim ilgimiz normal eğri altındaki nispi yüksekliklerden daha çok, eğrinin altındaki *orantısal alanlara* yönelir. Normal eğri altındaki herhangi bir aralığa ilişkin olasılık, eğrinin altında o aralığa tekabül eden alanın eğrinin altındaki alanın bütününe orantısıdır.

$N(0,1)$ ile temsil edilen standart normal dağılım, normal dağılım ailesinin $\mu = 0$ ve $\sigma = 1$ olan üyesidir. Standart normal dağılım için belirli *z*-puanlarına ait aralıkların olasılıkları tablo haline getirilmiştir. Herhangi bir rastgele değişkenin dağılımı normale uyduğu zaman, bu dağılıma ait ham değerleri mütekabil *z*-puanlarına dönüştürmek suretiyle standart normal dağılım tablosunu kullanarak arzu edilen olasılıklar hesaplanabilmektedir. Tablo B de bir standart normal dağılım tablosu (veya *z*-tablosu) mevcuttur ki bir *z*-puanı ve $-\infty$ aralığının olasılığını vermektedir.

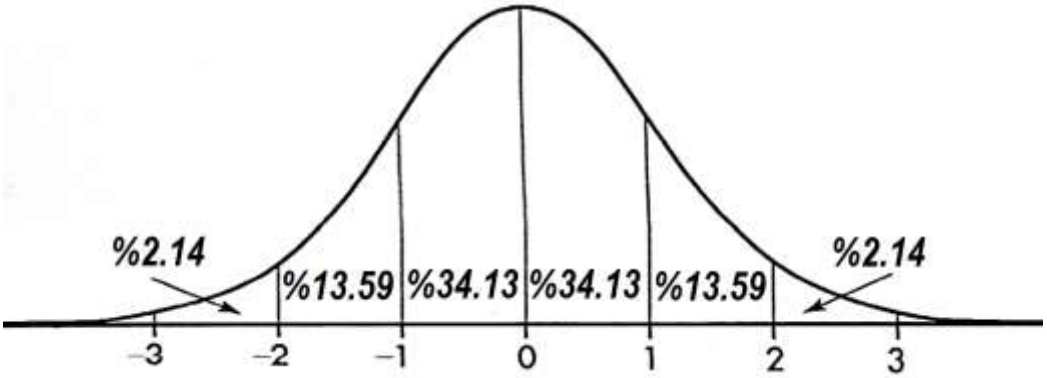
Daha yukarıdaki normal dağılım fonksiyonunda μ yerine sıfır ve σ yerine 1 yazarsan, yukarıdaki standart normal dağılım fonksiyonunu elde edersin.

Normal eğrinin altındaki alanın %68.26 sı (ve böylece gözlemlerin %68.26 sı) aritmetik ortalamanın birer standart sapma civarına düşer, %95.44 ü (ve böylece gözlemlerin %95.44 ü) aritmetik

ortalamanın ikişer standart sapma civarına düşer ve %99.72 si (ve böylece gözlemlerin %99.72 si) aritmetik ortalamanın üçer standart sapma civarına düşer. Daha önce buna ampirik kural adı verildiğini ve bu yüzdelerin kabaca 68-95-99.7 olarak ifade edildiğini öğrenmiştik.

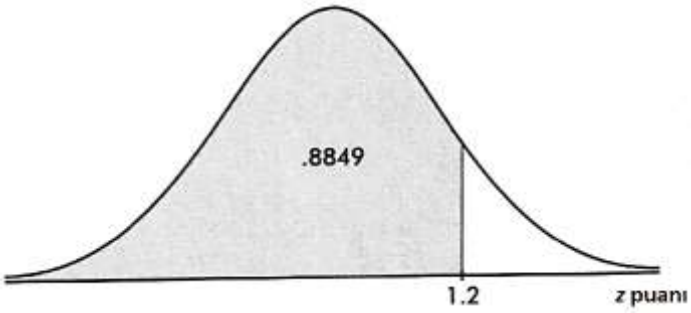


Standart normal dağılım için bu olasılıkların daha tafsilatlı bir taksimi de şöyledir:

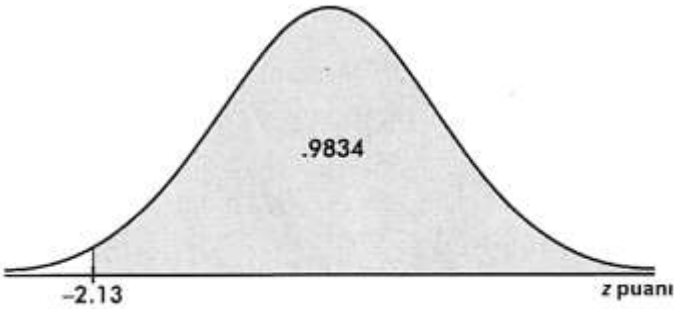


Normal Dağılıma Ait Tabloların Kullanımı

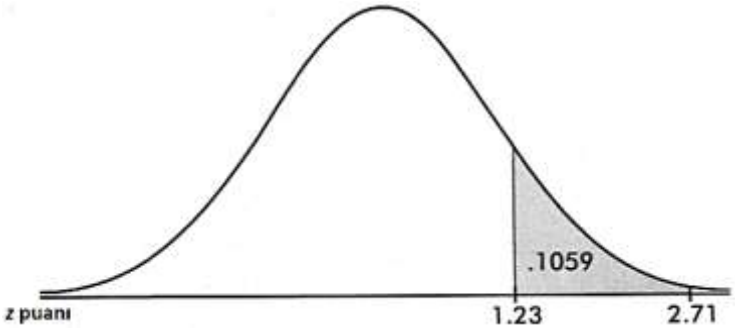
Tablo B *z*-puanları için normal eğri altındaki orantısal alanları gösteriyor. Spesifik olarak, bu tablo yatay eksen üzerinde belirli bir *z*-puanının soluna düşen alanı göstermektedir. Mesela, aşağıdaki diyagrama bakarsan 1.2 lik *z*-puanının solundaki alanın orantısı 0.8849 dur:



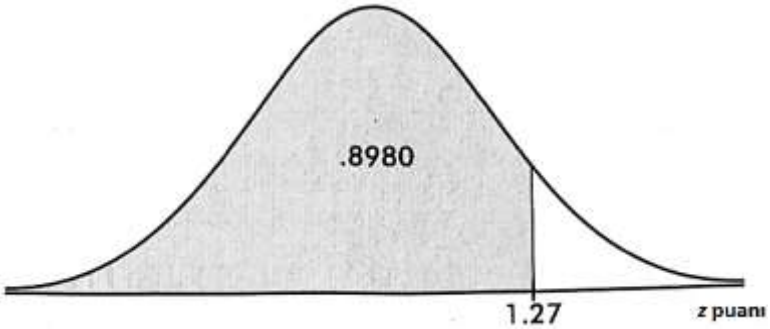
Eğrinin altındaki alanın toplamı 1 olduğuna göre, yatay eksen üzerinde bir *z*-puanının sağına düşen alanı hesaplamak için bu noktanın soluna düşen alanı 1 den çıkartırız. Mesela, –2.13 lük *z*-puanının sağına düşen alanın orantısı $1 - 0.0166 = 0.9834$ tür. (Alternatif olarak, simetriden dolayı, –2.13 değerinin sağındaki alanın +2.13 değerinin solundaki alana eşit olduğunu söyleyebiliriz ve buna göre cevabı doğrudan Tablo B den okuyabiliriz.)



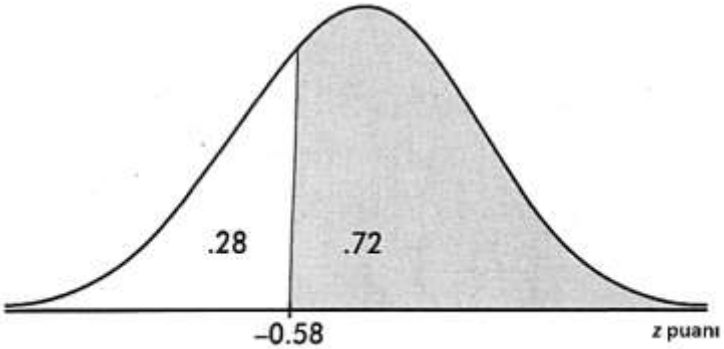
Yatay eksen üzerindeki iki *z*-puanı arasındaki alan bir çıkartma işlemiyle bulunabilir.Mesela, eğrinin 1.23 ve 2.71 *z*-puanları arasındaki alanı şöyle hesaplarız: $0.9966 - 0.8907 = 0.1059$.



Bir olasılık verildiği zaman, Tablo B yi tersinden kullanarak bu olasılığa tekabül eden z-puanını bulabiliriz. Mesela, solundaki alan 0.8982 olan z-puanını bulmak için, tablo içeriğinde 0.8982 ye en yakın olasılığı araştırırız ve tekabül eden z-puanını okuruz. z-puanının tam kısmını ve onda birlik kısmını tablonun solundaki ilk sütundan yüzde birlik kısmını ise üstündeki ilk satırdan okuruz. Bu durumda, 0.8982 olasılığına tablodaki en yakın değer 0.8980 dir ve tekabül eden z-puanı $1.2 + 0.07 = 1.27$ dir.



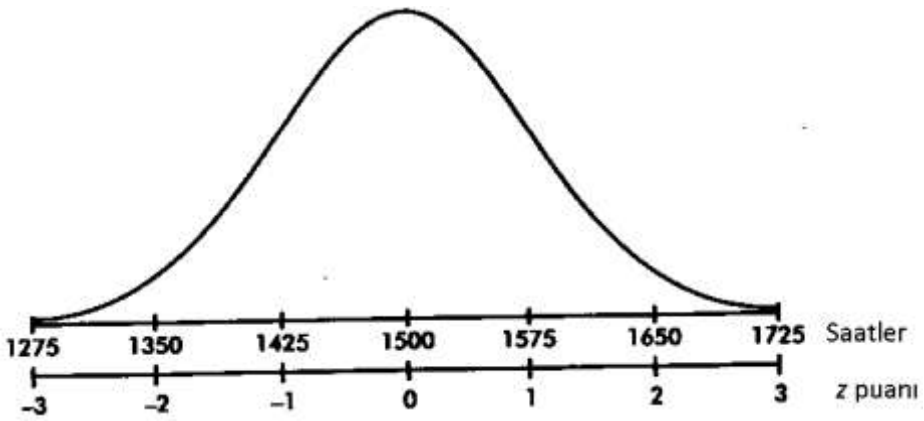
Sağındaki alan 0.72 olan bir z-puanını bulmak için, tabloda $1 - 0.72 = 0.28$ olasılık değerini veya ona en yakın değeri ararız. Tabloda 0.2810 olasılık değerine –0.58 z-puanı karşılık gelmektedir.



Ölçümlerin Bir Modeli Olarak Normal Dağılım

Bir problemi çözmek maksadıyla normal dağılımı bir model olarak kullanmak istediğimiz zaman, yatay düzlemde ham değerleri ve z-puanlarını beraber gösterdiğimiz bir normal eğri çizmek genellikle daha yararlıdır.

Misal 1 Belirli bir marka ampulün ömrü 1500 saatlik aritmetik ortalama ve 75 saatlik standart sapma ile normal dağılmaktadır.

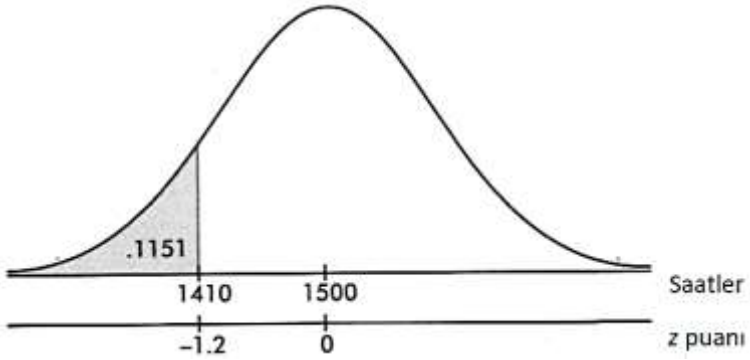


a. Rastgele seçilecek bir ampülün ömrünün 1410 saaten az olması olasılığı nedir?

Cevap: Önce, 1410 saat değerini z-puanına dönüştürelim:

$$z = \frac{1410 - 1500}{75} = -1.2$$

Sonra, Tablo B den -1.2 lik z-puanının solundaki alanı bulalım. Aradığımız olasılık olarak, bu alanın değeri 0.1151 dir.

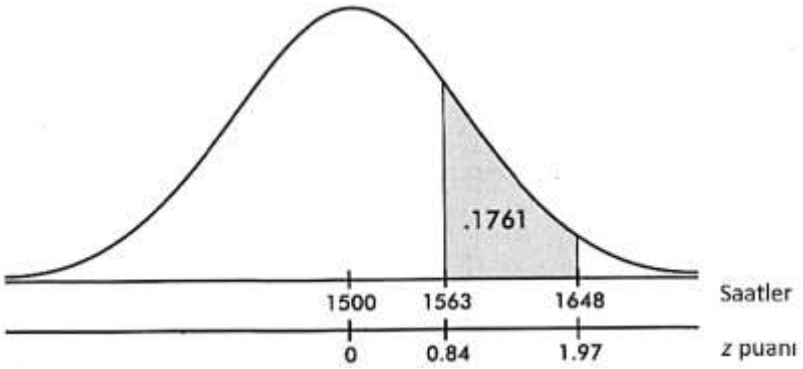


b. Rastgele seçilecek bir ampulün ömrünün 1563 saat ve 1648 saat arasında olması olasılığı nedir?

Cevap: 1563 saat ve 1648 saat değerlerini z-puanlarına dönüştürelim:

$$z_1 = \frac{1563 - 1500}{75} = 0.84 \qquad z_2 = \frac{1648 - 1500}{75} = 1.97$$

Tablo B de 0.84 lük z_1 ve 1.97 lik z_2 puanlarının herbirinin solundaki alanları sırasıyla 0.7995 ve 0.9756 olarak buluruz. Buna göre, rastgele seçilecek bir mpulün ömrünün 1563 saat ve 1648 saat arasında olması olasılığı $0.9756 - 0.7995 = 0.1761$ dir.

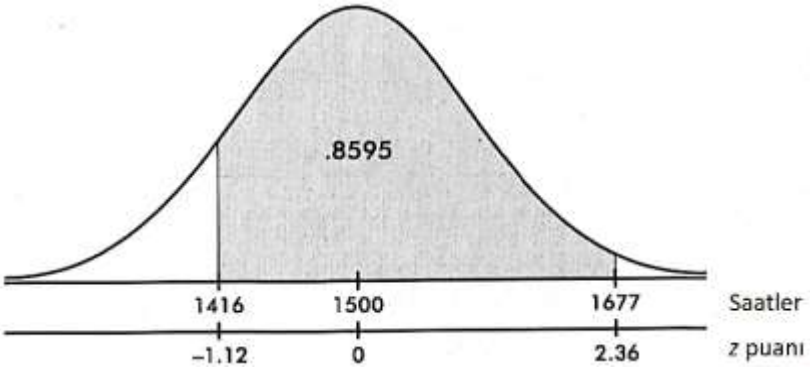


c. Rastgele seçilecek bir ampulün ömrünün 1416 saat ve 1677 saat arasında olması olasılığı nedir?

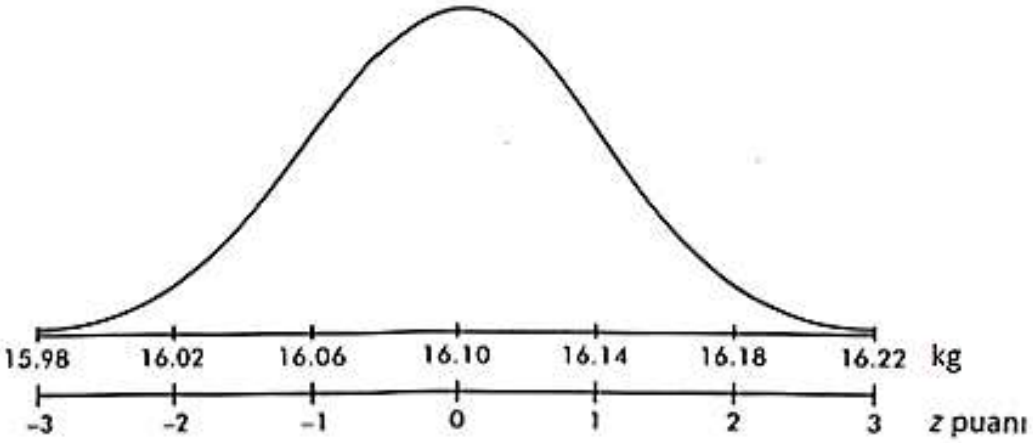
Cevap: 1416 saat ve 1677 saat değerlerini z-puanlarına dönüştürelim:

$$z_1 = \frac{1416 - 1500}{75} = -1.12 \qquad z_2 = \frac{1677 - 1500}{75} = 2.36$$

Tablo B de -1.12 lik z_1 ve 2.36 lık z_2 puanlarının herbirinin solundaki alanları sırasıyla 0.1314 ve 0.9909 olarak buluruz. Buna göre, rastgele seçilecek bir ampulün ömrünün 1416 saat ve 1677 saat arasında olması olasılığı $0.9909 - 0.1314 = 0.8595$ tir.



Misal 2 Bir paketleme makinası her bir torbaya 16.1 kg pirinç doldurmak üzere kurulmuştur. Torbalardaki pirinç miktarının standart sapması 0.04 kg olan bir normal dağılıma sahip olduğunu farzedelim.

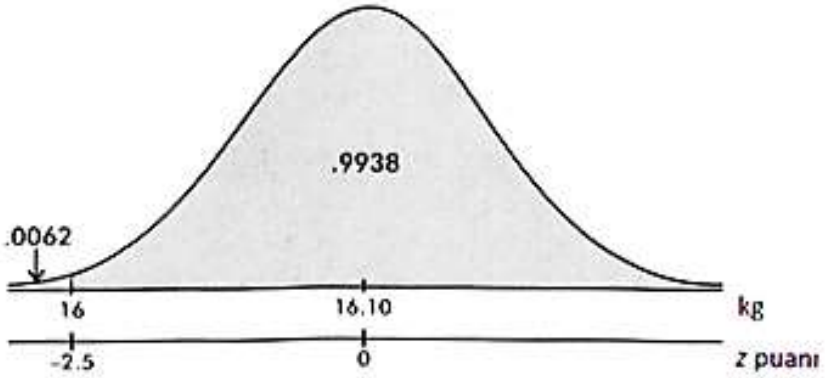


a. Torbaların yüzde kaçındaki pirinç miktarı en az 16 kg olacaktır?

Cevap: 16 kg değerini *z*-puanına dönüştürelim:

$$z = \frac{16 - 16.1}{0.04} = -2.5$$

Tablo B de -2.5 luk *z*-puanının solundaki alanı 0.0062 olarak buluruz. Buna göre aradığımız olasılık $1 - 0.0062 = 0.9938$ dir. Hesapladığımız olasılığı %99.38 olarak ta yazabiliriz.



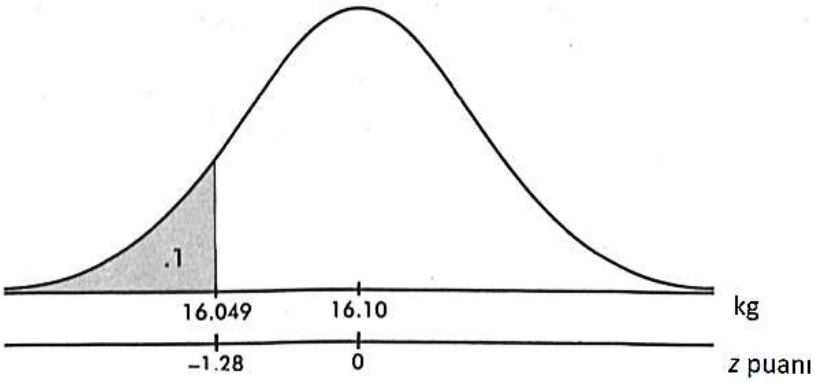
b. Torbaların %10 undaki pirinç miktarı kaç kilogramdan daha az olacaktır?

Cevap: Tablo B ye baktığımızda görüyoruz ki 0.1 lik (aslında 0.1003 lük) alan -1.28 lik z-puanının solundaki alandır. Şimdi z-puanını ham puana dönüştürelim:

$$-1.28 = \frac{y - 16.1}{0.04}$$

eşitliğinde gerekli değerleri yerlerine yazdıktan sonra,

$$\begin{aligned} y &= -1.28(0.04) + 16.1 \\ &= 16.049 \text{ kg.} \end{aligned}$$



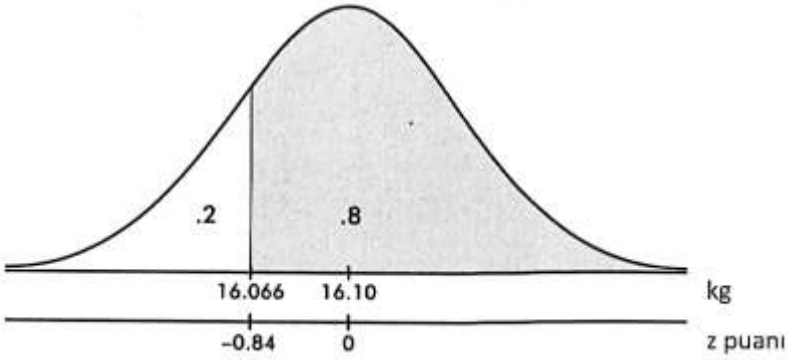
c. Torbaların %80 indeki pirinç miktarı kaç kilogramdan daha çok olacaktır?

Cevap: Tablo B ye baktığımızda görüyoruz ki $1 - 0.8 = 0.2$ lik (aslında 0.2005 lik) alan -0.84 lük z-puanının solundaki alandır ve böylece -0.84 lük z-puanının sağındaki alan eğri altındaki toplam alanın %80 idir. Bulduğumuz -0.84 lük z-puanını ham puana dönüştürelim:

$$-0.84 = \frac{y - 16.1}{0.04}$$

ve bu eşitlikten,

$$\begin{aligned} y &= -0.84(0.04) + 16.1 \\ &= 16.066 \text{ kg.} \end{aligned}$$



d. Torbaların ortadaki %90 ı hangi ağırlıklar arasında olacaktır?

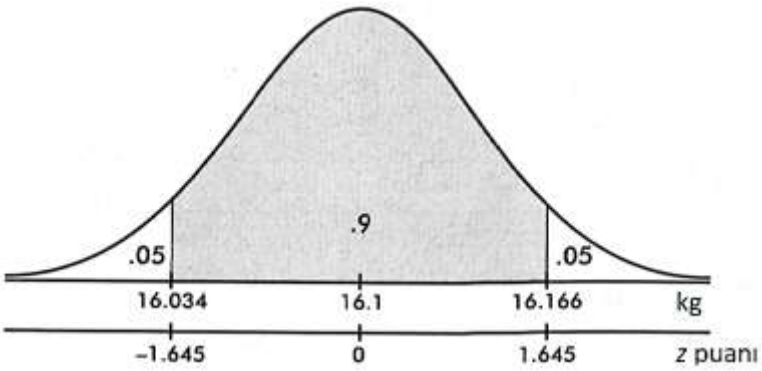
Cevap: Ortadaki yüzde doksanlık alan her bir kuyrukta %5 lik iki alan bırakır. Tablo B de 0.0505 lik alanın -1.64 lük z-puanının solundaki alan olduğunu görüyoruz. Aynı şekilde 0.495 lik alan -1.65 lik z-puanının solundaki alandır. Standart normal eğri altında ortaya düşen %90 lık alanın simetrik olması gerektiğini biliyoruz. Böylece alt sınır olarak -1.645 lik z-puanını ve üst sınır olarak ta 1.645 lik z-puanını kullanacağız. Bulduğumuz z-puanlarını ham puanlara dönüştürürsek, alt sınırı,

$$\begin{aligned} y_{Alt} &= -1.645(0.04) + 16.1 \\ &= 16.034 \text{ kg.} \end{aligned}$$

ve üst sınırı,

$$\begin{aligned} y_{Üst} &= 1.645(0.04) + 16.1 \\ &= 16.166 \text{ kg.} \end{aligned}$$

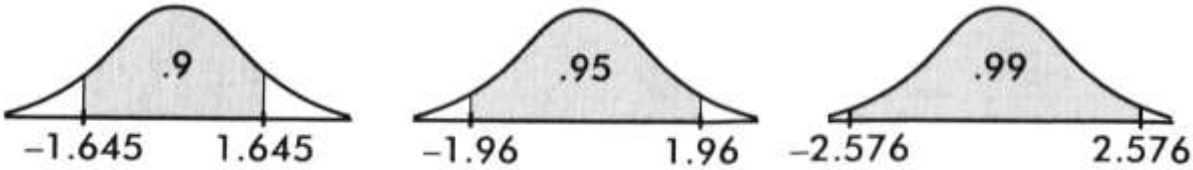
olarak hesaplarız.



Eğer elindeki verinin dağılımı simetrik ve tek-tepeli değilse normal modeli kullanma!

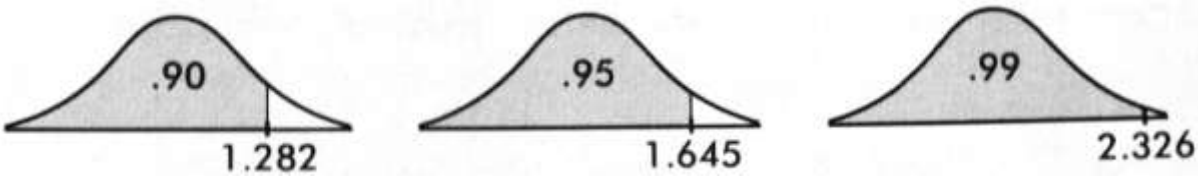
Müştereken Kullanılan Olasılıklar ve Mütekabil *z*-Puanları

Misal 2 nin d. şıkkında görüldüğü gibi çoğu zaman verinin ortaya düşen belirli bir yüzdesini sınırlayan değerleri bulmak arzu edilir. İleride kolaylık sağlaması bakımından en çok kullanılan sınır değerleri *z*-puanları cinsinden aşağıda not edilmiştir.



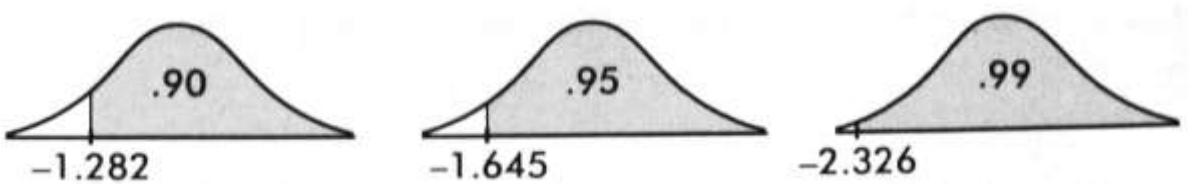
Değerlerin yüzde doksanı −1.645 ve 1.645 lik *z*-puanları arasındadır, değerlerin yüzde doksan beşi −1.96 ve 1.96 lık *z*-puanları arasındadır ve değerlerin yüzde doksan dokuzu −2.576 ve 2.576 lık *z*-puanları arasındadır.

Bazen belirli yüzdeliklerin değerlerini bulmak arzu edilir. Mesela,



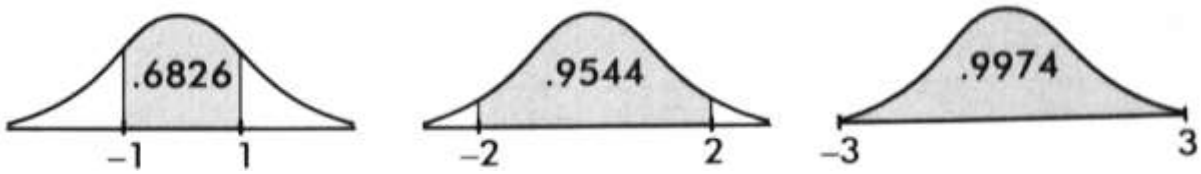
Değerlerin yüzde doksanı 1.282 lik *z*-puanının altındadır, değerlerin yüzde doksan beşi 1.645 lik *z*-puanının altındadır ve değerlerin yüzde doksan dokuzu 2.326 lık *z*-puanının altındadır.

Negatif *z*-puanları için mütekabil neticeler söz konusudur:



Değerlerin yüzde doksanı −1.282 lik *z*-puanının üstündedir, değerlerin yüzde doksan beşi −1.645 lik *z*-puanının üstündedir ve değerlerin yüzde doksan dokuzu −2.326 lık *z*-puanının üstündedir.

Bazı tamsayılı *z*-puanları arasına düşen değerlerin yüzdelerini not etmek te bazen yararlı olabilmektedir.



Daha evvel de 68-95-99.7 kuralında ifade edilmiş olduğu üzere, değerlerin %68.26 sı −1 ve +1 lik

z-puanları arasındadır, değerlerin %95.44 ü -2 ve $+2$ lik z-puanları arasındadır ve değerlerin %99.74 ü -3 ve $+3$ lük z-puanları arasındadır.

Misal 3 Belirli bir yörede yetişkin erkeklerin boy aritmetik ortalaması 170 cm ve standart sapması 6.5 cm olsun.

a. Eğer dağılım normal ise, erkeklerin ortadaki %95 i hangi iki boy uzunluğu arasındandır?

Cevap: Daha evvel not edildiği üzere, bu durum için kritik z-puanları ± 1.96 dır ve böylece boy için iki sınır değeri aritmetik ortalamadan 1.96 standart sapma uzaklıktaki değerler olacaktır. Buna göre alt sınır,

$$170 - 1.96(6.5) = 157.26 \text{ cm}$$

ve üst sınır,

$$170 + 1.96(6.5) = 182.74 \text{ cm.}$$

b. Boyların yüzde doksanı hangi değerin altındadır?

Cevap: Bu durumda kritik z-puanı 1.282 dir ve böylece aradığımız değer,

$$170 + 1.282(6.5) = 178.33 \text{ cm.}$$

c. Boyların yüzde doksan dokuzu hangi değerin üstündedir?

Cevap: Bu durumda kritik z-puanı -2.326 dır ve böylece aradığımız değer,

$$170 - 2.326(6.5) = 154.88 \text{ cm.}$$

d. Boyların ± 1, ± 2 ve ± 3 lük z-puanları arasına düşen yüzdeleri sırasıyla kaçtır?

Cevap: Sırasıyla %68.26, %95.44 ve %95.74.

Aritmetik ortalamaları ve Standart sapmaları Bulma

Eğer bir dağılımın normal olduğunu biliyorsak, μ yü ve σ yı hesaplamak için popülasyonun yüzdelerini kullanabiliriz.

Misal 4 Aritmetik ortalaması 25 olan bir popülasyon veriliyor. Eğer tüm olanaklı gözlem değelerinin %18 inin 29 un üstünde olduğunu biliyorsak, standart sapma kaçtır? Dağılımın normale uyduğunu farzediyoruz.

Cevap: Tablo B de 0.82 lik olasılığa karşılık gelen z-puanı 0.92 dir. Buna göre, 0.92 lik standart sapmaya,

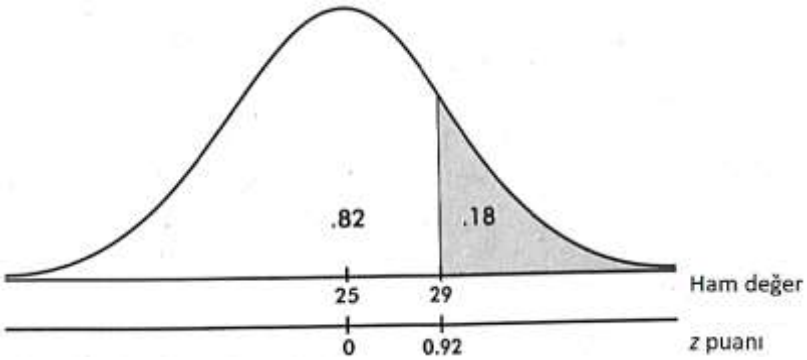
$$29 - 25 = 4$$

ham değeri karşılık gelmektedir. Yani,

$$0.92\sigma = 4$$

ve

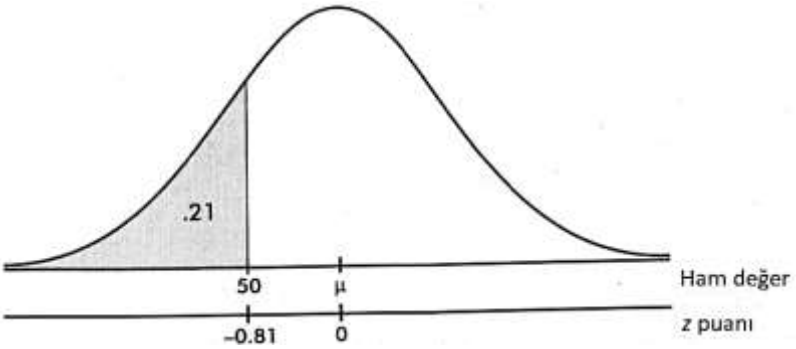
$$\sigma = \frac{4}{0.92} = 4.35.$$



Misal 5 Standart sapması 10 olan bir normal dağılım veriliyor, eğer değerlerin %21 i 50 den küçükse, aritmetik ortalama kaçtır?

Cevap: Tablo B de 0.21 olasılığına tekabül eden z-puanı -0.81 dir. Böylece 50 değerinin aritmetik ortalamaya olan uzaklığı -0.81 lik standart sapmaya tekabül eder. Buna göre,

$$\mu = 50 + 0.81(10) = 58.1.$$



Misal 6 Bir normal dağılıma ait değerlerin %80 i 125 in üstünde ve %90 ı 110 un üstündedir. Aritmetik ortalama ve standart sapma kaçtır?

Cevap: Tablo B den kritik z-puanlarını -0.84 ve -1.28 olarak buluruz. Böylece,

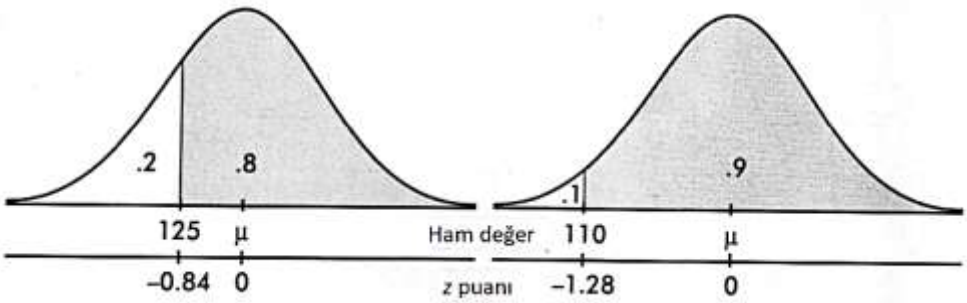
$$\frac{125-\mu}{\sigma} = -0.84 \qquad \text{ve} \qquad \frac{110-\mu}{\sigma} = -1.28$$

denklemlerini elde ederiz.

$$125 - \mu = -0.84\sigma$$
$$110 - \mu = -1.28\sigma$$

denklem sistemini eş zamanlı çözersek,

$$\mu = 153.64 \quad \text{ve} \quad \sigma = 34.09.$$



Binomiyale Normal Yaklaşım

Binomiyal dağılımın pek çok uygulamasında örnek hacmi n sonsuza gider. Böyle olmakla beraber, büyük n durumunda binomiyal olasılıkların hesaplanması pek güçleşir. Normal dağılım binomiyal dağılımın bir limiti olarak görülebileceği için, bu durumda normal dağılımı kullanarak binomiyal dağılıma yaklaşmak doğal bir çözümdür.

Binomiyal dağılım yalnızca tam sayı değerler alır. Buna karşın, normal dağılım aralıklar üzerindeki alanlara tekabül eden olasılıklarla süreklidir. Bu yüzden, bir dağılımdan diğerine dönüştürmemiz gerekir. Bir yaklaşım olarak şöyle yaparız. Her binomiyal olasılık, arzu edilen tam sayı değeri ortalayan birim aralığın normal olasılığına tekabül eder. Böylece, mesela, 8 başarının binomiyal olasılığına yaklaşmak için 7.5 ve 8.5 arasındaki normal olasılığı belirleriz.

Misal 7 Belirli bir deterjan markası kutularının %15 inde hediye çeki var. Kırk kutu deterjan satın alan bir kişiye tam olarak 5 hediye çeki çıkması olasılığı nedir?

Cevap: Binomiyal cevap şudur:

$$\left(\frac{40!}{35!\,5!}\right)(0.15)^5(0.85)^{35}.$$

Fakat, hesaplaması biraz meşakkatli görünüyor. Normal dağılımı kullanarak cevaba yaklaşmak için, önce aritmetik ortalama ve standart sapmayı hesaplayalım:

$$\mu = n\pi = 40(0.15) = 6$$

$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{40(0.15)0.85} = 2.258.$$

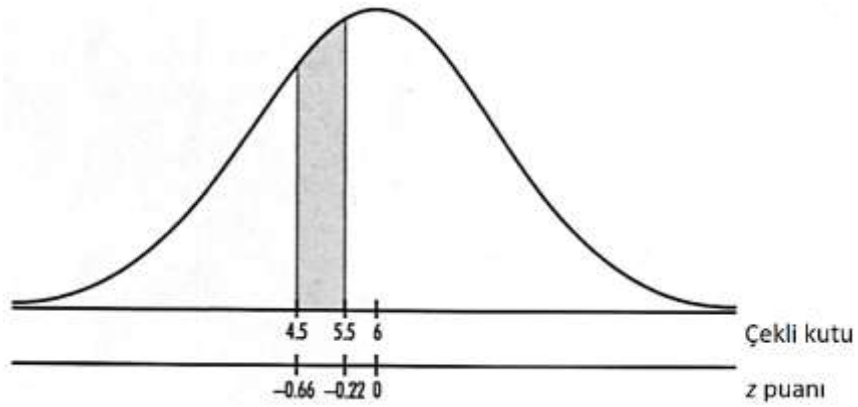
Sonra mütenasip z-puanlarını hesaplarız:

$$\frac{4.5-6}{2.258} = -0.66 \qquad \text{ve} \qquad \frac{5.5-6}{2.258} = -0.22.$$

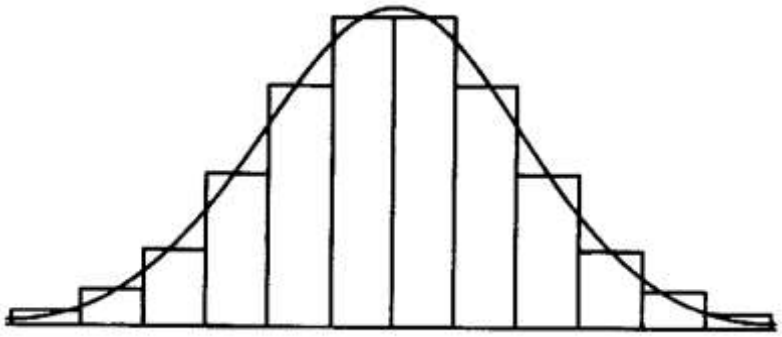
Tablo B de mütekabil olasılıkları bularak nihai cevabımıza ulaşırız:

$$0.4129 - 0.2546 = 0.1583.$$

(Binomiyal dağılımı kullanmış olsaydık, eğer, gerçek cevap 0.1692 olacaktı. İki sonuç neden farklı? Çünkü bir yaklaşımda bulunduk.)



Normalin binomiyale yaklaşımını kullanmak, birden çok aralığın olasılığı söz konusu olduğu zaman daha yararlı hale gelir.



Misal 8 Eğer popülasyonun %60 ı mevcut hükümetin vergi politikalarını destekliyor ise, 250 kişi ile yapılan bir yoklamada en çok 155 kişinin koyulan vergileri destekliyor olması olasılığı nedir?

Cevap: Gerçek cevap 156 adet binomiyal ifadenin toplamıdır:

$$(0.4)^{250} + \cdots + \frac{250!}{155!\,95!}0.6^{155}0.4^{95}$$

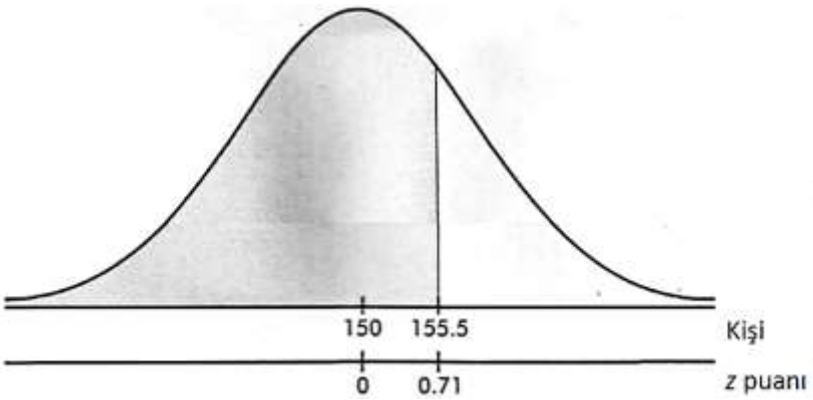
Mamafih, normal dağılımı kullanmak suretiyle, daha çabukça ve daha kolayca iyi bir yaklaşımda bulunabiliriz. Aritmetik ortalamayı ve standart sapmayı hesaplayalım:

$$\mu = n\pi = 250(0.6) = 150$$
$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{250(0.6)0.4} = 7.746$$

En çok 155 başarının binomiyali 155.5 un altının normal olasılığına tekabül eder. 155.5 ham değerinin z-puanı olarak karşılığı,

$$\frac{155-150}{7.746}=0.71$$

Tablo B ye baktığımızda nihai cevabı 0.7611 olarak buluruz. (Binomiyal dağılımı kullanmış olsaydık, eğer, gerçek cevap 0.7605 olacaktı.)



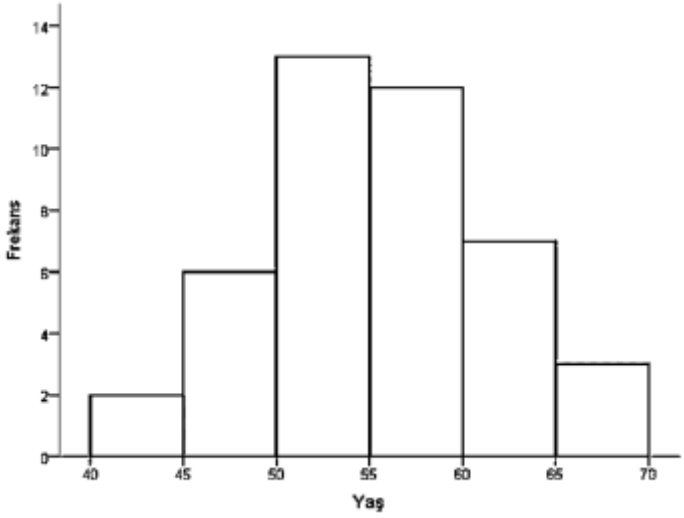
Binomiyal dağılıma normal dağılımla yaklaşmak iyi midir? Cevap, elbette, belirli durumlardaki hata toleranslarına bağlıdır. Bir genel kural olarak, hem $n\pi$ hem de $n(1-\pi)$ değerleri 10 dan büyük ise binomiyale normalle yaklaşmak iyidir.

Normalliğin Kontrolü

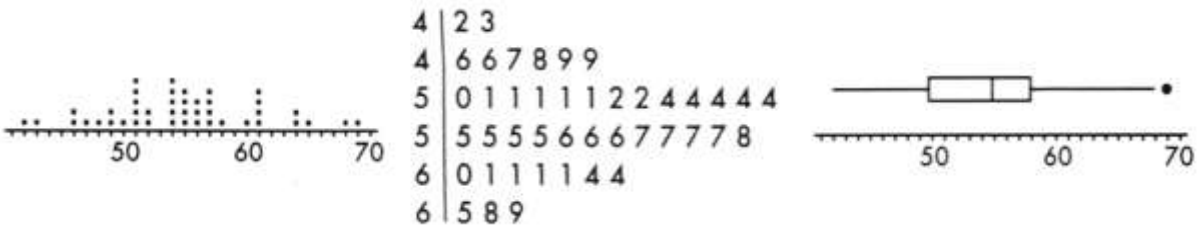
Bir öğrenme-çalışmasının verisi toplandığı zaman, veri analizi sfhasında normal modele dayalı prosedürleri kullanmadan önce verinin normal bir popülasyondan gelip gelmediğine karar vermek gerekir. Sonraki bölümlerde işleyeceğimiz veri analizi tekniklerinin bir çoğu verinin dağılımının normale uyduğunu varsaydığı için, normalliğin kontrolü bu bakımdan da önemlidir. Ancak, şunu aklından çıkarmamalısın ki bir örneğin dağılımının bu örneğin çekildiği popülasyonun dağılımını gerçekten yansıtabilmesi örnek hacmi ile ilişkilidir. Yani yeterince geniş hacimli bir örnek yoksa elinde, bu örnekten elde edeceğin verinin frekans dağılımı, popülasyonun dağılımına benzemeyebilir.

Misal 9 Bir ülkede bakanlık yapmış 43 kişinin makama çıkma yaşları şöyledir: {57, 61, 57, 57, 58, 57, 61, 54, 68, 51, 49, 64, 50, 48, 65, 52, 56, 46, 54, 49, 51, 47, 55, 55, 54, 42, 51, 56, 55, 51, 54, 51, 60, 61, 43, 55, 56, 61, 52, 69, 64, 46, 54}. Dağılımın kabaca normal olduğu neticesine varabilir miyiz?

Cevap: Verinin dağılımını bir histogramla betimleyebiliriz:



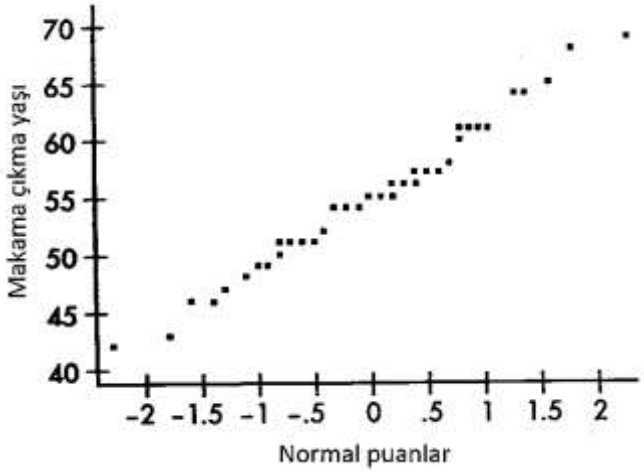
Frekans dağılımını görmek için, alternatif olarak nokta-grafiğini, dal-yaprak grafiğini ve kutu grafiğini de çizebilirdik:



Grafiklerin tümü bize dağılımın tek-modlu ve simetrik olduğunu ve bir çan eğrisine benzediğini gösteriyor.

Normal dağılımı sayısal ölçülerle de test edebiliriz: Elimizdeki verinin aritmetik ortalaması 54.81 yıl, ortancası ise 54.67 yıldır. Bu iki merkez ölçüsünün bir normal dağılımda, dağılımın simetrik yapısı gereğince, biribirine eşit olması gerektiğini biliyoruz. Öyleyse, bu bakımdan, makama çıkma yaşı verisinin dağılımı normal dağılıma oldukça yakındır. Çünkü verinin aritmetik ortalaması ve ortancası biribirine çok yakın.

Normalliği kontrol etmeye yarayan başka bir özel grafiksel gösterim *normal olasılık grafiği*dir. Verinin dağılımı yaklaşık normal olduğu zaman, bu grafik sol alt köşeden sağ üst köşeye uzanan kabaca düz bir çizgidir. Bu grafik, bir histograma nazaran, normallikten sapmaları daha belirgince göstermekle beraber, onun kadar kolay anlaşılabilir değildir. Öte yandan, normal olasılık grafiğini elle hesaplamak ve çizmek zordur. Mamafih, teknolojiyi kullanırsan eğer, mesela bir istatistiksel yazılım paketi ile çizmek kolaydır. Misal 9 daki verinin normal olasılık grafiği şudur:



Normal dağılım gerçek dünyada ölçülen değişkenlerin bir çoğunun dağılımı için güzel bir modelidir. Mesela, bir çok kişiye uygulanan eğitimle ilgili veya psikolojik testler, aynı fiziksel niceliğin tekrarlayan ölçümleri veya bir biyolojik popülasyonun bireyleri üzerinde tekrarlanan ölçümler gibi. Diğer bazı değişkenlerin dağılımı ise normal eğri ile modellenemez. Bunlar genelde sağa veya sola doğru ağır çarpık veya çok-tepeli dağılımlara sahiplerdir. Mesela bir ülkedeki kişilerin yıllık geliri normal bir dağılıma sahip değildir, sağa ağır çarpıktır.

Sürekli Rastgele Değişkenler için Beklenen Değerler

Önceki bölümlerde, örnek verideki "merkez" ve "yayılım" nosyonlarının, örnek verinin içinden çekildiğini varsaydığımız kesikli rastgele değişkenlere nasıl uzatıldığını gördük. Bilhassa, bir kesikli rastgele değişkenin beklenen değeri (ki aynı zamanda aritmetik ortalama deriz) bir örnek verinin aritmetik ortalaması fikrinin bir uzantısıydı. Ve benzerce, bir kesikli rastgele değişkenin kendi aritmetik ortalamasından kareli sapmasının beklenen değeri (ki aynı zamanda varyans deriz) bir örnek verinin varyansı fikrinin bir uzantısıydı.

Bu, sürekli rastgele değişkenler için de doğrudur. Herhangi bir Y sürekli rastgele değişkeni için, Y nin beklenen değeri (yani aritmetik ortalaması) o değişkenin merkezinin bir ölçüsüdür. Eğer Y nin

örnekleri çekilirse, bu örneklerin aritmetik ortalamaları Y nin aritmetik ortalamasının etrafında salınacaktır. Bir sürekli rastgele değişkenin beklenen değeri (veya aritmetik ortalaması), aynı kesikli rastgele değişkenlerinki gibi, çoğu zaman $E(Y)$ veya μ ile gösterilir.

Benzerce, herhangi bir Y sürekli rastgele değişkeni için, Y nin kendi aritmetik ortlamasından kareli sapmasının beklenen değeri (yani varyansı) o değişkenin yayılımının bir ölçüsüdür. Eğer Y nin örnekleri çekilirse, bu örneklerin varyansları Y nin varyansının etrafında salınacaktır. Bir sürekli rastgele değişkenin kendi aritmetik ortlamasından kareli sapmasının beklenen değeri (veya varyansı), aynı kesikli rastgele değişkenlerinki gibi, çoğu zaman $E[(Y-\mu_Y)^2]$ veya σ^2 ile gösterilir.

Bir sürekli rastgele değişkenin beklenen değerleri nosyonu kesikli rastgele değişkenlerinki ile aynı olmakla beraber, matematiksel tanımları farklıdır. Burada bu ölçülerin matematiksel tanımlarına girmeyeceğiz. Öğrenmeye-çalıştığımız yegane sürekli olasılık dağılımı olan normal dağılım için, beklenen değelerin hesaplanması bizim önceliğimiz olmayacak. Mamafih, kesikli olasılık dağılımlarının beklenen değerlerine ilişkin kurallar sürekli olasılık dağılımları için de aynen geçerlidir. Şöyle ki herhangi bir Y sürekli rastgele değişkeni için,

$$\mu_{Y\pm a} = \mu_Y \pm a$$

$$\mu_{bY} = b\mu_Y$$

$$\sigma^2_{Y\pm a} = \sigma^2_Y$$

$$\sigma^2_{bY} = b^2\sigma^2_Y$$

ve X, Y bağımsız sürekli rastgele değişkenleri için,

$$\mu_{X\pm Y} = \mu_X \pm \mu_Y$$
$$\sigma^2_{X\pm Y} = \sigma^2_X + \sigma^2_Y$$

kuralları geçerlidir.

Misal 10 Bir makine elemanı bir çubuk ve bunun iki ucundaki başlıklardan mürekkeptir. Bu üç parça bağımsızca üretilmekte ve bunlar kendi içlerinde bir miktar değişkenlik göstermekte. Çubuğun uzunluğu 12 santimetrelik aritmetik ortalamaya ve 0.004 milimetrelik standart sapmaya sahip. Başlığın uzunluğu 2 santimetrelik aritmetik ortalamaya ve 0.001 milimetrelik standart sapmaya sahip. Mevzubahis makine elemanının toplam uzunluğunun aritmetik ortalaması ve standart sapması nedir?

Cevap: Elemanların aritmetik ortalama uzunluğunu şöyle beyan edebiliriz:

$$\mu_{eleman} = \mu_{çubuk} + 2\mu_{başlık}$$
$$\mu_{eleman} = 12 \text{ cm} + 2(2 \text{ cm})$$
$$\mu_{eleman} = 16 \text{ cm}$$

Parçaların üretimi biribirinden bağımsız olduğuna göre, elemanların varyansını şöyle beyan edebiliriz:

$$\sigma^2_{eleman} = \sigma^2_{çubuk} + 2\sigma^2_{başlık}$$
$$\sigma^2_{eleman} = 0.004^2 + 2(0.001^2)$$
$$\sigma^2_{eleman} = 0.000018 \text{ mm}$$

ve elemanların standart sapması,

$$\sigma_{eleman} = \sqrt{0.000018 \text{ mm}} = 0.004243 \text{ mm}$$

Misal 11 Misal 10 dan devamla, makine parçalarının boyutlarının çoğunlukla kabaca normal dağılım gösterdiği biliniyor.

a. Çubukların %95 inin uzunluğu hangi $12 \pm t_{çubuk}$ ve başlıkların %95 inin uzunluğu hangi $2 \pm$

$t_{başlık}$ aralığındadır? Hesaplayacağın $\pm t_{çubuk}$ ve $\pm t_{başlık}$ değerleri genellikle parçaların "doğal toleransları" olarak adlandırılır.

Cevap: Bir normal dağılımda aritmetik ortalamanın etrafındaki %95 lik alanın sınırları $\mu \pm 1.96\sigma$ dır. Böylece, çubukların %95 inin uzunluğu şu aralıktadır,

$$12 \pm t_{çubuk} = 12 + 1.96\sigma_{çubuk} = 12 \pm 1.96(0.04 \text{ cm}) = 12 \pm 0.0784 \text{ cm}$$

Başlıkların %95 inin uzunluğu şu aralıktadır,

$$2 \pm t_{başlık} = 1.96\sigma_{başlık} = 2 \pm 1.96(0.01 \text{ cm}) = 2 \pm 0.0196 \text{ cm}$$

b. İstatistiksel teori bize söylemekte ki bağımsız herhangi normal rastgele değişkenlerin toplamı yine bir normal dağılıma sahiptir. Öyleyse, elemanın toplam uzunluğu da kabaca normaldir. Toplam uzunluğun doğal toleransı nedir?

Cevap: Misal 11 de elemanların aritmetik ortalamasını 16 cm ve standart sapmasını 0.004243 mm olarak hesapladık. Buna göre, toplam uzunluğun doğal toleransı,

$$\pm 1.96(0.04243 \text{ cm}) = \pm 0.000832 \text{ cm}.$$

Dikkat edersen, elbette ki $t_1 + 2t_2 = 0.0784 + 2(0.0196) = 0.1176$ cm değil.

Mini Proje 11

MP 01 de topladığın veri kümesinde yer alan nicel değişkenlerin nispi frekans dağılımlarının normal dağılıma uyup uymadığını olanaklı tüm vasıtaları kullanmak suretiyle denetle.

BÖLÜM 12 | ÖRNEKLEME DAĞILIMLARI

Giriş

İlgilendiğimiz bireylerin tümünün kümesi bir popülasyondur. Bir örnek ise popülasyonun bir parçasıdır ki popülasyonu temsil etmek için kullanırız. Popülasyon aritmetik ortalaması μ, popülasyon standart sapması σ ve popülasyon orantısı π popülasyon parametreleri için birer misaldir. Bunları ancak sözkonusu popülasyona ait tüm olanaklı gözlemlerden hesaplayabiliriz. Örnek aritmetik ortalaması $\bar{y}$, örnek standart sapması s ve örnek orantısı p örnek istatistikleri için birer misaldir. Bunları sözkonusu popülasyondan çektiğimiz bir örnekten hesaplarız.

Bir örnek istatistiğini o örneğin içinden çekildiği popülasyon parametresinin bir *kestirmi* olarak kullanırız. Popülasyon parametresi sabit (aynı zamanda meçhul) bir değer olmakla beraber, örnek istatistiği seçilecek örneğe göre değişebilir.

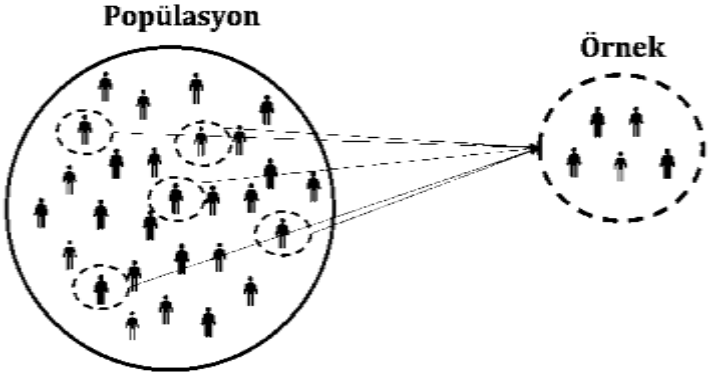
Bir istatistiğin nasıl değiştiğini gösteren olasılık dağılımına, *örnekleme dağılımı* adı verilir. Eğer örnekleme dağılımının aritmetik ortalaması ilişkin olduğu popülasyonun parametresine eşitse, örnekleme dağılımı *yansızdır*.

Bir örnekleme dağılımı, bir örneğin dağılımı ile aynı şey değildir.

İnsanların ortalama vücut ağırlığı nedir? Yüksek kolesterole sahip insanların orantısı nedir? Bir popülasyon hakkında önemli bir hakikati bilmek istiyoruz, fakat bu hakikat fiilen bilinebilir değildir. Yapabileceğimiz; mümkün olduğunca geniş hacimde ve mümkün olduğunca temsil kabiliyeti yüksek bir grup bireyden dikkatlice veri toplamak ve bu verinin sağladığı enformasyonu popülasyon parametresinin değerini kestirmek için kullanmaktır. Peki, hakikate ne kadar yakın olabiliriz? Biliyoruz ki farklı örnekler farklı kestirimler verebilecektir ve böyle olunca, *örnekleme hatası* kaçınılmazdır. Harika olan ve bu bölümde öğreneceğin şey, bu örnekleme hatasını niceleyebiliyor (yani sayısal olarak ölçebiliyor) olabildiğimizdir. Şunlara benzer ifadelerde bulunabiliyoruz: "kadınların aritmetik ortalama ağırlığı %99 olasılıkla 61 kg $\pm$ 4 kg arasındadır" veya "yüksek kolesterolü olan insanların orantısı %95 olasılıkla %41 $\pm$ %3 arasındadır".

Örnek Aritmetik ortalamasının Örnekleme Dağılımı

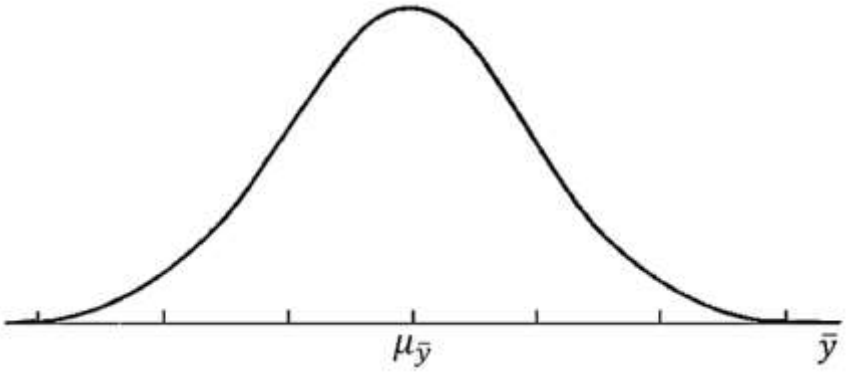
Farzedelim ki bir popülasyonun aritmetik ortalamasını kestirmek istiyoruz. Kestirimimiz için basitçe popülasyondan bir gözlem değerini rastgele seçebiliriz, yani $n = 1$. Fakat böyle yaparsak, cevabımıza çok az güvenebiliriz. Böyle yapmak yerine, farzedelim ki popülasyondan 100 gözlem değerini çekiyoruz ve aritmetik ortalamasını hesaplıyoruz, yani $n = 100$. Şunu sezgisel olarak biliriz ki çektiğimiz daha geniş hacimli bir örneğin aritmetik ortalamasının popülasyonun aritmetik ortalamasına daha yakın olması şansı, popülasyonun tek bir üyesininkinden daima daha fazladır.



Bir örnek çekersek ve $\bar{y}$ aritmetik ortalamasını hesaplarsak, popülasyondaki bireylerin bir kısmını rastgele örneklemiş olacağımıza göre, tüm olanaklı örnek aritmetik ortalamaları uzayından bir

aritmetik ortalamayı rastgele seçmiş oluruz. Tek bir örneğin aritmetik ortalamasının anlamlılığını değerlendirmek için, örnek aritmetik ortalamalarının nasıl değiştiğini bilmemiz gerekir.

Aynı hacimdeki tüm olanaklı örneklerin aritmetik ortalamaları kümesini göz önüne alalım. Gayet aşikar ve makul ki $\bar{y}$ örnek aritmetik ortalamalarının dağılımı, bu aritmetik ortalamalar popülasyonunun aritmetik ortalaması $\mu_{\bar{y}}$ etrafında öbekleşecektir ve dahası örnek aritmetik ortalamalarının bu öbekleşmesi popülasyona ait orjinal gözlem değerlerininkinden daha sıkı olacaktır (Niçin daha sıkı olacaktır? Sebep?). Esasen, çekilen örneğin hacmi ne kadar geniş olursa, öbekleşme de o kadar sıkı olacaktır. (Zaten, biliyoruz ki örnek aritmetik ortalaması popülasyon aritmetik ortalamasının yansız bir kestirimidir.) Eğer orjinal popülasyonun dağılımı normal (yani, çan eğrisi şeklinde) olsaydı, bu popülasyondan çekilecek aynı hacimdeki örneklerin aritmetik ortalamalarının dağılımın da normal olmasını beklemez miydin?



Aritmetik ortalamalar kümesinin $\sigma_{\bar{y}}$ standart sapmasını nasıl hesaplarız? Varyansı σ^2 olan bir popülasyondan hacmi *n* olan bir örnek çekiyoruz. Örnek aritmetik ortalamasını hesaplarken, önce *n* tane gözlem değerini topluyoruz ve sonra *n* e bölüyoruz. Toplamların kümesine ait varyans, matematiksel olarak, orijinal gözlem değerleri kümelerine ait varyansların toplamına eşittir. Bu durumda,

$$\sigma^2_{Toplam} = \sigma^2 + \cdots + \sigma^2 = n\sigma^2.$$

Bir kümenin her elemanı bir sabite bölünürse, matemtiksel olarak, yeni varyans eski varyansın sabitin karesine bölümüdür. Örnek aritmetik ortalamaları, toplamların *n* e bölünmesiyle elde edildiğine göre, matematiksel olarak, aritmetik ortalamaların varyansı da toplamların varyansını n^2 ye bölmek suretiyle elde edilir. Böylece, örnek aritmetik ortalamalarının, $\sigma^2_{\bar{y}}$ ile simgelediğimiz varyansını bulmuş oluruz,

$$\sigma^2_{\bar{y}} = \frac{\sigma^2_{Toplam}}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Standart sapma cinsinden ifade edersek,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}.$$

Örnekleme dağılımı teorisini, daha iyi anlaşılabilmesi bakımından, bir örnek aritmetik ortalamasının örnekleme dağılımının elde edilmesi üzerinden şöyle senaryolaştırabiliriz:

- Elimizde bir popülasyon var ve bu popülasyonun aritmetik ortalaması μ ve standart sapması σ olarak veriliyor.
- Bu popülasyondan çektiğimiz hacmi *n* olan tüm olanaklı örneklerin aritmatik-ortalamalarını hesaplıyoruz.
- Aritmetik ortalamalar kümesinin aritmetik ortalaması $\mu_{\bar{y}}$ popülasyonun aritmetik ortalaması μ ye eşit çıkar. Aritmetik ortalamalar kümesinin standart sapması $\sigma_{\bar{y}}$ ise popülasyonun standart sapmasının örnek hacminin kareköküne bölümüne eşit çıkar, yani $\sigma/\sqrt{n}$.

Bu bir senaryo, çünkü rastgele bir örneğin aritmetik ortalamasının örnekleme dağılımını elde etmek için, aynı hacimdedeki tüm olanaklı örnekleri çekmenin pratik bir gerçekliği yok elbette. Kaldı ki bir

popülasyonun aritmetik ortalaması, diğer bütün parametreleri gibi, sabit ama çoğu zaman pratikte meçhuldur.

Şuna dikkatini çekmek isterim ki örnek aritmetik ortalamaları kümesinin varyansı orijinal popülasyonun varyansı ile düz ve örneklerin hacmi ile ters yönde değişirken, örnek aritmetik ortalamaları kümesinin standart sapması orijinal popülasyonun standart sapması ile düz ve örneklerin hacminin kare kökü ile ters yönde değişir. Mesela, belirli bir insan topluluğunda, rastgele seçilen bir genç kızın santimetre cinsinden boyu $N(166.3,\ 6.3)$ dağılımına göre değişir, yani $Y \sim N(166.3, 6.3)$. Bu toplulukta, rastgele seçilen 100 genç kızın boyunun aritmetik ortalaması ise $N(166.3,\ 0.63)$ dağılımına göre değişir, yani $\bar{Y} \sim N(166.3, 0.63)$. Şöyle ki:

$$\mu_{\bar{y}} = \mu = 166.3$$

ve

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{6.3}{\sqrt{100}} = 0.63$$

Netice itibarıyla, bir genç kızın boyu popülasyon aritmetik ortalaması etrafında çok fazla değişirken, 100 genç kızdan oluşan örneğin aritmetik ortalaması popülasyon aritmetik ortalaması etrafında bunun onda biri kadar bir değişkenliğe sahiptir (bu misale münhasıran).

Misal 1 Belirli bir cinse ait elmaların ortalama ağırlığı 150 gr ve standart sapması 30 gramdır. Bu elmalar bir manavda her birinde 12 elma bulunan kutularda satılmaktadır. Eğer müşteriler satın aldıkları her kutunun içindeki elmaların ortalama ağırlığını belirleselerdi, bu ortalamaların aritmetik ortalaması ve standart sapması ne olurdu?

Cevap: Hacimleri 12 birey olan örneklere sahibiz. Bu örneklerin aritmetik ortalamalarının aritmetik ortalaması popülasyonun aritmetik ortalamasına eşit olacaktır, yani 150 gram Bu örnek aritmetik ortalamalarının standart sapması ise,

$$\frac{30}{\sqrt{12}} = 8.66 \text{ gram.}$$

Dikkat ettiysen, aritmetik ortalamalar kümesinin aritmetik ortalamasını ve standart sapmasını verirken, dağılımın şeklini betimlemedik. Eğer, aynı zamanda, orijinal popülasyonun normal olduğu bize söylenmiş olsaydı, o zaman aritmetik ortalamalar kümesinin bir normal dağılıma sahip olduğu neticesine varabilirdik.

Misal 2 Bir üniversitenin öğrencilerinin bir aylık zorunlu harcamaları, aritmetik ortalaması 650 TL ve standart sapması 120 TL olan bir normal dağılıma sahiptir. Bu üniversitenin öğrencilerinin 10 kişilik bir basit rastgele örneğinin bir aylık zorunlu harcamalarının aritmetik ortalamasının 600 TL ve 700 TL arasında olması olasılığı nedir?

Cevap: Hacmi 10 kişi olan tüm olanaklı örneklerin aritmetik ortalamaları kümesinin aritmetik ortalaması ve standart sapması şöyledir:

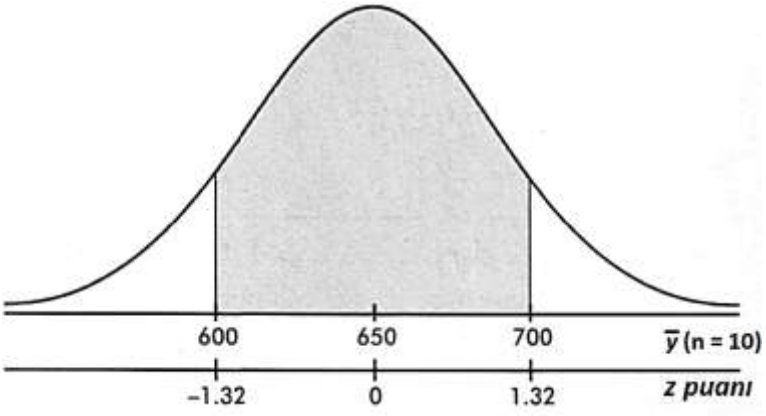
$$\mu_{\bar{y}} = 650 \text{ TL} \qquad \text{ve} \qquad \sigma_{\bar{y}} = \frac{120}{\sqrt{10}} = 37.95 \text{ TL.}$$

Dağılımın normal olduğunu bildiğimize göre, 600 TL ye ve 700 TL ye tekabül eden z-puanları sırasıyla şöyledir:

$$\frac{600 - 650}{37.95} = -1.32 \qquad \text{ve} \qquad \frac{700 - 650}{37.95} = 1.32.$$

Tablo B yi kullanarak, aradığımız olasılığı şöyle buluruz:

$$0.9066 - 0.0934 = 0.8132$$



Merkezi Limit Teoremi

Hemen yukarıdaki misalde orijinal popülasyonun dağılımının normal olduğunu varsaydık. Ne var ki, bırakalım tüm popülasyonların normal dağılmalarını, bazı popülasyonların dağılımı normale yaklaşır. Böyle olmakla beraber, orijinal popülasyonun dağılımı her ne olursa olsun, eğer *n* yeterince büyükse (en az 30), örnek aritmetik ortalamaları kümesinin dağılımının normale yaklaştığı matematiksel olarak gösterilebilir. Mesela, farklı kişilerin bakkal dükkanlarında harcadığı para miktarının normal dağıldığını farzetmek için makul bir sebebimiz olmayabilir. Mamafih, eğer her gün bir bakkaldan çıkan 30 kişiyi yoklar ve bakkal faturalarının aritmetik ortalamalarını hesaplarsak, göreceğiz ki bu günlük aritmetik ortalamalar yaklaşık olarak bir normal dağılıma sahip olacaktır.

Yeniden ifade edersek, herhangi bir örnek aritmetik ortalamasının örnekleme dağılımı örnek hacmi büyüdükçe normale daha çok yaklaşır. Bilmemiz gereken şey gözlem değerlerinin bağımsız olması ve rastgeleleştirme ile toplanmış olmalarıdır. Orijinal popülasyonun dağılımının şekline dair bir şey bilmemiz gerekmiyor. Bu olgu, istatistik biliminde *merkezi limit teoremi* olarak adlandırılır. Merkezi limit teoremi bize şunu söyler: Aritmetik ortalaması μ ve standart sapması σ olan bir popülasyondan çektiğimiz *n* hacimli bir örneğin $\bar{y}$ aritmetik ortalamasının örnekleme dağılımı yaklaşık olarak μ aritmetik ortalama ve $\frac{\sigma}{\sqrt{n}}$ standart sapma ile normaldir. Örnek hacmi genişledikçe bu yaklaşım da artar.

Yeterince geniş bir örnek hacmi için yukarıda kabaca andığımız $n \geq 30$ sınırı esasen çoğu zaman $n \geq 40$ olarak kullanılır ve eğer kestirimde daha fazla isabetlilik gerekiyorsa veya eğer orijinal popülasyonun şekli normalden aşırı uzaksa *n* daha büyük seçilmelidir. Elbette basit rastgele örnek varsayımı geçerli olmalı ve *n* popülasyonun %10 undan daha büyük olmamalıdır.

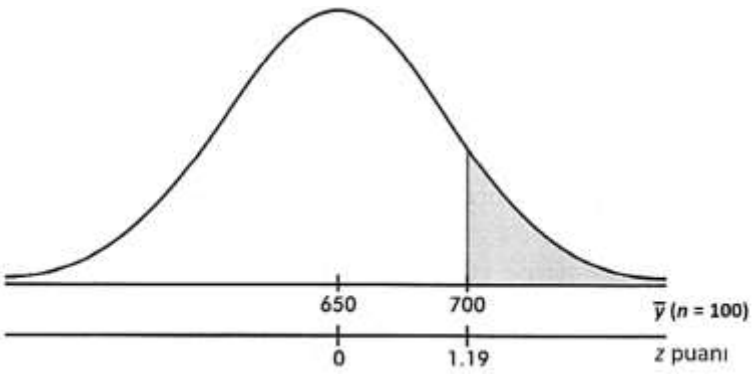
Misal 3 Geçmiş deneyimden, biliniyor ki bir bölgede hane başına günlük ekmek tüketimi 650 gram ve standart sapması 420 gramdır. Bu bölgede 100 hanelik bir basit rastgele örnek çekilirse, aritmetik ortalama ekmek tüketiminin 700 gramdan çok olması olasılığı nedir?

Cevap: Örnek hacmimiz 30 dan geniş ve bir basit rastgele örneğimiz var. Merkezi limit teoremine göre örnek aritmetik ortalamaları kümesinin dağılımı 650 gramlık aritmetik ortalama ve $420/\sqrt{100} = 42$ gramlık standart sapma ile yaklaşık olarak normaldir. Örneğimizin aritmetik ortalamasına tekabül eden *z*-puanı,

$$\frac{700 - 650}{42} = 1.19$$

olarak hesapladıktan sonra, Tablo B yi kullanarak örnek aritmetik ortalamasının 700 gramdan çok olması olasılığını buluruz:

$$1 - 0.8830 = 0.1170 \text{ veya } \%11.7$$



Misal 4 Bol çiçekli bir yaylada kovan başına bal üretiminin aritmetik ortalaması 25 kilogram ve standart sapması 2.3 kilogramdır. Kırk kovanlık bir basit rastgele örneğin kovan başına aritmetik ortalama bal üretiminin 24.5 kilogram ve 25.5 kilogram arasında olması olasılığı nedir?

Cevap: Örnek aritmetik ortalamasının örnekleme dağılımının aritmetik ortalaması ve standart sapması, sırasıyla,

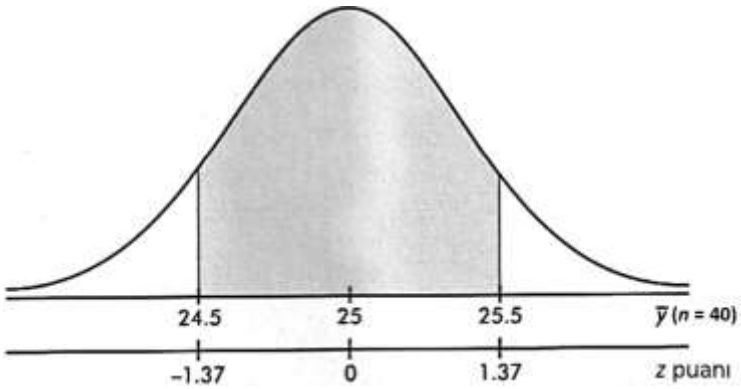
$$\mu_{\bar{y}} = 25 \text{ kilogram} \qquad \text{ve} \qquad \sigma_{\bar{y}} = \frac{2.3}{\sqrt{40}} = 0.364 \text{ kilogram.}$$

Örnek aritmetik ortalamasının bulunmasını arzu ettiğimiz aralığı sınırlayan değerlere tekabül eden z-puanları, sırasıyla,

$$\frac{24.5 - 25}{0.364} = -1.37 \qquad \text{ve} \qquad \frac{25.5 - 25}{0.364} = 1.37.$$

Tablo B yi kullanarak, örnekte kovan başına bal üretiminin aritmetik ortalamasının 24.5 kilogram ve 25.5 kilogram arasında olması olasılığını şöyle buluruz:

$$0.9147 - 0.0853 = 0.8294$$



Bu bölümde birden çok dağılımdan bahsediliyor olması kafaları karıştırmasın. Birinci olarak, orijinal popülasyonun dağılımı var ki şekli belki üniform, belki çan eğrisi, belki çarpık olabilir. İkinci olarak, örnekteki verinin dağılımı var ve örnek hacmi arttıkça bu dağılım popülasyonun dağılımına benzer. Üçüncü olarak, çekilen örnekle aynı hacimdeki olanaklı örneklerin aritmetik ortalamalarının dağılımı var ki bu örnekleme dağılımına, orijinal popülasyonun dağılımının şekline bakılmaksızın, bir normal dağılım modeli ile yaklaşılabilir.

Örnek Orantısının Örnekleme Dağılımı

Aritmetik ortalama esasen nicel bir ölçü(m) olmakla beraber, orantı esasen bir nitel yaklaşımı temsil eder. Üzerinde durduğumuz şey, bir özelliğin basitçe varlığı veya yokluğudur. "Evet" yanıtlarını sayarız ve bir orantı teşkil ederiz. Mesela, emniyet kemeri takan sürücülerin orantısı nedir? Bozuk aşıların orantısı nedir? Çimlenmeyen tohumların orantısı nedir?

Popülasyonun bu şekilde "var" ve "yok" diye ayrılması, daha önce öğrenmiş olduğumuz binomiyal dağılımdan yararlanabilmek için bize bir fırsat yaratır. Şunu da önceden yine biliyoruz ki *n* (yani

örnek hacmi) yeterince büyük ise, binomiyal dağılıma normal dağılım ile yaklaşabiliriz.

Bu kısımda, çekilen tek bir örneğin orantısı p yi dikkate almak suretiyle bir popülasyonun orantısı π nin kestirimini ele alacağız. Bir örneğin orantısı, tüm örnek orantıları uzayının yalnızca bir elemanıdır ve bir örnek orantısının anlamlılığını değerlendirmek için örnek orantılarının nasıl değiştiğini bilmemiz gerekir.

Belirli bir n hacimdeki tüm olanaklı örneklerin orantılarının kümesini mülahaza edelim. Bu orantıların popülasyon orantısı etrafında öbekleneceklerini ve örnek hacmi ne kadar geniş olursa, bu öbeklenmenin o kadar sıkı olacağını düşünmek makuldür. (Esasında, örnek orantısı popülasyon orantısının yansız bir kestirimidir.)

Örnek orantıları kümesinin aritmetik ortalamasını ve standart sapmasını nasıl hesaplarız? Örnek hacmi n ve hakiki popülasyon orantısı π olsun. Binomiyal dağılımı öğrenme-çalışmamızdan biliyoruz ki verilen bir örnekteki başarı sayısının aritmetik ortalaması ve standart sapması sırasıyla $n\pi$ ve $\sqrt{n\pi(1-\pi)}$ dir. Yine biliyoruz ki büyük n durumunda dağılımın biçimi "normal"e yaklaşır.

Burada, mamafih, başarıların sayısından ziyade orantıyla ilgileniyoruz. Yine önceki bölümlerden biliyoruz ki her gözlem değerini bir sabitle çarpar veya bölersek, hem aritmetik ortalamayı ve hem de standart sapmayı aynı sabitle çarpmış veya bölmüş oluruz. Bu durumda, başarıların sayısını (np) başarıların orantısına (p) dönüştürmek için n e böleriz:

$$\mu_p = \frac{n\pi}{n} = \pi \qquad \text{ve} \qquad \sigma_p = \frac{\sqrt{n\pi(1-\pi)}}{n} = \sqrt{\frac{\pi(1-\pi)}{n}}.$$

Dahası, eğer yaklaşık normal bir dağılımın her gözlem değerini aynı sabite bölersek, dağılımın biçiminin yine yaklaşık normal olacağını düşünmek makuldur.

Netice itibarıyla, bir p örnek orantısının örnekleme dağılımı yaklaşık olarak π aritmetik ortalama ve $\sqrt{\frac{\pi(1-\pi)}{n}}$ standart sapma ile normaldir. Örnek hacmi genişledikçe bu yaklaşım da artar.

Binomiyal dağılıma normal dağılımla yaklaşmayı kullanacağımız için, hem $n\pi$ hem de $n(1-\pi)$ en az 10 olmalıdır. Daha önemlisi, belirli bir örnek için hesaplamalar yaparken ve neticelere varırken, bu örneğin bir *basit rastgele örnek* olması gerektiğini unutmamaktır.

Nihayetinde, örnekleme genellikle *yerine koymadan* yapıldığı için, örnek hacminin popülasyonun hacminin %10 unu geçmesi durumunda *gözlemlerin bağımsızlığı varsayımının* ihlal ediliyor olmasından endişelenmek gerekir. Bu durumda, nispeten daha dar ama mutlaka rastgele seçilmiş bir örnek ile gözlemlerin bağımsızlığı varsayımı yerine getirilebilir.

Elbette örneğin geniş hacimli olması her zaman arzu ediliyor olsa bile, örnek hacmi popülasyona nispetle geniş olduğu zaman bu dersin kapsamına girmeyen daha münasip istatistiksel tekniklerin kullanılması söz konusudur.

Mesela, insan türü için, doğumda kız bebeklerin π orantısı 0.5 tir. Rastgele 100 doğumda kız bebeklerin p orantısı, $N(0.5, 0.05)$ dağılımına göre değişir. Şöyle ki:

$$\mu_p = \pi = 0.5$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.5(1-0.5)}{100}} = 0.05$$

Misal 5 Bir hastahanenin kayıtlarına göre, tüm diyaliz hastalarının %70 i en az 5 yıl yaşamaktadır. Yeni teşhis koyulan diyaliz hastalarının 100 kişilik bir basit rastgele örneğinde, en az 5 yıl yaşayacakların orantısının %80 den fazla olması olasılığı nedir?

Cevap: Hem $n\pi = 100(0.7) = 70$ hem de $n(1-\pi) = 100(0.3) = 30$ 10 dan büyük. Diğer taraftan örnek hacmimizin tüm diyaliz hastalarının sayısının %10 undan küçük olduğu aşikar. Öyle ise, örnek

orantıları kümesi $\mu_p = 0.70$ lik aritmetik ortalama ve

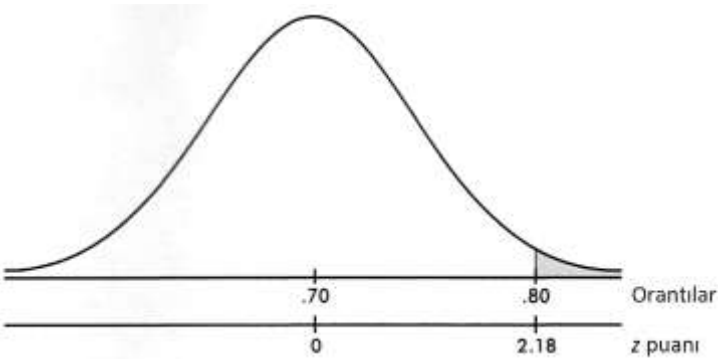
$$\sigma_p = \sqrt{\frac{0.7(0.3)}{100}} = 0.0458$$

lik standart sapma ile yaklaşık normal dağılmaktadır. Yüzde 80 lik örnek orantısına tekabül eden z-puanını,

$$\frac{0.80 - 0.70}{0.0458} = 2.18$$

olarak hesapladıktan sonra, Tablo B yi kullanarak örnek orantısının %80 den büyük olması olasılığını hesaplarız:

$$1 - 0.9854 = 0.0146$$



Misal 6 Otomobil sürücülerinin %48 inin emniyet kemeri kullandığı, bir öğrenme-çalışması ile kestirilmiştir. Bir yol kontrolünde tamamen rastgele durdurulan 400 otomobil sürücüsünden kemer takanların orantısının %45 ve %55 arasında olması olasılığı nedir?

Cevap: Hem $n\pi = 400(0.48) = 192$ hem de $n(1-\pi) = 400(0.52) = 208$ 10 dan büyüktür. Diğer taraftan örnek hacmimizin otomobil kullanıcılarının sayısının %10 undan küçük olduğu aşikar. Öyle ise, örnek orantıları kümesi $\mu_p = 0.48$ lik aritmetik ortalama ve

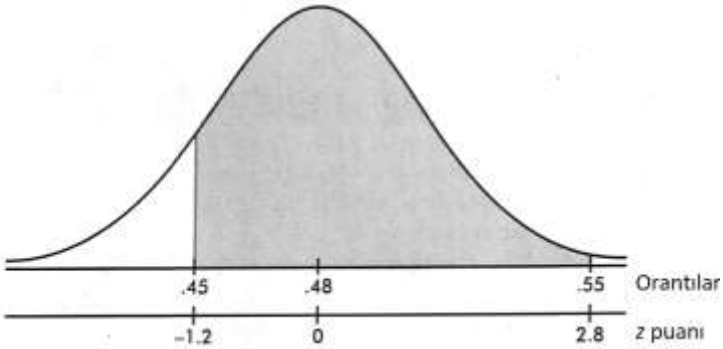
$$\sigma_p = \sqrt{\frac{0.48(0.52)}{400}} = 0.025$$

lik standart sapma ile yaklaşık normal dağılmaktadır. Örnek orantısının bulunmasını arzu ettiğimiz aralığı sınırlayan değerlere tekabül eden z-puanları, sırasıyla,

$$\frac{0.45 - 0.48}{0.025} = -1.2 \qquad \text{ve} \qquad \frac{0.55 - 0.48}{0.025} = 2.8.$$

Tablo B yi kullanarak, örnek orantısının %45 ve %55 arasında olması olasılığını şöyle buluruz:

$$0.9974 - 0.1151 = 0.8823$$



Bağımsız İki Örnek Aritmetik ortalaması Arasındaki Farkın Örnekleme Dağılımı

İstatistik biliminin pek çok gerçek hayat uygulaması iki popülasyonun aritmetik ortalamalarını karşılaştırmayı kapsar. Mesela, spesiyal bir diyetle beslenen laboratuvar farelerinin ortalama ağırlığı standart diyetle beslenen laboratuvar farelerininkinden farklı mıdır? Bankacılık sektöründe mi yoksa sağlık hizmetleri sektöründe mi ortalama ücret daha yüksektir? İnşaat işçilerinin ömrü doktorların ömründen daha kısa mıdır?

Önce örneklerin aritmetik ortalamalarını nasıl karşılaştıracağımızı mülahaza edelim. İki popülasyondan birer örneğimiz var. Bu iki örneğin aritmetik ortalamaları arasında gözlenen fark ne zaman anlamlıdır? Bu farkın bir farklar kümesinden çekilmiş olduğunu kavradığımız taktirde cevap aşikar olur. Yani, bir kümedeki örnek aritmetik ortalamalarını başka bir kümedeki örnek aritmetik ortalamalarından çıkartmak suretiyle elde edilen tüm olanaklı farklar kümesinin mevcudiyeti söz konusudur.

Belirli bir farkın anlamlılığını değerlendirmek için, önce tüm olanaklı farkların kendi aralarında nasıl değiştiğini belirlememiz gerekir. Şu olgu bize ışık tutacaktır ki bir farklar kümesinin varyansı, ilişkin kümelerin varyanslarının toplamına eşittir. Böylece,

$$\sigma^2_{\bar{y}_1-\bar{y}_2} = \sigma^2_{\bar{y}_1} + \sigma^2_{\bar{y}_2}.$$

Şimdi,

$$\sigma_{\bar{y}_1} = \frac{\sigma_1}{\sqrt{n_1}} \qquad \text{ve} \qquad \sigma_{\bar{y}_2} = \frac{\sigma_2}{\sqrt{n_2}}$$

olduğuna göre,

$$\sigma^2_{\bar{y}_1-\bar{y}_2} = \frac{\sigma_1^2}{\sqrt{n_1}} + \frac{\sigma_2^2}{\sqrt{n_2}} \qquad \text{ve} \qquad \sigma_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

İki örnek aritmetik ortalaması arasındaki farkın örnekleme dağılımı yaklaşık normaldir. Örneklerin çekildiği popülasyonlardan herhangi birinin dağılımı normalden uzaklaştıkça, mütekabil örnek hacminin artırılması gerekir. Aynı zamanda bağımsız basit rastgele örnekler varsayımının geçerli olması gerekir. Ayrıca örnek hacimlerinin popülasyonların %10 unu geçmemesi gerekir.

Misal 7 A gübresi ile muamele edilen bitkilerin verimi parsel başına 4.5 kg ve standart sapması 0.7 kg iken, B gübresi ile muamele edilen bitkilerin parsel başına verimi 4.3 kg ve standart sapması 0.4 kg dır. A gübresi 45 parsele ve B gübresi 50 parsele uygulandığına göre, verim ortalamaları arasındaki farkın negatif çıkması olasılığı nedir? Yani A gübresi uygulanan parsellerin ortalama veriminin B gübresi uygulanan parsellerin ortalama veriminden az olması olasılığı nedir?

Cevap: Geniş hacimli ($n_1 = 45$ ve $n_2 = 50$) basit rastgele örneklere sahibiz ve tüm olanaklı parsellerin %10 undan daha az sayıda oldukları da açık. Farkların aritmetik ortalaması

$$\mu_1 - \mu_2 = 4.5 - 4.3 = 0.2 \text{ kg}$$

ve standart sapması

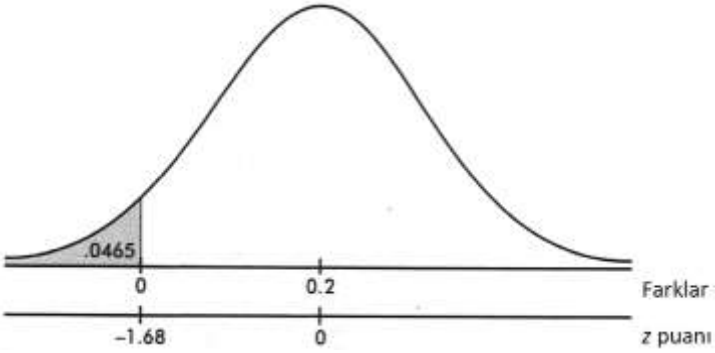
$$\sigma_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{0.7^2}{45} + \frac{0.4^2}{50}} = 0.119 \text{ kg}.$$

Sıfırlık bir farka tekabül eden z-puanı,

$$\frac{0 - 0.2}{0.119} = -1.68.$$

Tablo B de, -1.68 in solundaki alan 0.0465 e eşittir. Buna göre, A gübresi uygulanan örnek parsellerin

ortalama veriminin B gübresi uygulanan örnek parsellerin ortalama veriminden az olması olasılığı %4.65 tir.



Misal 8 Devlet okullarına kayıtlı ilkokul öğrencilerinin devamsızlık ortalaması 8.5 gün ve standart sapması 4.1 gün iken, özel okullara kayıtlı ilkokul öğrencilerinin devamsızlık ortalaması 5.3 gün ve standart sapması 2.9 gündür. Devlet okullarına kayıtlı ilkokul öğrencilerinin 200 kişilik bir basit rastgele örneği ve özel okullara kayıtlı ilkokul öğrencilerinin 150 kişilik bir basit rastgele örneği alındığında, devamsızlıklarının aritmetik ortalamaları arasındaki fark (devlet okulu – özel okul), 0.95 lik bir olasılıkla hangi değerin üstündedir?

Cevap: Geniş hacimli ($n_1 = 200$ ve $n_2 = 150$) basit rastgele örneklere sahibiz ve tüm olanaklı devlet okulu ve özel okul öğrencilerinin %10 undan daha az sayıda oldukları da açık. Farkların aritmetik ortalaması

$$\mu_1 - \mu_2 = 8.5 - 5.3 = 3.2 \text{ gün}$$

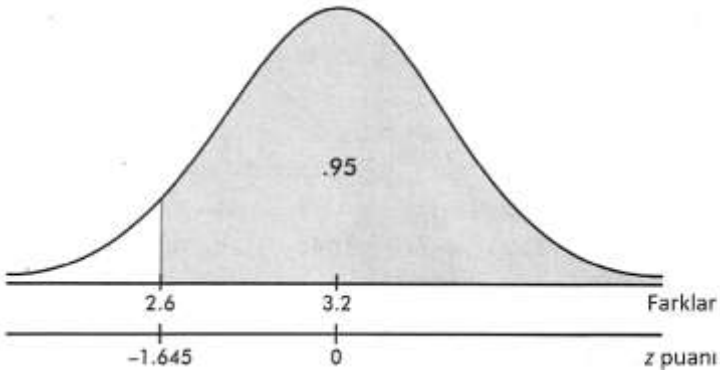
ve standart sapması

$$\sigma_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{4.1^2}{200} + \frac{2.9^2}{150}} = 0.374 \text{ gün}.$$

Tablo B de, 0.95 lik alanı sağdan sınırlayan z-puanı 1.645 e tekabül etmekte. Dağılım simetrik olduğu için 0.95 lik bir alan -1.645 lik z-puanının sağında kalacaktır. Buna göre, devamsızlıkların aritmetik ortalamaları arasındaki farkı, 0.95 lik bir olasılıkla

$$3.2 - 1.645(0.374) = 2.6 \text{ gün}$$

olarak hesaplarız. Yani, devlet okulu ve özel okul örneklerinin devamsızlık aritmetik ortalamaları arasındaki fark, 0.95 lik bir olasılıkla 2.6 günden fazla olacaktır.



Bağımsız İki Örnek Orantısı Arasındaki Farkın Örnekleme Dağılımı

İstatistik biliminin sayısız önemli uygulamaları iki popülasyon orantısının karşılaştırılmasını kapsar. Mesela, Türk malı buzdolaplarını kullananların memnuniyet orantısı Güney Kore malı buzdolaplarını kullananlarınkinden az mıdır? Başarılı bir girişimci olmak için öğretimi önceleyen işadamlarının yüzdesi deneyimi önceleyen işadamlarının yüzdesi ile karşılaştırıldığında nasıldır? Özel okuldan mezun öğrencilerin üniversiteye girenlerinin orantısı ile devlet okulundan mezun öğrencilerin

üniversiteye girenlerinin orantısı arasındaki fark için ne söylenebilir?

İki örnek orantısını karşılaştıracak bir prosedürü konu ediyoruz, burada. Ayrı popülasyonlardan çekilen iki örnek orantısı arasındaki fark ne zaman anlamlıdır? Bir kümedeki örnek orantılarını başka bir kümedeki örnek orantılarından çıkartmak suretiyle elde edilen tüm olanaklı farklar kümesinden çekilen bir farkı ele alıyoruz. Belirli bir farkın anlamlılığını değerlendirmek için, önce farkların kendi aralarında nasıl değiştiğini belirlememiz gerekir. Biliyoruz ki bir farklar kümesinin varyansı, ilişkin kümelerin varyanslarının toplamına eşittir. Yani,

$$\sigma_{p_1-p_2}^2 = \sigma_{p_1}^2 + \sigma_{p_2}^2$$

Şimdi,

$$\sigma_{p_1} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1}} \qquad \text{ve} \qquad \sigma_{p_2} = \sqrt{\frac{\pi_2(1-\pi_2)}{n_2}}$$

olduğuna göre,

$$\sigma_{p_1-p_2}^2 = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}$$

ve

$$\sigma_{p_1-p_2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}\,.$$

İki örnek orantısı arasındaki farkın örnekleme dağılımı yaklaşık normaldir. Burada binomiyal dağılıma normal dağılımla yaklaştığımız için $n_1\pi_1$, $n_1(1-\pi_1)$, $n_2\pi_2$ ve $n_2(1-\pi_2)$ değerlerinden her biri en az 10 olmalıdır. Aynı zamanda, bağımsız basit rastgele örnekler varsayımının geçerli olması gerekir. Ayrıca örnek hacimlerinin popülasyonların %10 unu geçmemesi gerekir.

Misal 9 Atatürk Üniversitesinde erkek öğrencilerin %23 ünün ve kız öğrencilerin %12 sinin Erzurumspor taraftarı olduğu biliniyor. Erkek öğrencilerin 100 kişilik bir rastgele örneğinde ve kız öğrencilerin 125 kişilik bir rastgele örneğinde Erzurumspor taraftarı olan öğrencilerin orantıları arasındaki farkın %10 dan daha çok olması olasılığı nedir?

Cevap: Görüyoruz ki $n_1\pi_1 = 23$, $n_1(1-\pi_1) = 77$, $n_2\pi_2 = 15$, $n_2(1-\pi_2) = 110$ ve hepsinin değeri 10 dan büyük; bağımsız basit rastgele örnekler alınmış ve örnek hacimleri çekildikleri popülasyonların hacminin %10 undan küçük. Örnek orantıları arasındaki farkın örnekleme dağılımının aritmetik ortalaması,

$$\mu_{p_1-p_2} = \mu_{p_1} - \mu_{p_2} = 0.23 - 0.12 = 0.11$$

ve standart sapması,

$$\sqrt{\frac{0.23(0.77)}{100} + \frac{0.12(0.88)}{125}} = 0.0511\,.$$

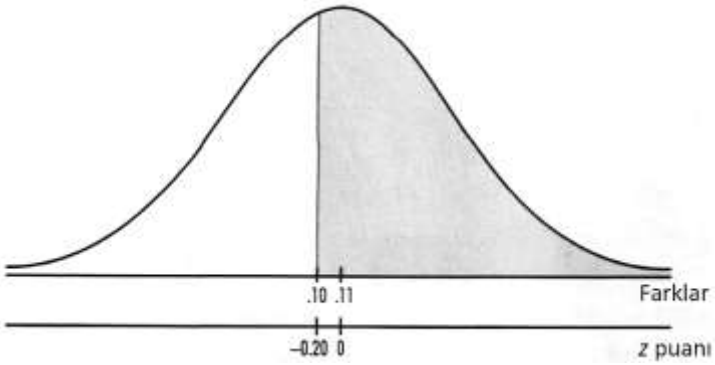
Yüzde onluk farka tekabül eden *z*-puanı,

$$\frac{0.10 - 0.11}{0.0511} = -0.20\,.$$

Tablo B yi kullanarak, farkın %10 dan çok olması olasılığını,

$$1 - 0.4207 = 0.5793$$

olarak buluruz.



Misal 10 Şehirde yaşayan evli çiftlerin %43 ü herhangi bir evcil hayvana sahipken bekarların %19 u evcil hayvan sahibidir. İki yüz evli çiftten oluşan bir basit rastgele örnekte ve bekarların 180 kişilik bir basit rastgele örneğinde evcil hayvana sahip olanların yüzdeleri arasındaki fark 0.90 lık bir olasılıkla hangi değerden daha büyüktür?

Cevap: Görüyoruz ki $n_1\pi_1 = 86,\ n_1(1-\pi_1) = 114,\ n_2\pi_2 = 34.2,\ n_2(1-\pi_2) = 145.8$ ve hepsinin değeri 10 dan büyük; bağımsız basit rastgele örnekler alınmış ve örnek hacimleri çekildikleri popülasyonların hacminin %10 undan küçük. Örnek orantıları arasındaki farkın örnekleme dağılımının aritmetik ortalaması,

$$\mu_{p_1-p_2} = \mu_{p_1} - \mu_{p_2} = 0.43 - 0.19 = 0.24$$

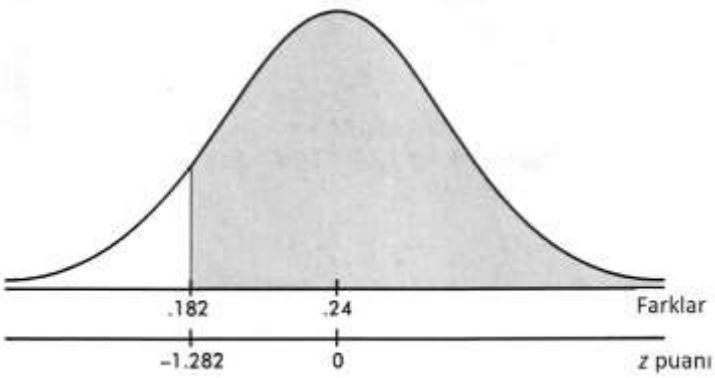
ve standart sapması,

$$\sqrt{\frac{0.43(0.57)}{200} + \frac{0.19(0.81)}{180}} = 0.0456\,.$$

Tablo B de, 0.90 lık alanı sağdan sınırlayan z-puanı 1.282 ye tekabül etmektedir. Dağılım simetrik olduğu için 0.90 lik bir alan -1.282 lik z-puanının sağında kalacaktır. Buna göre, orantıların farkının 0.90 lık bir olasılıkla daha büyük olacağı değeri,

$$0.24 - 1.282(0.0456) = 0.182$$

olarak hesaplarız. Yani, evli çiftlere ait ve bekarlara ait örneklerde evcil hayvana sahip olma yüzdeleri arasındaki fark, 0.90 lık bir olasılıkla %18.2 den daha büyük olacaktır.



Standart Hata

Bir istatistiğin örnekleme dağılımının bir rastgele örnekten hesaplanan standart sapmasına aynı zamanda *standart hata* adı verilir. Hem aritmetik ortalamalar hem de orantılar söz konusu olduğu zaman, pek çok durumda popülasyon parametrelerini bilmeyiz. Daha önce söylemiş olduğumuzu bir kere daha burada tekrar edecek olursak, popülasyon parametrelerinin değerleri sabit fakat genellikle meçhuldür. Böyle olunca, bir örnekleme modelinin (veya dağılımının) standart sapmasını, yani eldeki istatistiğin standart hatasını hesaplamada, çoğu zaman popülasyonun standart sapması (σ) malum olabilmeyeceği için, çektiğimiz örneğin standart sapmasını (s) kullanarak bir kestirimde bulunuruz. Yani, örnek-ortalamaları için,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \quad \text{yerine} \quad s_{\bar{y}} = \frac{s}{\sqrt{n}}.$$

Benzerce, orantılar için,

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \quad \text{yerine} \quad s_p = \sqrt{\frac{p(1-p)}{n}}.$$

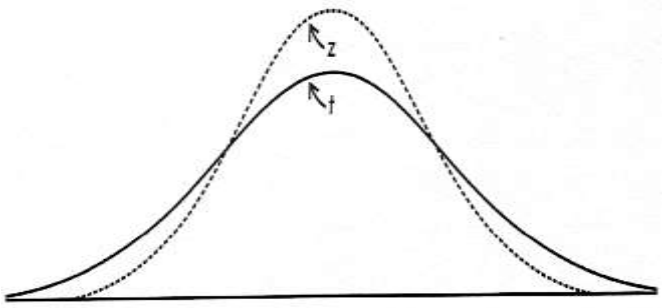
Student *t*-Dağılımı

Popülasyon standart sapması σ bilinmediği zaman, örnek standart sapması s yi σ nın bir kestirimi olarak kullanırız. Fakat bu taktirde $\frac{\bar{y}-\mu}{s/\sqrt{n}}$ normal dağılım göstermez. Mamafih, eğer **orijinal popülasyon normal dağılıyor** ise, $s/\sqrt{n}$ oranları ile uyarlı bir dağılım vardır. Bu dağılıma *Student t-dağılımı* adı verilir.

Böylece, normal dağılıma sahip bir popülasyondan bir örnek çektiğimiz zaman, *z*-puanı yerine *t*-puanı hesaplarız:

$$t = \frac{\bar{y}-\mu}{s/\sqrt{n}}$$

Aynı örnek hacmi için, *t*-puanlarının dağılımı da çan eğrisi şeklinde ve simetriktir, fakat *z*-puanlarının dağılımına göre merkezde daha basık ve kuyruklarda daha diktir. Yani, bu demektir ki *t*-dağılımı *z*-dağılımına göre daha yayvandır.



Binomiyal dağılım gibi, *t*-dağılımı da *n* in farklı değerleri için farklıdır. Bu farklı *t*-dağılımları, *serbestlik derecelerine* göre ayrı ayrı tablolaştırılmıştır.

Serbestlik derecesi örnekleme dizaynına bağlı olarak örnek hacminden belli bir sayı eksiltilerek hesaplanır ve *sd* ile gösterilir. Mesela popülasyon aritmetik ortalamasının kestirimi maksadıyla alınan bir rastgele örnek için serbestlik derecesi örnek hacminin bir eksiğidir. Serbestlik derecesinin değeri ne kadar küçülürse, dağılımın yayılımı da o kadar artar. Serbestlik derecesinin değeri ne kadar büyürse, dağılım normale o kadar yaklaşır.

Her *sd* değeri için ayrı bir *t*-dağılımı olması hasebiyle, bu tabloların eksiksiz hali sayfalarca yer işgal eder. Bu yüzden, Tablo C de yalnızaca farklı örnek hacimleri için çok müştereken kullanılan kuyruk olasılıklarına karşılık gelen *t*-değerleri verilmektedir. Tablo C nin son satırı standart normal dağılımdır ki *n* sonsuz olduğu zaman *t*-dağılımının özel bir durumudur. Pratikte, her iki dağılım $n \geq 30$ durumunda çok yakınlaşır. Bununla beraber, daha fazla hassasiyet istenirse, *n* in 30 dan daha büyük değerleri için de Student *t*-dağılımı kullanılabilir.

Şuna dikkat edilmelidir ki Tablo B belirli bir *z*-puanının solunda normal eğri altında kalan alanı verirken, Tablo C belirli pozitif bir *t*-puanının sağında eğri altında kalan alanı verir. Mesela, örnek hacmini 20 farzedelim ve böylece $sd = 20 - 1 = 19$ olacaktır. O zaman, kuyruktaki 0.05 lik bir olasılık, 1.729 luk bir *t*-değerine tekabül edecektir. Yine kuyruktaki 0.01 lik bir olasılığın tekabül ettiği *t*-değeri 2.539 dur.

Bu manada, popülasyon standart sapması σ her ne zaman bilinmiyor ise, t-dağılımı münasip bir tercihtir. Gerçek dünyada σ neredeyse her zaman bilinmezdir ve böylece neredeyse her zaman t-dağılımını kullanacağız. Günümüze değin bu pek kolay değildi, çünkü hassas hesaplamalar için çeşitli t-dağılımlarına ait yoğunluklu tablolara gereksinim duyulurdu. Oysa, günümüzde kullanılan hesap makinaları ve kompüter yazılımları ile bu sorun hemen hemen ortadan kalkmıştır. Artık, n in 30 dan daha büyük olduğu hallerde t-dağılımının z-dağılımına yeterince yaklaştığını varsaymak zorunda değiliz.

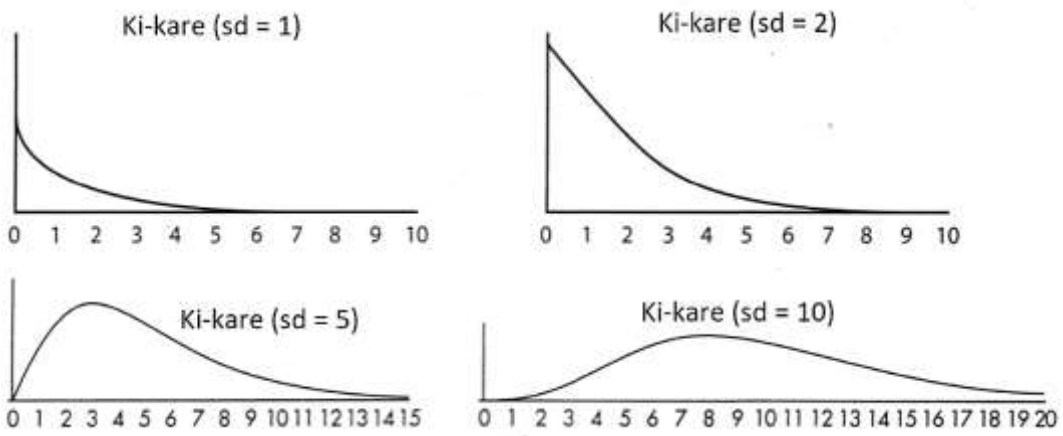
İstatistik bilimcilerinin örnek hacmine ilişkin tartışmalarını aşağıdaki gibi rafine edebiliriz:

Hacmi n olan bir basit rastgele örnek için t-dağılımını kullandığımız taktirde;	geniş hacimli, $n \geq 40$ ise,	orijinal popülasyon hakkında herhangi bir varsayımda bulunmaya gerek yok.
	orta hacimli, n 15 ve 40 arasında ise,	örnekte aşırı değerler olmamalı ve eğer varsa çarpıklık az olmalıdır; veya orijinal popülasyonun dağılımının normal olduğu varsayılmalıdır.
	dar hacimli, $n \leq 15$ ise,	örnekte dışlaklar bulunmamalı ve çarpıklık görülmemelidir; veya orijinal dağılımın normal olduğu varsayılmalıdır.

Dağılımı normale yakın olmayan bir popülasyondan çekilen dar hacimli örneklerle çalıştığımız zaman, parametrik-olmayan (non-parametrik te denir) metotları kullanırız ki onlar bu dersin kapsamı dışındadır.

Ki-kare Dağılımı

Sonraki konularımızda, özellikle bazı hipotezlerin testinde kullanacağımız bir örnekleme dağılımı modeli de ki-kare dağılımıdır. Ki-kare dağılımının parametresi serbestlik derecesidir (sd). Küçük serbestlik dereceleri için dağılım sağa çarpıktır. Mamafih, serbestlik derecesi büyüdükçe dağılım simetrikleşir ve çan eğrisi şekline kavuşur (t-dağılımında olduğu gibi). Serbestlik derecesi 1 veya 2 iken, dağılımın tepesi 0 civarındadır, fakat 3 veya daha çok serbestlik derecesi için tepesi $sd - 2$ civarındadır.

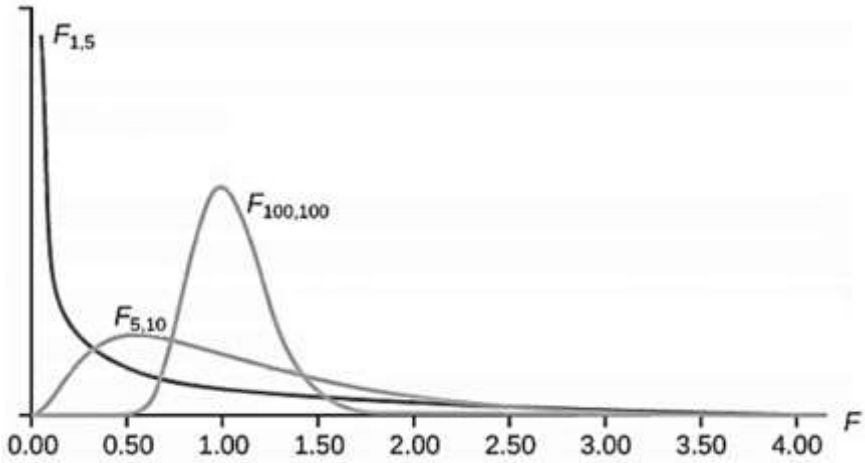


Ki-kare dağılımı, t-dağılımı gibi, her farklı sd için ayrı bir eğriye sahiptir. Bu yüzden Ki-kare tablolarında (Tablo D) yalnızca müştereken kullanılan kuyruk olasılıklarına ait kritik ki-kare değerleri verilir. Ki-kare dağılımının sayısız uygulamaları vardır. En iyi bilinen uygulamalarından birisi uyum iyiliği testidir ki gözlenen bir dağılımın teorik bir dağılıma uyumunu test etmek için kullanılır. Diğer birisi ise bağımsızlık testidir ki nitel değişkenler arasındaki ilişkiyi test etmek için kullanılır. Her ne

kadar sonraki konularımızda ki-kare dağılımını kategorik veri için kullanacak olsak ta, ki-kare dağılımı yukarıda görüldüğü üzere sürekli bir dağılımdır ve sayım verisine (yani frekanslara) uygulanması sadece bir yaklaşımdır. Ki-kare, χ^2 şeklinde de yazılabilmektedir.

F-Dağılımı

Başka bir örnekleme dağılımı modeli de *F*-dağılımıdır. Sonraki konularda, *F*-dağılımını varyansların oranını test etmek için kullanacağız. *F*-dağılımının iki parametresi, oranlanan varyansların serbestlik dereceleridir ($n_1 - 1$ ve $n_2 - 1$). Dağılımın şekli daima sağa çarpıktır ve serbestlik dereceleri arttıkça simetrikleşir. Dağılımın beklenen değeri 1 civarındadır.



F-dağılımı da *t*-dağılımı ve Ki-kare dağılımı gibi, varyansların serbestlik derecelerinin her farklı kombinasyonu için ayrı bir eğriye sahiptir. Bu yüzden *F*-dağılımı tablolarında da (Tablo E) yalnızca müştereken kullanılan kuyruk olasılıklarına ait kritik *F*-değerleri verilir.

Mini Proje 12

Aşağıdaki görsel $N = 400$ hacimli bir popülasyonu temsil ediyor. Bireylerin 160 ı (%40) kalıtsal bir hastalığa ait geni taşıyor ki bunların simgesi ● ve geriye kalan 240 ı hastalık genini taşımıyor ki bunların simgesi ○.

○●●●○○○○●●●●●○○○○○○○○○○○○○●○●○●○●●●●●●●●●●●●●●●○○○○
○○○○○○○○○○○○●●●●●●●●●●●●●●○○○○○○○○○○○○○●●●●●●○○○○○
○●●○○○○○○○○○●○○●○●○●○●○●○●○●●●●●●●●●●●●●●●●●●●●●●●○○○●●
●●●●●●●○○○○○○○○○○○○○○○●●●○○○○○○○○○○○○○○●●●●●●●●●●●●●○○
○○○○○○○○○○○○○●●●●●○○○○○○○○○○○○○○○○○○○○○○○○●●●●●●●●●○●○
○○○○○●○○○○○○●●○○○○○○○○○○●●●●○○○○○●●●●●○○○○○○○○○○○○○○○○
○○○○○○○○○○●●●●●●●○○●○○○○○○○○○○○○○●●●●●●○○○○○○○●●●●○○○○
○○●●○○○○○○○○●●●●●○○○○○○○○○○○○○●○●○○○○○○○○○●●●○○○○○○○○○○○
○○○○●○○○

Bu popülasyondan 20 ayrı örnek çekeceksin. Şu adımları takip et:

Adım 1. Bu popülasyondan basit rastgele örnekler çekmek için bir metot geliştir. Mesela bir metot simgeleri keserek bir torbaya koyup, iyice karıştırıp, torbadan çekmek olabilir. Daha az zahmetli metotlar da var elbette. Fakat gerçekten rastgele örnekler çekebileceğinden emin olmalısın. *Metodunu açıkla*.

Adım 2. $n = 15$ hacimli bir rastgele örnek çek ve hastalık genini taşıyanların sayısını ve yüzdesini kaydet.

Adım 3. Adım 2 yi 20 defa tekrar et. Böylece, toplamda 20 ayrı örneğin olacak. Her defasında örnekleme prosedürünü baştan alman çok önemli. Mesela eğer simgeleri torbadan çekme metodunu kullanıyorsan, her 15 hacimlik örneği çektikten sonra, çıkan simgeleri torbaya geri koymalısın ki bir sonraki örnek için de müsait olsunlar.

Adım 4. Yirmi örneğin orantılarının bir dal-yaprak grafiğini veya histogramını yap. Aritmetik ortalamayı hesapla.

Adım 5. Binomiyal dağılıma normal dağılımla yaklaşmayı göz önüne alırsan, hesapladığın örnek orantıları için nasıl bir dağılım beklersin?

Adım 6. Kendi sonuçlarını örnek orantılarının dağılımına normal yaklaşıma ait beklentinle karşılaştır. Aritmetik ortalamayı, standart sapmayı, dağılımın şeklini ve örnek orantılarının %68 inin, % 95 inin ve 99.7 sinin içine düşmesini beklediğin aralıkları belirtmeyi unutma.

BÖLÜM 13 | GÜVEN ARALIKLARI

Giriş

Çıkarımsal istatistiğin ilgi alanlarından biri *popülasyon parametrelerinin örnek istatistikleri vasıtasıyla kestirimidir*. Buradaki sırayı anlamak önemlidir. Örnek istatistiği örnek veriden hesaplanır ve popülasyonun meçhul parameteresi bu örnek istatistiği ile kestirilir (veya daha geniş manada ifade eedersek, bir çıkarımda bulunulur). Tekrar etmek gerekirse, istatistikler hesaplanır, parametreler kestirilir. Ancak, bir örnekten hesaplanan bir istatistik ile, bir popülasyonun aritmetik ortalaması veya orantısı gibi herhangi bir parametrenin hakiki değerinin tam olarak kestirilip kestirilmediğini asla bilemeyiz. Buna **nokta kestirimi** denir. Bunun yerine, şunu da yapabiliriz: belli bir *güvenle* popülasyonun parametresini içine alacak bir aralığı, bir örneği kullanarak, hesaplayabiliriz. Buna da **aralık kestirimi** denir. Bir aralık kestirimi aritmetik ortalama veya orantı gibi herhangi bir örnek istatistiğini merkeze alır ve buna bir *hata payının* eklenmesi ve çıkartılması ile ifade edilir.

Güven Aralığının Manası

Örnekleme dağılımları hakkında öğrendiklerimizi kullanarak, bir örnek aritmetik ortalamasının veya orantısının içinde yer aldığı, popülasyon aritmetik ortalaması veya orantısı etrafındaki bir aralığı belli bir güvenle tesis edebiliriz. Bunu yapabiliyorsak, o zaman, aynı güvenle, bir popülasyonun aritmetik ortalamasının veya orantısının içinde yer aldığı, örnek aritmetik ortalaması veya orantısı etrafındaki bir aralığı da belirtebiliriz. (Yani, mesela, Erzurumun'un Ankara'ya uzaklığı, Ankara'nın Erzurum'a uzaklığıyla aynıdır. Bölüm 13, 14 ve 15.)

Çoğu zaman %90, %95 ve %99 luk güven aralığı kestirimleri göz önüne alınıyor olsa da, *güven düzeyi* olarak her hangi bir yüzdeyi seçmek mümkündür. Bu yüzde, yani güven düzeyi, örnek kestiriminin belli bir hata payı sağında ve solunda kalan alan içerisinde meçhul μ yü veya π yi bulunduracak örneklerin yüzdesidir. Bunu söylerken, μ nün veya π nin hesaplanan bir örnek aritmetik ortalaması veya bir örnek orantısının belli bir hata payı civarına düşmesi olasılığının 0.90, 0.95 veya 0.99 olduğunu söylemiyoruz. Hesaplanan bir aralık içerisinde μ veya π ya vardır ya da yoktur ve böylece olasılık ya 1 ya da 0 dır. Yani, parametrenin değeri meçhul fakat sabit olduğu için, burada değişken olan paramtrenin değeri değil, hesaplanan güven aralığıdır.

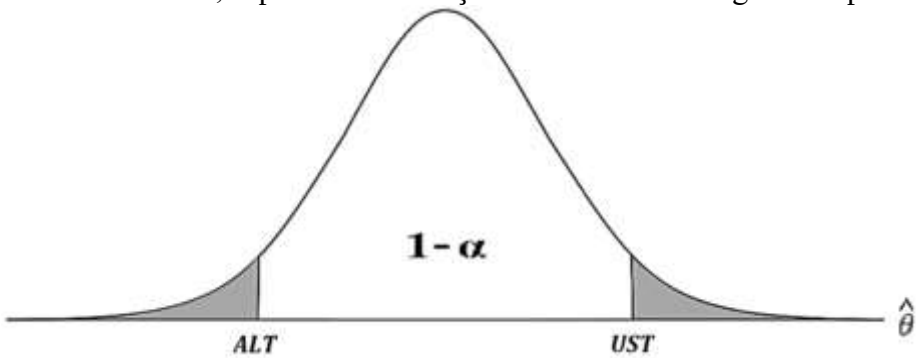
Görüleceği üzere, bu kavramlaştırmaya dair iki hususiyet mevcuttur. Birincisi, *güven aralığıdır* ki çoğunlukla şu formda ifade edilir:

$$\text{kestirim} \pm \text{hata payı}\,.$$

İkincisi, bu *prosesin başarı yüzdesidir*; yani, aynıyla tekrarlanacak örnekleme proseslerinde, hesaplanan güven aralıklarının popülasyon parametresini ihtiva edenlerinin orantısı. Bu manada, güven aralığı sabit ama meçhul bir popülasyon parametresini belirli bir olasılıkla içinde bulundurabilecek bir aralık kestirimini tanımlar. Güven aralığının alt sınırını ALT ve üst sınırını da UST ile gösterirsek, bir parametre için güven düzeyi, parametrenin hakiki değeri θ ise,

$$P(ALT \leq \theta \leq UST) = 1 - \alpha$$

olarak yazılır. Bu tanımlamada, θ parametresi meçhul fakat sabit bir değere sahiptir.



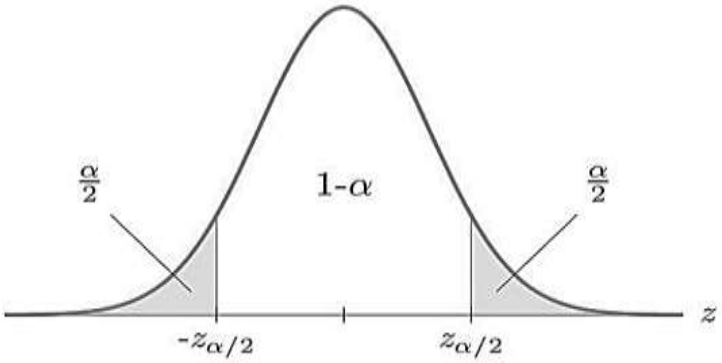
Güven düzeyi $(1-\alpha) = \%95$ olarak seçilirse, $\alpha = \%5$ olur ve güven aralığının θ yı ihtiva

etmemesi olasılığını verir. Bir güven aralığının inşası, hesaplanan istatistiğin (yukarıdaki notasyona göre $\hat{\theta}$) örnekleme dağılımına dayanılarak gerçekleştirilir.

Popülasyon Aritmetik ortalaması İçin Güven Aralığı

Çekilen bir örneğin aritmetik ortalaması $\bar{y}$ yi göz önüne almak suretiyle, bir popülasyonun aritmetik ortalaması μ yü kestirmek istiyoruz. Bu örnek aritmetik ortalaması, olanaklı tüm örnek aritmetik ortalamaları uzayının yalnızca bir elemanıdır ve bir önceki bölümden biliyoruz ki eğer popülasyon normal dağılımlı veya *n* yeterince büyük ise,

1. tüm örnek aritmetik ortalamaları kümesi yaklaşık normal dağılır,
2. tüm örnek aritmetik ortalamaları kümesinin aritmetik ortalaması orijinal popülasyonun aritmetik ortalaması μ ye eşittir,
3. tüm örnek aritmetik ortalamaları kümesinin standart sapması $\sigma_{\bar{y}}$ yaklaşık olarak $\sigma/\sqrt{n}$ e eşittir; yani, orijinal popülasyonun standart sapmasının örnek hacminin kareköküne bölümüdür.



Buna göre, *z*-dağılımını kullanmak suretiyle, bir popülasyoun aritmetik ortalaması için güven düzeyini

$$P(\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

olarak yazabiliriz. Başka bir deyişle, μ için $\%100(1-\alpha)$ güven aralığı şudur:

$$\bar{y} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Böylece hesaplanacak bir güven aralığının interpretasyonu ise şöyle yapılır: Popülasyon parametresi μ nün $\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ ve $\bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ arasında olduğuna $\%100(1-\alpha)$ güvenimiz var.

Çoğu zaman, öğrenme-çalışmalarımızda popülasyonun standart sapması σ nın değerini bilmeyiz. Bu durumda, σ nın bir kestirimi olarak *örneğin standart sapması* s yi kullanmamız icabeder. Böyle olunca, $s_{\bar{y}} = s/\sqrt{n}$ i $\sigma_{\bar{y}} = \sigma/\sqrt{n}$ in bir kestirimi olarak kullanırız. Yani, $\sigma_{\bar{y}} \cong s_{\bar{y}}$ kabul ederiz. Yine bir önceki bölümden biliyoruz ki σ bilinmediği zaman, örnek hacmi ne olursa olsun, popülasyonun yaklaşık normal dağıldığını varsaymak suretiyle, standart normal eğri yerine *t*-dağılımını kullanabiliriz.

Spesifik bir örneğe ilişkin hesaplamalar yaparken ve neticelere varırken, önemli olan, yine biliyoruz ki örneğin bir basit rastgele örnek olması ve içinden çekildiği popülasyonun % 10 undan daha geniş olmaması gerektiğidir.

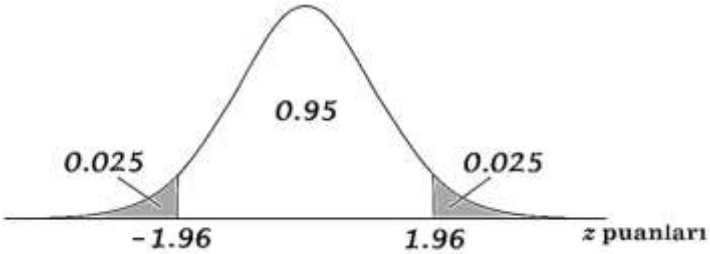
Misal 1 Bir şişeleme makinası 3.4 ml lik bir standart sapma ile çalışıyor. Otuz altı şişelik bir basit rastgele örnekte, makina her şişeyi ortalama 457.45 ml doldurmuştur.

a. Bu makina tarafından doldurulan tüm şişelerin içeriklerinin aritmetik ortalamasını kestir. Daha spesifik olarak, öyle bir aralık tespit et ki aritmetik ortalamanın bu aralıkta olduğuna %95 güvenelim.

Cevap: Örnek hacmi 36 olduğuna göre, örnek aritmetik ortalamaları,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{3.4}{\sqrt{36}} = 0.57 \text{ ml}$$

lik bir standart sapma ile yaklaşık normal dağılımlıdır. Daha önceki bölümlerde öğrenmiştik ki örnek aritmetik ortalamalarının %95 i popülasyon aritmetik ortalamasının sağlı-sollu 1.96 standart sapma civarında bulunmaktadır.



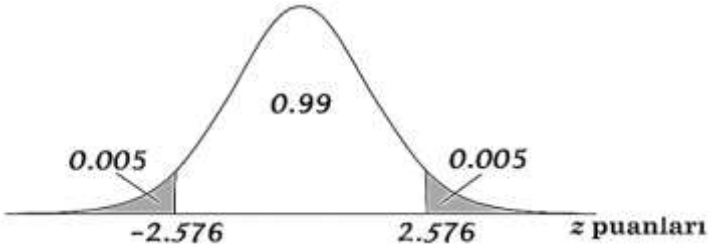
Diğer bir ifadeyle, popülasyon aritmetik ortalamasının herhangi bir örnek aritmetik ortalamasının sağlı-sollu 1.96 standart sapma civarında bulunduğuna %95 güvenimiz var. O halde,

$$457.45 \pm 1.96(0.57) = 457.45 \pm 1.117 \text{ ml}$$

ve tüm şişelerin içeriklerinin aritmetik ortalamasınının 456.33 ml ve 458.57 ml arasında olduğuna %95 güvenimiz var. Bunu *%95 lik güven aralığı kestirimi* olarak adlandırıyoruz.

b. Yüzde 99 luk güven aralığı kestirimi nedir?

Cevap: Yine biliyoruz ki örnek aritmetik ortalamalarının %99 u popülasyon aritmetik ortalamasının sağlı-sollu 2.576 standart sapma civarında bulunmaktadır.



Şimdiki halde,

$$457.45 \pm 2.576(0.57) = 457.45 \pm 1.468 \text{ ml}$$

ve tüm şişelerin içeriklerinin aritmetik ortalamasınının 455.98 ml ve 458.92 ml arasında olduğuna %99 güvenimiz var. Güven düzeyi tanımını kullanarak, aynı zamanda, eğer örnekleme prosedürü pek çok kere tekrarlanmış olsaydı, hesaplanacak güven aralıklarının yaklaşık %99 unun gerçek popülasyon aritmetik ortalamasını içinde bulunduracağını söyleyebiliriz (ki böylece metodun bizim hesapladığımız aralık için doğru sonuç vereceğine dair güvenimiz %99 olmaktadır).

Dikkatini çekmek isterim ki daha yüksek bir kesinlik (%95 yerine %99) istediğimiz zaman, daha geniş ve dolayısıyla daha az spesifik bir aralığa ($\pm$1.117 yerine $\pm$1.468) razı olmak zorundayız.

Misal 2 Belli bir fabrikanın ürettiği pillerin ömrüne ait varyans 5.76 ay^2 dir. Altmış dört pilden oluşan bir basit rastgele örnekte, pillerin ömürlerinin aritmetik ortalaması 12.35 ay olarak bulunmuştur.

a. Bu fabrika tarafından üretilen tüm pillerin ömürlerine ait %90 lık bir güven aralığı kestirimi belirt.

Cevap: Popülasyonun standart sapması $\sigma = \sqrt{5.76} = 2.4$ ve örnek aritmetik ortalamalarının standart sapması,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{2.4}{\sqrt{64}} = 0.3 \text{ ay.}$$

Popülasyon aritmetik ortalamasının %90 lık güven aralığı kestirimi,

$$12.35 \pm 1.645(0.3) = 12.35 \pm 0.4935 \text{ ay.}$$

Buna göre, %90 güvenimiz var ki pillerin aritmetik ortalama ömrü 11.8565 ay ve 12.8435 ay arasındadır.

b. Eğer 12.35 ay lık bir örnek aritmetik-ortalamsı, 100 pillik bir örneğe ait olsaydı, %90 lık güven aralığı kestirimi ne olurdu?

Cevap: Örnek aritmetik ortalamalarının standart sapması, o zaman,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{2.4}{\sqrt{100}} = 0.24 \text{ ay}$$

olurdu ve %90 lık güven aralığı kestirimi de,

$$12.35 \pm 1.645(0.24) = 12.35 \pm 0.3948 \text{ ay}$$

olurdu.

Dikkatini çekmek isterim ki örnek hacmi artırıldığında (64 yerine 100), örnek aritmetik ortalaması için daha dar ve dolayısıyla daha spesifik bir aralık ($\pm$0.4935 yerine $\pm$0.3948) elde ediyoruz.

Misal 3 Yeni geliştirilen bir ilacın, kalp atış hızını, dakikada 2.49 atışlık bir standart sapmayla düşürdüğü bildirilmiştir.

a. Eğer 50 hastalık bir basit rastgele örneğin kalp atış hızındaki düşüşüşün aritmetik ortalaması 5.32 atış/dakika ise, tüm hastaların kalp atış hızındaki düşüşüşün aritmetik ortalaması için %95 lik bir güven aralığı kestirimi belirt.

Cevap: Örnek aritmetik ortalamalarının standart sapması,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{2.49}{\sqrt{50}} = 0.352 \text{ atış/dakika.}$$

Yüzde 90 güvenimiz var ki bu ilacı kullanan hastaların kalp atış hızındaki düşüşüşün aritmetik ortalaması,

$$5.32 \pm 1.96(0.352) = 5.32 \pm 0.69$$

aralığında veya 4.63 atış/dakika ve 6.01 atış/dakika arasındadır.

b. Bu yeni ilacın kalp atış hızını 5.32$\pm$0.75 atış/dakika aralığında bir aritmetik ortalama ile düşürdüğünü, hangi kesinlikle (yani güven düzeyinde) söyleyebiliriz?

Cevap: Yapmamız gereken, $\pm$0.75 lik hata payına tekabül eden *z*-puanlarını hesaplamak,

$$\frac{\pm 0.75}{0.352} = \pm 2.13.$$

Tablo A da, bu *z*-puanları sırasıyla 0.0166 ve 0.9834 lük olasılıkları vermektedir. Buna göre, aradığımız güven düzeyi,

$$0.9834 - 0.0166 = 0.9668.$$

Başka bir ifadeyle, 5.32$\pm$0.75 atış/dakika, yeni ilacın verildiği hastaların kalp atış hızındaki düşüşün aritmetik ortalaması için %96.68 lik bir güven aralığı kestirimidir.

Anımsanacağı üzere, σ meçhul olduğu zaman (ki bu neredeyse her zamandır), *z*-dağılımı yerine *t*-dağılımını kullandığımızı söylemiştim. Dahası, eğer *t*-dağılımını kullanacaksak, orijinal popülasyonun dağılımının normal veya en azından yaklaşık normal olması gerektiğini de biliyoruz.

Örnek verinin içinden çekildiği popülasyonun dağılımının normal olduğu bilinmiyor ise, şöyle davranmamız gerekir. Eğer örnek geniş hacimliyse ($n \geq 30$), orijinal popülasyonun dağılımı ne olursa olsun *t*-dağılımını kullanabiliriz. Eğer örnek orta hacimliyse ($15 < n < 30$), *t*-dağılımını

kullanabilmemiz için örnek verinin tek-tepeli ve simetriğe yakın olması ve aynı zamanda aşırı gözlemleri barındırmaması gerekir. Eğer örnek dar hacimliyse ($n \leq 15$), t-dağılımını kullanabilmemiz için örnek verinin tek-tepeli ve simetrik olması ve aynı zamanda dışlakları barındırmaması gerekir.

Örnek verinin dağılımının normale uyup uymadığını kontrol etmek için frekans dağılımını gösteren herhangi bir grafiğine bakmak, aritmetik ortalamasını ve ortancasını karşılaştırmak, çarpıklık ve basıklık ölçülerini değerlendirmek veya başka uyarlı metotlara başvurmak yoluna gidilebilir, elbette. Fakat eğer elimizdeki örneğin hacmi 30 dan ve hatta 15 ten daha dar ise, bu pek te manalı bir çaba olmaz. Çünkü, bir örneğin içinden çekilmiş olduğu popülasyonun dağılımını gösterebilmesi için geniş hacimli bir örnek olması icabeder. Yani, hacmi en az 30 olmalıdır. Aksi halde, dar örneğin kendi dağılımının yerine, içinden çekilmiş olduğu popülasyonun dağılımının normale uyup uymadığını, olanaklar dairesinde, kontrol etmek daha makul bir tavırdır.

Misal 4 Yeni geliştirilen bir yumurtacı tavuk hattından alınan 10 tavuğun aylık yumurta verimlerinin aritmetik ortalaması 27.2 adet ve standart sapması 1.8 adet olarak bulunmuştur. Bu hattın aylık yumurta veriminin %95 lik bir güven aralığı kestirimi nedir? Bu hatta ait tüm tavukların aylık yumurta verimlerinin yaklaşık normal dağıldığını varsayıyoruz.

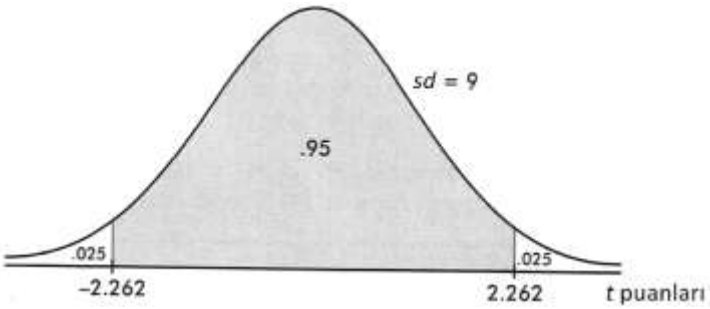
Cevap: Popülasyonun standart sapması meçhul ve bu sebeple t-dağılımını kullanacağız. Örnek aritmetik ortalamasının standart hatası,

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{1.8}{\sqrt{10}} = 0.569.$$

Serbestlik derecesi $10 - 1 = 9$ ve her bir kuyrukta %2.5 olduğuna göre, Tablo B yi kullanarak t-puanlarını ± 2.262 olarak buluruz. Böylece, yeni hattın aylık yumurta veriminin,

$$27.2 \pm 2.262(0.569) = 27.2 \pm 1.3$$

aralığında veya 25.9 adet ve 28.5 adet arasında olduğuna %95 güvenimiz var.



Misal 5 Altı evcil güvercinin gaga uzunlukları (burun deliğinin önünden gaga kemiğinin ucuna kadar) ölçülmüş ve milimetre cinsinden şu veri elde edilmiştir: {4.3, 5.2, 4.6, 4.9, 6.0 ve 5.6} Evcil güvercinlerin gaga uzunluklarının aritmetik ortalaması için %90 lık bir güven aralığı kestirimi hesapla. (Varsayıyoruz ki evcil güvercinlerin gaga uzunlukları yaklaşık normal dağılmaktadır.)

Cevap: Öncelikle, örnek aritmetik ortalamasını, standart sapmasını ve sonrasında örnek aritmetik ortalamasının standart hatasını (yani, örnekleme dağılımının standart sapmasının bir kestirimini) hesaplayalım:

$$\bar{y} = \frac{\sum y}{n} = \frac{4.3 + 5.2 + 4.6 + 4.9 + 6.0 + 5.6}{6} = \frac{30.6}{6} = 5.1 \text{ mm}$$

$$s = \sqrt{\frac{\sum (y - \bar{y})^2}{n - 1}} = \sqrt{\frac{0.8^2 + 0.1^2 + 0.5^2 + 0.2^2 + 0.5^2}{5}} = 0.632 \text{ mm}$$

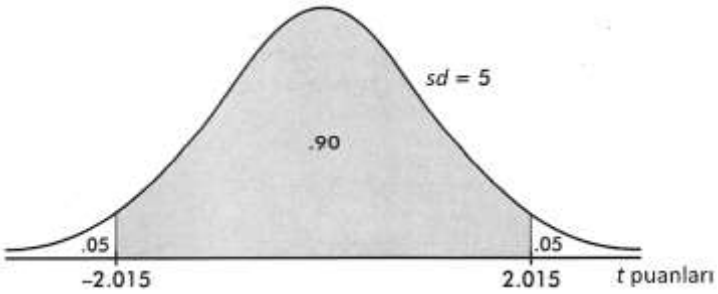
$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{0.632}{\sqrt{6}} = 0.258 \text{ mm}$$

Serbestlik derecesi $6 - 1 = 5$ ve her bir kuyrukta %5 olduğuna göre, Tablo B yi kullanarak

t-puanlarını ± 2.015 olarak buluruz. Böylece, %90 güvenle belirtebiliriz ki evcil güvercinlerin gaga uzunluklarının aritmetik ortalaması,

$$5.1 \pm 2.015(0.258) = 5.1 \pm 0.25$$

aralığında veya 4.58 mm ve 5.62 mm arasındadır.



Misal 6 Yüz yirmi beş bin ailenin yaşadığı bir şehirde, 250 aile ile bir bir örnek yoklaması yapılıyor. Aile başına aylık gelir ortalaması 3540 TL ve standart sapması 1150 TL olarak bulunuyor. Şehirdeki tüm ailelerin toplam aylık geliri için %99 luk bir güven aralığı kestirimi belirt. (Tüm ailelerin aylık gelirleri popülasyonunun yaklaşık normal dağıldığını varsay.)

Cevap: Verilenler $\bar{y} = 3540$ TL ve $s = 1150$ TL. Popülasyon standart sapması σ için s yi (yani örnek standart sapmasını) bir kestirim olarak kullanacağız. Buna göre, örnek aritmetik ortalamasının standart hatası,

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{1150}{\sqrt{250}} = 72.73 \text{ TL.}$$

Student *t*-dağılımını kullanacağız, çünkü σ bilinmiyor. Serbestlik derecesi $250 - 1 = 249$ ve her bir kuyrukta %0.5 olduğuna göre, Tablo B yi kullanarak *t*-puanlarını ± 2.596 olarak buluruz. Böylece, %99 güvenle belirtebiliriz ki ailelerin aylık gelirlerinin aritmetik ortalaması,

$$3540 \pm 2.596(72.73) = 3540 \pm 188.8$$

aralığındadır. Gel gör ki aradığımızı henüz bulmuş değiliz. Ailelerin aylık gelirlerinin toplamı için bir aralık arıyoruz. Öyle ise, şehirdeki tüm ailelerin toplam aylık geliri için %99 luk bir güven aralığı kestirimi,

$$125000 \times (3540 \pm 188.8) = 442500000 \pm 23600000$$

olarak elde edilir. Yani, bu şehirde yaşayan tüm ailelerin aylık toplam kazancının 418900000 TL ve 466100000 TL arasında olduğuna %99 güvenimiz var.

Görüldüğü üzere, bir güven aralığını iki yoldan ifade edebiliyoruz. Birincisinde, güven *aralığının* interpretasyonunu yapabiliriz ve popülasyon aritmetik ortalamasının mesala %90 güvenle 45.7 ile 51.2 arasında olduğunu söyleriz. İkincisinde, güven *düzeyininin* interpretasyonunu yapabiliriz ve eğer bu örnekleme prosedürü pek çok kere tekrarlanmış olsaydı, elde edeceğimiz güven aralıklarının yaklaşık mesela %90 ının popülasyon aritmetik ortalamasını içeriyor olacağını söyleriz. Doğru olmayan ifadelere de sıkça rastlanmaktadır. Bunlardan bazıları şöyledir: "Popülasyon aritmetik ortalaması 45.7 ve 51.2 arasındadır.", "Popülasyonun aritmetik ortalamasının 45.7 ve 51.2 arasında olması olasılığı 0.90 dır." veya "Eğer bu örnekleme prosedürü pek çok kere tekrarlanmış olsaydı, örnek aritmetik ortalamalarının yaklaşık %90 ı 45.7 ve 51.2 arasında olacaktı."

İstatistiksel prensipler sadece verinin analizi safhasında değil, aynı zamanda verinin toplanması safhasında da elzemdir ve bu safhadaki önemli mülahazalardan birisi örnek hacmidir. Popülasyon aritmetik ortalamalarına ait aralık kestirimleri yaparken, gördük ki her çıkarıma bir güven düzeyi ifadesi eşlik etmek zorundadır. Genel olarak, eğer daha dar, daha kesin bir aralık kestirimi istiyorsak, ya güven düzeyini azaltmamız ya da örnek hacmini genişletmemiz gerekir. Benzer şekilde, eğer güven düzeyini artırmak istiyorsak, ya daha geniş bir aralık kestirimine razı olmamız ya da örnek hacmini genişletmemiz gerekir. Bu minvalde, daha geniş bir örnek hacmi seçmenin her zaman arzu

edilir olduğu düşünülse bile, gerçek dünyada zaman ve maliyet telakkileri daima baş rolü oynar.

Misal 7 Bir fabrikada optik depolama cihazları için üretilen millerin çapına ait standart sapma 0.025 santimetredir. Eğer mil çapının aritmetik ortalamasının içinde bulunduğu ± 0.01 santimetrelik aralığı %99 güvenle bilmek istiyorsak, örnek hacmi ne olmalıdır?

Cevap: Elimizde olan,

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{0.025}{\sqrt{n}}$$

ve hata payını 0.01 santimetreden daha küçük yapacak örnek hacmini bulmak istiyoruz. Yani, her bir kuyrukta %0.5 olduğuna göre, Tablo A yı kullanarak z-puanlarını ± 2.576 olarak buluruz ve şöyle yazabiliriz:

$$2.576\sigma_{\bar{y}} \leq 0.01.$$

Böylece,

$$2.576\frac{0.025}{\sqrt{n}} \leq 0.01$$

Cebirsel olarak, ifadeyi yeniden düzenlersek, şunu buluruz:

$$\sqrt{n} \geq \frac{2.576(0.025)}{0.01} = 6.44$$

ve buna göre, $n \geq 41.5$. Öyleyse, seçeceğimiz örneğin hacmi 42 dir.

Yukarıdaki misalde örnek hacmini bulmak için yürüttüğümüz mantığı, örnek aritmetik ortalaması için yapılacak güven aralığı kestirimlerinde, arzu edilen güven düzeyinde ve arzu edilen hata payı ile, gerekli olacak örnek hacmini hesaplamak üzere formülleştirirsek, şöyle yazabiliriz:

$$n \geq \left(\frac{z\sigma}{\text{hata payı}}\right)^2$$

Popülasyon Orantısı İçin Güven Aralığı

Tek bir örneğin orantısı p yi göz önüne almak suretiyle bir popülasyonun orantısı π yi kestirmek istiyoruz. Bu örnek orantısı, tüm örnek orantıları uzayının yalnızca bir elemanıdır ve bir önceki bölümden biliyoruz ki eğer n yeterince büyük ise,

1. tüm örnek orantıları kümesi yaklaşık normal dağılır,
2. tüm örnek orantıları kümesinin aritmetik ortalaması μ_p orijinal popülasyonun aritmetik ortalaması π ye eşittir,
3. tüm örnek orantıları kümesinin standart sapması σ_p yaklaşık olarak $\sqrt{\pi(1-\pi)/n}$ e eşittir.

Popülasyon orantısı π nin güven aralığını bulurken, π yi bilmediğimize göre, $\sqrt{\pi(1-\pi)/n}$ i nasıl bulacağız? En mantıklı prosedür elbette π nin bir kestirimi olarak örnek orantısı p yi kullanmaktır. Buna göre örnek orantısının standart hatası,

$$s_p = \sqrt{\frac{p(1-p)}{n}}.$$

Önceki bölümlerden biliyoruz ki hem np hem de $n(1-p)$ 10 dan büyük olduğunda, normal dağılım binom dağılımı için çok yararlı bir yaklaşımdır. Spesifik bir örneğe ilişkin hesaplamalar yaparken ve neticelere varırken, önemli olan, yine biliyoruz ki örneğin bir basit rastgele örnek olması ve içinden çekildiği popülasyonun % 10 undan daha geniş olmaması gerektiğidir. (Eğer popülasyon dar ise ve örnek hacmi popülasyonun %10 unu geçerse, o halde normalden başka modeller daha mütenasiptir.)

Misal 8 Bir şehirdeki seçmenlerin 550 kişilik bir basit rastgele örneğinin %64 ü mevcut belediye başkanının politikalarını tasvip ettiklerini belirtmişlerdir. Söz konusu şehirdeki tüm seçmenlerin belediye başkanının politikalarını tasvip edenlerinin orantısı için %99 luk bir güven aralığı kestirimi belirle.

Cevap: Bakıyoruz, $np = 550(0.64) = 352 > 10$ ve $n(1-p) = 550(0.36) = 198 > 10$, örnek bir basit rastgele örnek ve şehirdeki tüm seçmenlerin sayısı verilmemekle beraber 550 nin bu sayının %10 undan daha büyük olmadığını varsayabiliriz. Örnek orantısı $p = 0.64$ olduğuna göre, örnek orantısının standart hatası,

$$s_p = \sqrt{\frac{0.64(0.36)}{550}} = 0.0205.$$

Daha önce öğrenmiş olduğumuz üzere, örnek orantılarının %99 u popülasyon orantısının sağlı-sollu 2.576 standart sapma civarında yer alır. Aynı manada, %99 güvenimiz var ki popülasyon orantısı herhangi bir örnek orantısının sağlı-sollu 2.576 standart sapma civarında yer alır. Böylece, popülasyon orantısının %99 luk güven aralığı kestirimi,

$$0.64 \pm 2.576(0.0205) = 0.64 \pm 0.053.$$

Bu sonuca göre, *hata payı* 0.053 tür. Yani, belediye başkanının politikalarını tasvip edenlerin orantısı, %99 güvenle, %58.7 ve %69.3 arasındadır.

Güven düzeyi tanımını kullanarak, aynı zamanda, söyleyebiliriz ki eğer bu anket prosedürü (yani, 550 kişilik basit rastgele örnek) pek çok kere tekrarlanmış olsaydı, elde edilecek güven aralıklarının yaklaşık %99 u gerçek popülasyon orantısının kapsayacaktı.

Misal 9 Bir üretim bandından alınan bir basit rastgele örnekteki 225 parçanın 18 inde kusur tespit edilmiştir. Bu bantta üretilen tüm kusurlu parçaların orantısı için %95 lik bir güven aralığı kestirimi belirt.

Cevap: Bakıyoruz, $np = 18 > 10$ ve $n(1-p) = 225 - 18 = 207 > 10$, örnek bir basit rastgele örnek ve söz konusu üretim bandında üretilen tüm parçaların sayısı verilmemekle beraber örnek hacminin popülasyonun %10 undan daha dar olduğunu varsayabiliriz. Örnek orantısının standart hatası,

$$s_p = \sqrt{\frac{0.08(0.92)}{225}} = 0.0181.$$

Tablo A dan kritik *z*-puanını 1.96 olarak buluruz. Buna göre, popülasyon orantısının %95 lik güven aralığı kestirimi,

$$0.08 \pm 1.96(0.0181) = 0.08 \pm 0.035.$$

Böylece, %95 güvenimiz var ki söz konusu batta üretilen kusurlu parçaların orantısı 0.45 ve 0.115 arasındadır.

Eğer bu üretim bandında 100000 parça üretildiğini biliyor olsaydık, orantıları kusurlu parça sayılarına dönüştürebilirdik:

$$(0.045)100000 = 4500 \qquad \text{ve} \qquad (0.115)100000 = 11500$$

Yani, söz konusu üretim bandında üretilen 100000 parça içerisinde kusurlu olanların sayısının 4500 ve 11500 arasında olduğunu %95 güvenle söyleyebiliriz.

Misal 10 Ülke çapında 1000 hemşire ile yapılan bir örnek yoklamasında, hemşirelere mesleklerinden memnun olup olmadıkları soruldu.

a. Yoklanan hemşirelerin 832 si mesleğinden memnuniyetini dile getirdiğine göre, ülkede istihdam edilen tüm hemşirelerin $\%83.2 \pm \%3$ ünün mesleğinden memnuniyet duyduğunu hangi gevenle öne sürebiliriz?

Cevap: Görüldüğü üzere, $np = 832 > 10$ ve $n(1-p) = 1000 - 832 = 168 > 10$, bir basit rastgele örnek alındığını ve örnek hacminin muvazzaf hemşireler popülasyonunun %10 undan daha dar olduğunu varsayıyoruz.

$$p = \frac{832}{1000} = 0.832 \qquad \text{ve} \qquad s_p = \sqrt{\frac{0.832(0.168)}{1000}} = 0.0118\,.$$

Hata payını ve standart hatayı bildiğimize göre, kritik z-puanlarını şöyle hesaplarız:

$$\pm\text{hata payı} = \pm z(s_p) \qquad \text{ise} \qquad \pm\frac{0.03}{0.0118} = \pm 2.54.$$

Tablo A bu kritik z-puanları için bize 0.0055 ve 0.9945 olasılıklarını verir. Böylece, aradığımız cevap $0.9945 - 0.0055 = 0.989$ dur. Başka bir ifadeyle, $\%83.2 \pm \%3$, muvazzaf hemşireler popülasyonu içerisinde mesleğinden memnun olanların yüzdesi için %98.9 luk bir güven aralığı kestirimidir.

b. Ülkemizde muvazzaf hemşirelerin sayısının kabaca 200000 olduğunu farzederek, mesleğinden memnun olanların sayısını hesapla.

Cevap: Eğer muvazzaf hemşire popülasyonunun hacmi 200000 ise, mesleğinden memnun olanların sayısı, %98.9 luk bir güvenle,

$$(0.832 - 0.03)200000 = 160400$$

ve

$$(0.832 + 0.03)200000 = 172400$$

arasındadır.

Misal 11 Amerika Birleşik Devletlerinde yapılan bir örnek yoklamasına yanıt veren 6230 yetişkin işsiz, medeni hallerine, cinsiyetlerine ve ırklarına göre aşağıdaki gibi sınıflandırılmıştır (Dul: Eşi ölmüş, Bekar: Hiç evlenmemiş veya boşanmış):

Beyaz			
	Evli	Dul	Bekar
Adam	1090	337	1168
Kadın	952	423	632

Beyaz-olmayan			
	Evli	Dul	Bekar
Adam	266	135	503
Kadın	189	186	349

a. Evli işsiz adamların orantısı için %90 lık bir güven aralığı kestirimi hesapla.

Cevap: Her iki tablonun birinci satırlarını toplayarak, yoklamaya katılan adamların sayısını 3499 olarak buluruz; adamların evli olanlarının sayısı ise $1090 + 266 = 1356$ dır. Görüyoruz ki, $np = 1356 > 10$ ve $n(1-p) = 3499 - 1356 = 2143 > 10$; örneğin bir basit rastgele örnek olduğunu varsayıyoruz ve örnek hacminin tüm işsiz yetişkinler popülasyonun %10 undan daha dar olduğu açıktır. Buna göre,

$$p = \frac{1356}{3499} = 0.3845 \qquad \text{ve} \qquad s_p = \sqrt{\frac{0.3875(0.6125)}{3499}} = 0.08236\,.$$

Örnek orantılarının dağılımı yaklaşık normal ve her bir kuyrukta %5 olacağı için, Tablo A yı kullanarak kritik z puanlarını ± 1.645 olarak bulduktan sonra, %90 lık güven aralığı kestirimi,

$$0.3875 \pm 1.645(0.008236) = 0.3875 \pm 0.0135.$$

b. Bekar işsiz kadınların orantısına ait %98 lik bir güven aralığı kestirimi yap.

Cevap: Yoklamada bekarların sayısı $1168 + 632 + 503 + 349 = 2652$ dir. Bunların, $632 + 349 = 981$ i kadındır. Yine, $np = 981 > 10$ ve $n(1-p) = 2652 - 981 = 1671 > 10$; örneğin bir basit rastgele örnek olduğunu varsayıyoruz ve örnek hacminin tüm işsiz yetişkinler popülasyonun %10 undan daha dar olduğu açıktır. Buna göre,

$$p = \frac{981}{2652} = 0.3699 \qquad \text{ve} \qquad s_p = \sqrt{\frac{0.3699(0.6301)}{2652}} = 0.09375\,.$$

Her bir kuyrukta bu defa %1 olacağı için, Tablo A yı kullanarak kritik z-puanlarını ± 2.326 olarak bulduktan sonra, %98 lik güven aralığı kestirimi,

$$0.3699 \pm 2.326(0.009375) = 0.3699 \pm 0.0218.$$

Bir yoklamanın dizaynında önemli mülahazalardan birisi de örnek hacminin seçimidir. Örnek orantısının daha dar, daha kesin bir aralık kestirimini elde edebilmek için, ya güven düzeyini azaltmamız ya da örnek hacmini genişletmemiz gerekir. Benzer şekilde, eğer güven düzeyini artırmak istiyorsak, ya daha geniş bir aralık kestirimine razı olmamız ya da örnek hacmini genişletmemiz gerekir. Yine, daha geniş bir örnek hacmi seçmenin her zaman arzu edilir olduğu düşünülse bile, gerçek dünyada zaman ve maliyet telakkileri daima baş rolü oynar.

Popülasyon orantısının bir güven aralığı kestirimini elde etmek için dizayn edilecek bir yoklamada, örnek hacmini tayin ederken, araştırmacının σ_p için hangi değeri kullanacağı kritik bir sorudur. Bu soruya cevap vermek için, öncelikle $\sqrt{\pi(1-\pi)}$ nin ne kadar büyük olabileceğini mülahaza etmemiz gerekir. Çeşitli π değerleri için $\sqrt{\pi(1-\pi)}$ nin alabileceği değerler şöyledir:

π	:	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\sqrt{\pi(1-\pi)}$	:	0.3	0.4	0.458	0.49	0.5	0.49	0.458	0.4	0.3

Tabloya bakınca, görüyoruz ki aradığımız cevap 0.5 tir. Böylece, $\sqrt{\pi(1-\pi)/n}$ in alabileceği değer en büyük $0.5/\sqrt{n}$ olabilecektir. Misal 12 ve Misal 13 te örnek hacimlerini belirlerken ve benzeri problemlerin çözümünde bu olguyu kullanacağız.

Misal 12 Ailesel akdeniz ateşi çeşitli varyasyonları olan bir genetik mutasyondur. Akdeniz havzasında yaşayan toplumlarda yaygındır. Eğer bir bilimsel araştırmacı Türk toplumunda ailesel akdeniz ateşinin bilinen mutasyonlarının orantısını belirlemek isterse, %96 güvenle ± 0.03 lük bir güven aralığı kestirimi yapabilmesi için, teste tabi tutacağı örnekte kaç kişi olmalıdır?

Cevap: Her bir kuyrukta %2 olacağı için, Tablo A yı kullanarak aradığımız güven aralığını sınırlayan z-puanlarını ± 2.05 olarak bulduktan sonra, şunu bilmek istiyoruz:

$$2.05\sigma_p \leq 0.03.$$

Hemen yukarıda belirtmiş olduğumuz üzere, σ_p nin alabileceği en büyük değer $0.5/\sqrt{n}$ dir. Buna göre, düşüncemizi şu yönde geliştirmemiz gayet mantıklıdır:

$$2.05\frac{0.5}{\sqrt{n}} \leq 0.03.$$

Cebirsel olarak, ifadeyi n in çözümü için yeniden düzenlersem,

$$\sqrt{n} \geq \frac{2.05(0.5)}{0.03} = 34.17$$

ve

$$n \geq 1167.4.$$

Demek ki 1168 kişilik bir basit rastgele örnek seçmek, bilimsel araştırmacımıza ailevi akdeniz

ateşinin bilinen mutasyonlarına sahip kişilerin orantısını, %96 güvenle, ± 0.03 aralığında verecektir.

Yeri gelmişken, ifade etmekte yarar var ki *kestirimin isabetliliği popülasyonun yüzde kaçını örneklendiğimize bağlı değildir*. Kritik olan, örneğin mutlak hacmidir. Kullandığımız prosedürlerin manalı olabilmesi için gerekli olan bir minimum n değeri var mıdır? Binomiyale normal yaklaşımı kullandığımız için, hem $n\pi$ nin hem de $n(1-\pi)$ nin en az 10 olması gerekir.

Misal 13 Öğrenci merkezli öğretimin lehinde olan öğretmenlerin orantısını belirlemek üzere bir öğrenme-çalışmasına girişilecektir. Eğer ± 0.06 aralığında, %99 güven düzeyiyle bir kestirim arzu ediliyor ise, kaç öğretmenle anket yapılmalıdır? Aynı güven düzeyinde, ± 0.03 aralığı için kaç anket? ± 0.02 aralığı için kaç anket?

Cevap: Cebirsel olarak,

$$2.576\left(\frac{0.5}{\sqrt{n}}\right) \leq 0.06$$

ifadesini yeniden düzenlersek,

$$\sqrt{n} \geq \frac{2.576(0.5)}{0.06} = 21.5 \qquad \text{ve} \qquad n \geq 462.25.$$

Benzerce,

$$2.576\left(\frac{0.5}{\sqrt{n}}\right) \leq 0.03$$

ifadesini yeniden düzenlersek,

$$\sqrt{n} \geq \frac{2.576(0.5)}{0.06} = 42.9 \qquad \text{ve} \qquad n \geq 1840.4.$$

Yine aynı şekilde,

$$2.576\left(\frac{0.5}{\sqrt{n}}\right) \leq 0.02$$

ifadesini yeniden düzenlersek,

$$\sqrt{n} \geq \frac{2.576(0.5)}{0.02} = 64.4 \qquad \text{ve} \qquad n \geq 4147.4.$$

Böylece, aralık kestirimi için arzu edilen isabetliliğe bağlı olarak, 463, 1841 veya 4148 öğretmenle anket yapılmalıdır.

Dikkatinizi çekmek isterim ki aralık kestirimini yarıya indirmekle (yani, ± 0.06 dan ± 0.03 e), örnek hacmini dörde katlamış olduk. Aralık kestirimini üçte birine indirmekle (yani, ± 0.06 dan ± 0.02 ye), örnek hacmini dokuza katlamış olduk (yuvarlama hataları nedeniyle cevaplar tam değil).

Daha genel olarak, güven düzeyi aynı kaldığı halde, aralık kestirimini bir sabite bölersek, örnek hacmini o sabitin karesi kadar artırmamız gerekir.

Yukarıdaki misalde örnek hacmini bulmak için yürüttüğümüz mantığı, örnek orantısısı için yapılacak güven aralığı kestirimlerinde, arzu edilen güven düzeyinde ve arzu edilen hata payı ile, gerekli olacak örnek hacmini hesaplamak üzere formülleştirirsek, şöyle yazabiliriz:

$$n \geq \left(\frac{0.5z}{\text{hata payı}}\right)^2$$

İki Popülasyon Aritmetik ortalaması Arasındaki Fark İçin Güven Aralığı

Önceki bölümlerden, $\bar{y}_1 - \bar{y}_2$ nin örnekleme dağılımı hakkında şu enformasyona sahibiz:

1. tüm örnek aritmetik ortalamaları farkları kümesi yaklaşık normal dağılır,
2. tüm örnek aritmetik ortalamaları farkları kümesinin aritmetik ortalaması popülasyon aritmetik ortalamalarının farkı $\mu_1 - \mu_2$ ye eşittir,
3. tüm örnek aritmetik ortalamaları farkları kümesinin standart sapması $\sigma_{\bar{y}_1-\bar{y}_2}$ yaklaşık olarak $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ ye eşittir.

Spesifik örneklere ilişkin hesaplamalar yaparken ve neticelere varırken, örneklerin her ikisinin de birer *basit rastgele örnek* olmaları ve biribirlerinden *bağımsız olarak* çekilmeleri gerektiğini bilmek önemlidir.

Misal 14 Aynı endüstri kolunda üretim yapan iki ayrı fabrikada yürütülen 30 aylık bir öğrenme-çalışması ile, bu fabrikalarda meydana gelen aylık kaza sayıları arasındaki fark belirlenmek istenmektedir. Birinci fabrikanın aylık kaza sayılarının aritmetik ortalaması 12.3 ve standart sapması 3.5 iken ikinci fabrikanın aritmetik ortalaması 7.6 ve standart sapması 3.4 tür. Bu iki fabrikada meydana gelen aylık kaza sayıları arasındaki fark için %95 lik bir güven aralığı kestirimi belirle. (Varsayıyoruz ki iki popülasyon bağımsızdır ve yaklaşık normal dağılmaktalar.)

Cevap:

$$\begin{array}{ll} n_1 = 30 & n_2 = 30 \\ \bar{y}_1 = 12.3 & \bar{y}_2 = 7.6 \\ s_1 = 3.5 & s_2 = 3.4 \end{array}$$

$$s_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{3.5^2}{30} + \frac{3.4^2}{30}} = 0.89$$

Gözlenen fark, $12.3 - 7.6 = 4.7$, ve $(n_1 - 1) + (n_2 - 1) = 58$ lik serbestlik derecesi ile kritik *t*-puanı ± 2.00 dir. Böylece, güven aralığı kestirimi $4.7 \pm 2.00(0.89) = 4.7 \pm 1.78$ dir. Yüzde 95 güvenimiz var ki birinci fabrikada meydana gelen aylık kaza sayısı, ikinci fabrikaya göre, 2.92 ve 6.48 arasında daha çoktur. Serbestlik derecesi konusunda bazı istatistik bilimcileri daha tutucu davranılması ve serbestlik derecesinin $min(n_1 - 1, n_2 - 1)$ olarak alınmasını teklif ederler. Bu durumda, iki örneğin serbestlik derecelerinden en küçüğü tercih edileceğinden, serbestlik derecemiz $min(30 - 1, 30 - 1) = 29$ olur ve bu da bize $4.7 \pm 2.045(0.89) = 4.7 \pm 1.82$ lik bir güven aralığı kestirimi verir.

Misal 15 Akaryakıt istasyonlarındaki marketlerde satılan bakkaliye ürünlerinin şehir marketlerine göre daha pahalıya satıldığı iddiasını test etmek üzere bir yoklama yapılır. Önceden seçilen bir grup ürün rastgele örneklenen 35 akaryakıt istasyonu marketinden ve 45 şehir marketinden satın alınır. Şu veri elde edilir.

İstasyon marketleri	Şehir marketleri
$n_1 = 35$	$n_2 = 45$
$\bar{y}_1 = 39.4$ TL	$\bar{y}_2 = 36.52$ TL
$s_1 = 1.23$ TL	$s_2 = 1.1$ TL

Bakkaliye fiyatları arasındaki fark için %90 lık bir güven aralığı kestiriminde bulun. (Varsayıyoruz ki iki popülasyon bağımsızdır ve yaklaşık normal dağılmaktalar.)

Cevap:

$$s_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{1.23^2}{35} + \frac{1.1^2}{45}} = 0.265$$

Gözlenen fark, $39.4 - 36.52 = 2.88$, ve $(35 - 1) + (45 - 1) = 78$ lik serbestlik derecesi ile kritik *t*-puanı ± 1.664 tür. Böylece, güven aralığı kestirimi $2.88 \pm 1.664(0.265) = 2.88 \pm 0.44$ tür. Yüzde 90 güvenle söyleyebiliyoruz ki seçili ürünlerin fiyatı, şehir marketlerine göre, istasyon marketlerinde

2.44 TL ve 3.32 TL arasında daha pahalıdır.

Misal 16 İki ayrı marka boyanın kuruma süreleri arasındaki farkı belirlemek üzere dizayn edilecek bir öğrenme-çalışmasında, her iki popülasyon için kutuların standart sapması 2.5 dakika ise, farkın 1 dakikalık hata payı içerisinde olduğundan %98 güvenle emin olabilmek için, örnek hacmi ne olmalıdır?

Cevap:

$$\sigma_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{2.5^2}{n} + \frac{2.5^2}{n}} = \frac{3.536}{n}$$

Tablo A yı kullanarak kritik z-puanını 2.325 olarak bulduktan sonra, hata payı 1 dakika olduğuna göre,

$$2.325\frac{3.536}{\sqrt{n}} \leq 1$$

ifadesini yeniden yazarak, $\sqrt{n} \geq 8.22$ ve $n \geq 67.6$ değerini elde ederiz. Böylece, araştırmacı her markadan 68 kutuluk birer örneği test etmelidir.

Örneklerin bağımsızlığı durumunda, eğer popülasyonların hem normal dağılımlı olmaları hem de eşit varyanslı olmaları varsayımları geçerli ise, örnek varyanslarını *karmak* suretiyle, ortak popülasyon varyansının daha iyi bir kestirimi elde edilebilir.

Eşli Farklar

Bu alt bölümde, buraya kadar, yaptığımız analiz ve uyguladığımız prosedürler için bir varsayım olarak, iki örneğin biribirinden bağımsız olması gerektiğini ifade ettik. Mamafih, pek çok istatistiksel deney ve test verinin doğal olarak çiftler halinde bulunduğu iki popülasyonu karşılaştırmak üzere yürütülür. Böylesi *eşli farklar* durumunda, daha uygun bir prosedür, eşli verinin farklarından oluşan tek değişken üzerinde bir tek örnek analizi yapmaktır.

Misal 17 Doping etkisi yaptığı düşünülen bir ilaç 30 atlet üzerinde test edilmiştir. Müteakip tablo atletlerin ilaç almadan ve ilaç aldıktan sonra 5000 metre koşusu performanslarını (saniye) vermektedir.

Atlet	1	2	3	4	5	6	7	8	9	10
İlaçsız performas	1025	1085	1350	1202	982	950	1089	1159	1246	1135
İlaçlı performans	912	1025	1295	1123	875	890	1002	998	1235	1045
fark	-113	-60	-55	-79	-107	-60	-87	-161	-11	-90

Atlet	11	12	13	14	15	16	17	18	19	20
İlaçsız performas	1302	1010	1235	1103	1200	1187	910	1049	1123	901
İlaçlı performans	1215	948	1190	1077	1223	1100	842	985	1107	847
fark	-87	-62	-45	-26	23	-87	-68	-64	-16	-54

Atlet	21	22	23	24	25	26	27	28	29	30
İlaçsız performas	1005	1010	1015	1032	1310	1086	1276	1029	1221	874
İlaçlı performans	956	987	1028	954	1152	987	1228	1005	1166	808
fark	-49	-23	13	-78	-158	-99	-48	-24	-55	-66

Koşu performanslarındaki aritmetik ortalama artış için, %90 lık bir güven aralığı kestiriminde bulun.

Cevap: Koşu performanslarının iki kümesi (ilaçsız ve ilaçlı) bağımsız değil; atletler için eşlilik ilişkisine sahipler. Bu durumda münasip prosedür, farklar kümesinden oluşan tek bir değişken için

tek örnek analizidir. Böylece, atletlerin ilaçlı ve ilaçsız performans farkları kümesi için, $\bar{y} = -63.2$ saniye, $s = 41.89$ saniye ve

$$s_{\bar{y}} = \frac{41.89}{\sqrt{30}} = 7.648 \text{ saniye.}$$

Yüzde 90 lık güven aralığı kestirimi,

$$-63.2 \pm 1.699(7.648) = -63.2 \pm 13.$$

Yüzde 90 güvenle, koşu performanslarındaki aritmetik ortalama artış –50.2 saniye ve –76.2 saniye arasındadır.

Ne zaman iki örnek analizi ve ne zaman tek örnek analizi yapacağına karar verirken, kafa karışıklığı yaşayabilirsin. Ana fikir, grupların bağımsız olup olmadığıdır. Ayrımı yaratan, öğrenme-çalışmasının dizaynıdır. Kadınların gebelikten korunma yöntemleri hakkındaki bilgisini artırmaya yönelik bir eğitim programının etkinliğini ölçmek istiyoruz. Eğer deney dizaynımız, gönüllü kadınların rastgele seçtiğimiz yarısına eğitim programını uyguladıktan sonra, bu kadınların standart bir testten aldıkları puanları kadınların geri kalanlarının puanlarıyla karşılaştırmak ise, o zaman yapılması icabeden iki örnek analizidir. Mamafih, eğer deney dizaynımız, gönüllü kadınların tümünü standart teste tabi tuttuktan sonra, bu kadınların tümüne eğitim programını uygulamak ve yeniden standart teste tabi tutmak ise, o zaman yapılması icabeden farkların tek örnek analizidir. Eşleme yukarıda zikrettiğimiz üzere, aynı birey üzerinde iki ölçüm yapmakla meydana gelebildiği gibi ikiz kümeleri üzerinde ölçüm yapmakla ya da aynı konumu veya aynı durumu paylaşan çiftler üzerinde ölçüm yapmakla da meydana gelebilir. Ana fikir, ölçümlerin iki kümesinin bağımsız olup olmadığı, ya da mülahaza edilen soru bakımından bir şekilde ilişkili olup olmadıklarıdır.

İki Popülasyon Orantısı Arasındaki Fark İçin Güven Aralığı

Önceki bölümlerden, $p_1 - p_2$ nin örnekleme dağılımı hakkında şu enformasyona sahibiz:

1. tüm örnek orantıları farkları kümesi yaklaşık normal dağılır,
2. tüm örnek orantıları farkları kümesinin aritmetik ortalaması $\mu_{p_1-p_2}$ popülasyon orantılarının farkı $\pi_1 - \pi_2$ ye eşittir,
3. tüm örnek orantıları farkları kümesinin standart sapması $\sigma_{p_1-p_2}$ yaklaşık olarak $\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$ ye eşittir.

Unutma ki binomiyal dağılıma normal dağılımla yaklaşıyoruz ve bu sebeple n_1p_1, $n_1(1-p_1)$, n_2p_2 ve $n_2(1-p_2)$ değerlerinden her biri en az 10 olmalıdır. Spesifik örneklerle hesaplama yaptığımız ve netice çıkardığımız için, örneklerin hem basit rastgele örnekler olmaları hem de biribirinden bağımsızca çekilmiş olmaları önemlidir. Nihayetinde, orijinal popülasyonların da örnek hacimlerine kıyasla geniş olmaları gerekir.

Misal 18 Şehir hastahanelerinde gündüz vardiyasında çalışan 125 hemşirenin %84 ü işinden memnun olduğunu ifade ederken, gece vardiyasında çalışan 150 hemşirenin yalnızca %72 si memnuniyetini ifade etmiştir. İki örnek orantısı arasındaki fark için %90 lık bir güven aralığı kestirimi tesis et.

Cevap: Görüyoruz ki $n_1p_1 = 125(0.84) = 105$, $n_1(1-p_1) = 125(0.16) = 20$, $n_2p_2 = 150(0.72) = 108$, $n_2(1-p_2) = 150(0.28) = 42$ ve hepsinin değeri 10 dan daha büyük; bağımsız basit rastgele örnekler alınmış ve şehir hastahanelerinde çalışan hemşirelerin popülasyonunu geniş bir popülasyon olarak varsayıyoruz.

$$n_1 = 125, n_2 = 150$$
$$p_1 = 0.84, p_2 = 0.72$$

$$s_{p_1-p_2} = \sqrt{\frac{0.84(0.16)}{125} + \frac{0.72(0.28)}{150}} = 0.0492\,.$$

Gözlenen fark, $0.84 - 0.72 = 0.12$ ve kritik z puanları, ± 1.45 tir. Güven aralığı kestirimi,

$$0.12 \pm 1.645(0.0492) = 0.12 \pm 0.081.$$

Yüzde 90 güvenle, işinden memnun olan hemşirelerin orantısı, gündüz vardiyasında çalışanlarda, gece vardiyasında çalışanlara göre, %3.9 ve %20.1 arasında daha yüksektir.

Misal 19 Bir sığır ırkında, rastgele örneklenen 85 erkek sığırın 10 u boynuzsuz iken, rastgele örneklenen 92 dişi sığırın 37 si boynuzsuzdur. Boynuzsuz erkek ve dişi sığırların orantıları arasındaki fark için %95 lik bir güven aralığı kestiriminde bulun.

Cevap: Görüyoruz ki $n_1p_1 = 10$, $n_1(1-p_1) = 75$, $n_2p_2 = 37$, $n_2 \times (1-p_2) = 55$ ve hepsi 10 dan daha büyük; bağımsız basit rastgele örnekler alınmış ve erkek ve dişi sığır popülasyonları hayli geniş.

$$n_1 = 85, n_2 = 92$$
$$p_1 = \frac{10}{85} = 0.118, p_2 = \frac{37}{92} = 0.402$$

$$s_{p_1-p_2} = \sqrt{\frac{0.118(0.882)}{85} + \frac{0.402(0.598)}{92}} = 0.0619\,.$$

Gözlenen fark, $0.118 - 0.402 = -0.284$ ve kritik z-puanları, ± 1.96 dır. Buna göre, güven aralığı kestirimi,

$$-0.284 \pm 1.96(0.0619) = -0.284 \pm 0.121.$$

Boynuzsuz erkek sığırların orantısının, boynuzsuz dişi sığırların orantısından 0.163 ve 0.405 arasında daha düşük olduğunu, %95 güvenle söyleyebiliriz.

Misal 20 Bir yoklamacı bir sonraki seçimde sandığa gidecek kişiler içerisinde düşük gelir düzüyindekilerin ve yüksek gelir düzeyindekilerin orantıları arasındaki farkı belirlemek istiyor. Bu farkı 0.02 lik bir hata payı ile %95 lik güven düzeyinde kestirebilmek için, alınması gereken örneklerin hacimleri ne olmalıdır?

Cevap: Evvela şunu anımsayalım ki bir örneği çekerken, $\sqrt{\pi(1-\pi)}$ ifadesinin alabileceği en büyük değer 0.5 tir. Böylece, iki örnek orantısı arasındaki farkın standart sapması için, şunu yazabiliriz,

$$\sqrt{\pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \leq 0.5\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\,.$$

Buna göre, eğer basitçe durumu idealize ederek, $n_1 = n_2 = n$ koşulunu öne sürersek, yukarıdaki ifadeyi şuna indirgeriz:

$$0.5\sqrt{\frac{1}{n} + \frac{1}{n}} = 0.5\sqrt{\frac{2}{n}} = \frac{0.5\sqrt{2}}{\sqrt{n}}$$

Hülasa, çekilecek iki örneğin hacimlerinin eşit olacağını varsayarak,

$$\sigma_{p_1-p_2} \leq \frac{0.5\sqrt{2}}{\sqrt{n}}$$

ve

$$1.96\sigma_{p_1-p_2} \leq 0.02$$

yazabiliriz; ve bu iki ifadeyi yeniden düzenlersek,

$$\frac{1.96(0.5)\sqrt{2}}{\sqrt{n}} \leq 0.02.$$

Nihayetinde,

$$\sqrt{n} \geq \frac{1.96(0.5)\sqrt{2}}{0.02}$$

ve

$$n \geq 69.3^2 = 4802.5.$$

Böylece, yoklamacının kullanması gereken örnek hacmi, her bir örnek için en az 4803 kişi olmalıdır.

Popülasyon Regresyon Doğrusunun Eğimi İçin Güven Aralığı

Önceki bölümlerde, en küçük kareler regresyon doğrusunu tartışmıştık,

$$\hat{y} = a + bx$$

ki buradaki b doğrunun eğimidir. İki değişkeli veri için bu eğimi bir kompüter yazılımı ile halihazırda kolayca hesaplamak mümkün. Bir örnekten hesaplanan b değeri, bu örneğin içinden çekilmiş olduğu popülasyonuna ait regresyon doğrusunun eğimi β nın bir kestirimidir. Popülasyona ait β için bir güven aralığı kestiriminde bulunmak istediğimiz zaman, $(n-2)$ serbestlik dereceli t-dağılımını kullanırız. Örneğe ait regresyon doğrusunun eğimi b nin standart-hatası şöyle hesaplanır,

$$s_b = \frac{s_e}{s_x\sqrt{n-1}}$$

ki burada s_e kalıntıların standart sapması ve s_x X değişkeninin standart sapmasıdır. Çoğu zaman, regresyon denklemini hesaplatmakta kullandığımız bir bir kompüter yazılımı, b nin değerinin yanısıra s_b nin değerini de vermektedir.

Bir güven aralığını, β için tesis ederken, şu varsayımların geçerli olması gerekir: (1) örneğin rastgele seçilmiş olması, (2) saçılım grafiğinin yaklaşık doğrusal olması, (3) kalıntı grafiğinin belirgin bir biçim arzetmemesi ve (4) kalıntıların dağılımının yaklaşık normal olması.

Misal 21 Bir lisede öğrenim gören öğrencilerin 10 kişilik bir rastgele örneği, haftalık televizyon izleme süreleri (saat) ve not ortalamaları (4 üzerinden) için şu sonuçları verdi:

TV süresi (X):	12	21	8	20	16	16	24	0	11	18
Not ortalaması (Y):	3.1	2.3	3.5	2.5	3.0	2.6	2.1	3.8	2.9	2.6

Veriden şu istatistikler hesaplanmıştır: $\bar{x} = 14.6$, $s_x = 7.074$; $\bar{y} = 2.84$, $s_y = 0.53$; $r = 0.962$.

En küçük kareler doğrusunu belirle ve hakiki eğim için %95 lik bir güven aralığı kestirimi yap.

Cevap:

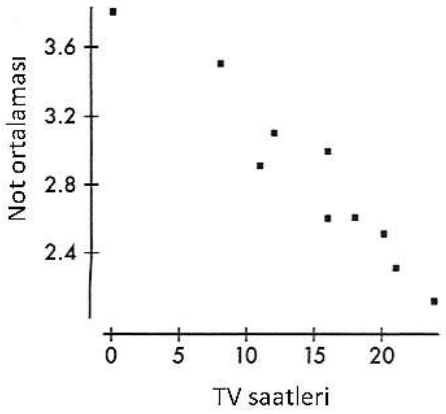
$$b = r\frac{s_y}{s_x} = 0.962\frac{0.53}{7.074} = -0.0721$$

$$a = \bar{y} - b\bar{x} = 2.84 - (-0.0721)14.6 = 3.893$$

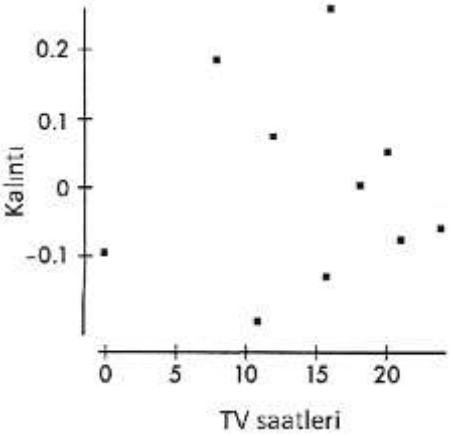
$$\hat{y} = 3.893 - 0.07202x$$

x_i:	12	21	8	20	16	16	24	0	11	18
y_i:	3.1	2.3	3.5	2.5	3.0	2.6	2.1	3.8	2.9	2.6
$\hat{y}_i$:	3.027	2.379	3.315	2.451	2.739	2.739	2.163	3.892	3.099	2.595
e_i:	0.073	-0.079	0.185	0.049	0.261	-0.139	-0.063	-0.092	-0.199	0.005

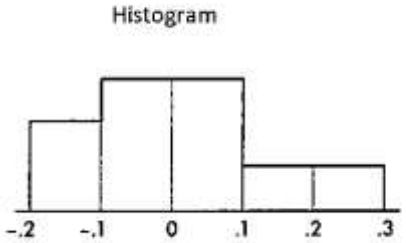
Öncelikle, kullanacağımız istatistiksel prosedürün varsayımlarını kontrol etmeliyiz: (1) Verinin öğrencilerin bir rastgele örneğinden elde edildiği bize söyleniyor. (2) Saçılım grafiği yaklaşık doğrusal.

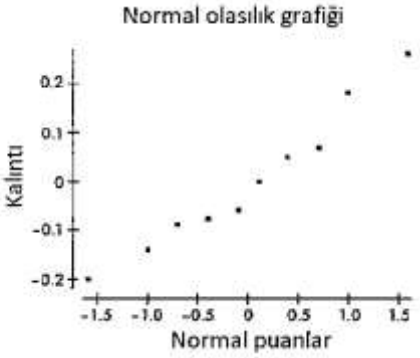


(3) Kalıntı grafiği belirgin bir biçim göstermiyor.



(4) Kalıntıların dağılımı yaklaşık normal.





Şimdi, kalıntıların standart sapmasını hesapladıktan sonra, en küçük kareler doğrusunun eğimi *b* nin standart hatasını hesaplayabiliriz,

$$s_e = \sqrt{\frac{\sum e_i^2}{n-2}} = 0.1531$$

$$s_b = \frac{s_e}{s_x\sqrt{n-1}} = \frac{0.1531}{7.074\sqrt{10-1}} = 0.0072$$

İki-değişkeli veri, örnekte on adet veri noktasına tekabül ettiği için, serbestlik derecesi,

$$sd = 10 - 2 = 8$$

ve kritik *t*-değerleri, her kuyrukta 0.025 olmak üzere, ± 2.306 dır. Gerçek eğim için, %95 lik güven

aralığı,

$$b \pm t(s_b) = -0.0721 \pm 2.306(0.0072) = -0.0721 \pm 0.166$$

Böylece, %95 lik güvenle şunu söyleyebiliyoruz ki televizyon karşısında geçirilen ilave her bir saat, not ortalamasında 0.055 ve 0.089 puan arasında bir düşüşü çağrıştırmaktadır. Ancak, şimdi ve her zaman, şunu hiç aklından çıkarmamalısın ki böyle durumlarda neden-ve-etkiye dair neticelere varırken dikkatli davranman gerekir. Belki de, düşük not ortalamasına sahip öğrenciler, bu ortalamaları, haftada kaç saat televizyon izlerlerse izlesinler, yine de alacaklardı. Yani burada etkili neden belki de öğrencilerin mesela zeka düzeyidir. Eğer bir deney dizaynı mevcut değilse, bir gözlemsel öğrenme-çalışmasıyla elde edilmiş veri bunu sana söyleyebilmez.

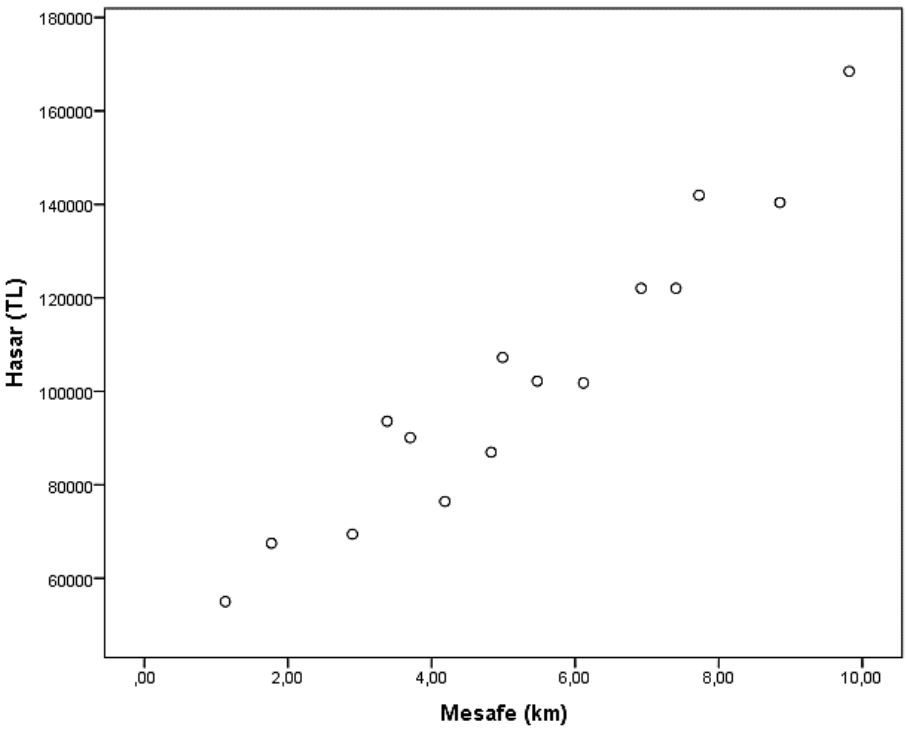
Misal 22 Bir sigorta şirketi, büyük bir şehirde son bir yılda gerçekleşen konut yangını vakaları arasından 15 ini rastgele seçerek, yangının neden olduğu hasar tutarı (TL) ve yangına maruz kalan konutun en yakın itfaiye istasyonuna mesafesini (km) belirlemiştir. Veri üzerinde bir istatistiksel yazılım paketi ile icra edilen doğrusal regresyon analizi sonucunda, kompüterin ürettiği çıktının bir kısmı şudur:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.961	0.923	0.918	9033.750

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	40083.921	5539.083		7.237	0.000
	Mesafe	11921.278	951.767	0.961	12.525	0.000



Regresyon denklemini yazabilir misin?

Cevap: Regresyon analizine dair tüm varsayımların yerine getirildiğini kabul ediyoruz; örneğin basit rastgele olduğu söylendi ve saçılım grafiğinin yaklaşık doğrusallığını görebiliyoruz. İstatistiksel paket yazılımlarının çoğunda olduğu gibi, bu çıktıda da "Coefficients" başlığı altında, "B" sütununda, uydurulan en küçük kareler doğrusunun *y*-kesimi ve eğimi veriliyor.

$$\widehat{\text{Hasar}} = 40083.921 + 11921.278(\text{Mesafe})$$

Regresyon doğrusunun eğimi için %95 lik bir güven aralığı kestirimi nedir?

Cevap: Eğimin standart hatası 951.767. On beş veri noktası olduğu için, $sd = 15 - 2 = 13$ ve kritik *t*-değerleri ± 2.16. Böylece, hakiki eğimin %95 lik güven aralığı,

$$b \pm t(s_b) = 11921.278 \pm 2.16(951.767) = 11921.278 \pm 2055.817$$

Yüzde 95 güvenimiz var ki konutun en yakın itfaiye istasyonuna olan uzaklığı 1 km arttığında, yangında meydana gelen hasarın tutarındaki ortalama artış 9865.5 TL ve 13977.1 TL arasındadır.

Mini Proje 13

Bir bilimsel dergi makalesi bul ki en az bir tane %95 lik güven aralığı kestirimini rapor ediyor olsun. Öğrenme-çalışmasının niçin yapılmış olduğunu, yani maksadını, kısaca açıkla. Makalede rapor edildiği üzere %95 lik güven aralıklarını yaz. Sonuçların interpretasyonunu yap. Yani, istatistiksel sonuçları, öğrenme-çalışması problemi bağlamında, herkesin anlayabileceği bir dilde yeniden ifade et. Makalenin niyet edilen maksada ulaştığını düşünüyor musun? Açıkla. Sonuçları, veri toplama prosesi ile alakalı potansiyel problemler bakımından irdele.

BÖLÜM 14 | ANLAMLILIK TESTLERİ: ARİTMETİK-ORTALAMA VE ORANTI

Giriş

Bir popülasyonun aritmetik ortalamasını veya orantısını kestirme problemi ile yakından ilişkili bir problem de bir popülasyonun aritmetik ortalaması veya orantısı hakkındaki bir hipotezi test etmedir. Mesela, bir biyoloji bilimcisi bir göldeki sazanların boylarının aritmetik ortalaması için bir aralık kestirimi belirleyebilir, ya da öteki taraftan, başka bir biyoloji bilimcisinin aynı göldeki sazanların boylarının aritmetik ortalaması hakkında ileri sürmüş olduğu iddiayı test edebilir. Bir tarım bilimcisi uyguladığı spesifik bir gübrenin neden olduğu ürün artışının aritmetik ortalaması için bir aralık kestirimi hesaplayabilir, ya da öteki taraftan, gübreyi geliştiren şirketin iddia etmiş olduğu aritmetik ortalama ürün artışını test edebilir. Rusya'daki bir seyahat acentası Antalya'daki güneşli günlerin orantısı için bir aralık kestiriminde bulunabilir, ya da öteki taraftan, turizm bürosunun güneşli günlerin orantısı hakkındaki iddiasını test edebilir. Bir iş sağlığı ve güvenliği uzmanı bir sektördeki ölümlü iş kazalarının orantısı için bir aralık kestirimi tespit edebilir, ya da öteki taraftan, ilgili işverenler örgütünün ölümlü iş kazalarının orantısı hakkındaki iddiasını test edebilir. Bu durumların her birinde, bilimsel araştırmacı popülasyon aritmetik ortalaması veya orantısının bir aralık kestirimi ile mi ilgilendiğine yoksa iddia edilen bir aritmetik ortalamanın veya orantının testi ile mi ilgilendiğine karar vermelidir.

Bir önceki bölümde, belli bir güvenle, bir popülasyonun aritmetik ortalamasının, orantısının ve regresyon doğrusunun eğiminin içinde yer aldığı, sırasıyla örnek aritmetik ortalaması, örnek orantısı ve örnek regresyon doğrusunun eğimi etrafındaki bir aralığın kestirimini yaptık. Bu ve bundan sonraki bölümlerde ise, mevzuyu, öteki tarafından bakarak ele alacağız. Yani, bir örnek istatistiğinin, iddia edilen bir popülasyon parametresi etrafında belli bir güvenle tesis edilecek bir aralık içinde yer alıp almadığını test edeceğiz.

Anlamlılık Testinin Mantığı

Bir istatistiksel test prosedürü, genel manada, şu aşamalardan geçer:

- *Sıfır hipotezi* adı verilen, test edilecek spesifik bir hipotezin seçilmesi,
- mütenasip bir rastgele örneğin alınması,
- örnekteki ölçümleri kullanarak, sıfır hipotezinin olabilirliğinin belirlenmesi.

Eğer örnek istatistiği, iddia edilen popülasyon parametresinden yeterince uzaksa, sıfır hipotezini reddetmek için yeterli kanıtın mevcut olduğunu söyleriz. Sıfır hipotezinin olasılıksızlığını göstermek suretiyle, kabul edilebilir olmadığını kanıtlamaya çalışırız.

Sıfır Hipotezi ve Alternatif Hipotez

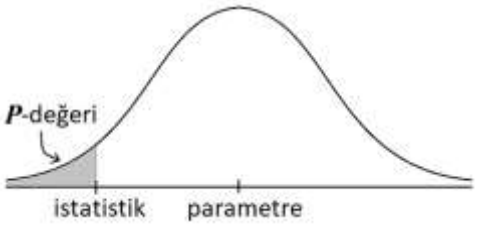
Örnek aritmetik ortalaması bağlamında düşünelim. Sıfır hipotezi H_0 ile gösterilir ve popülasyon aritmetik ortalaması hakkında bir eşitlik ifadesi formunda kurulur. Mesela, H_0: $\mu = 63.7$. Aynı zamanda, herhangi bir eşitsizlik formunda kurulmuş bir *alternatif hipotez* vardır ve H_1 ile gösterilir. Mesela, H_1: $\mu < 63.7$ veya H_1: $\mu > 63.7$ veya H_1: $\mu \neq 63.7$. Burada yürütülen mantığa göre, alternatif hipotezin vazifesi verinin temin ettiği kanıta dayanılarak reddedilecek bir sıfır hipotezinin yerini almaktır.

Dikkatini çekmek isterim ki H_0 ı ve H_1 i yalnızca μ ve π gibi popülasyon parametreleri üzerine kurabilirsin. Ne H_0 da ve ne de H_1 de $\bar{y}$ ve p gibi örnek istatistiklerini kullanamazsın.

Hiçbir zaman, hipotezlerini örnek veriden hesapladığın ölçülerin (yani örnek istatistiklerinin) üzerine kurma.

P-değeri

Sıfır hipotezinin doğru olduğu varsayımı altında, örnek istatistiğine ilişkin bir *P-değeri* hesaplanır. *P*-değeri, sıfır hipotezinin doğru varsayılması halinde örnekten hesaplanan istatistik kadar veya daha aşırı bir istatistik elde etme olasılığıdır. *P*-değeri ne kadar küçük ise, sıfır hipotezi ile örnek sonuçları (yani parametre için iddia edilen değer ve örnekte gözlenen değer) arasındaki fark o kadar anlamlıdır.



Demek ki, bir anlamlılık testinin sonucunda verilecek kararı *P*-değeri belirler. Örnek istatistiğine ilişkin *P*-değeri, H_0 hipotezi altında, eğer yüksek bir olasılık değerine sahip ise toplanan verinin sıfır hipotezinde ifade edilen model ile tutarlı olduğuna ve sıfır hipotezini reddetmek için bir sebep olmadığına hükmedilir. Fakat *bu, sıfır hipotezinin doğruluğunun ispatlandığı manasına gelmez*. Çünkü, toplanan veri için başka benzer hipotezlerle de aynı neticeye varılabilir. En fazla, sıfır hipotezinin yanlışlanamadığını söyleyebiliriz. Formal olarak, "sıfır hipotezi reddedilemedi" denir.

P-değeri yeterince küçük olduğu zaman, toplanan verinin sıfır hipotezinde ifade edilen model ile tutarlı olmayabileceğine ve sıfır hipotezinin reddedilmesine hükmedilir.

P-değeri, sıfır hipotezinin doğru olması halinde, toplanan verinin gözlenebilmesi olasılığıdır; yoksa, sıfır hipotezinin doğru olması olasılığı değildir.

I. Tip Hata ve II. Tip Hata

Bir rastgele örneğin sağladığı enformasyonu kullanarak, bir popülasyon parametresine ait bir sıfır hipotezini test ederken, iki tip hata yapmak olanaklıdır: (1) doğru olan bir sıfır hipotezini reddetmek ve (2) yanlış olan bir sıfır hipotezini reddedememek.

Eğer doğru bir sıfır hipotezi yanılgıyla reddedilir ise, işlenen hataya *I. tip hata* denir. Birinci tip hata işleme olasılığına, α-riski, aynı zamanda *anlamlılık düzeyi* denir. Eğer yanlış bir sıfır hipotezini reddetmekte yanılgıya düşülür ise, işlenen hataya *II. tip hata* denir. İkinci tip hata işleme olasılığı β ile gösterilir.

		Popülasyondaki hakikat	
		H_0 doğru	H_0 yanlış
Örneğe dayalı karar	H_0 reddedilir	I. tip hata (α-riski)	Doğru karar
	H_0 reddedilmez	Doğru karar	II. tip hata (β-riski)

Tek kuyruklu Test ve Çift Kuyruklu Test

Alternatif hipotez tek kuyruklu veya çift kuyruklu kurulabilir. Alternatif hipotez tek kuyruklu kurulduğu zaman, sıfır hipotezi değerinden yalnızca bir taraftaki sapmalara odaklanır. Tek kuyruklu bir alternatif hipotez ile yapılan anlamlılık testi için, *P*-değeri sıfır hipotezi değerinden yalnızca belirtilen taraftaki sapmaların olasılığıdır. Alternatif hipotez çift kuyruklu kurulduğu zaman, sıfır hipotezi değerinden her iki taraftaki sapmalarla da ilgilenir. Çift kuyruklu bir alternatif hipotez ile yapılan anlamlılık testi için, *P*-değeri sıfır hipotezi değerinden her iki taraftaki sapmaların olasılığıdır. Mesela, H_0: $\mu = 63.7$ ise, H_1: $\mu < 63.7$ ve H_1: $\mu > 63.7$ tek kuyruklu hipotezlerken, H_1: $\mu \neq 63.7$ çift kuyruklu bir hipotezdir.

Aynı veri kümesi için, çift kuyruklu P-değeri tek kuyruklu P-değerinin iki katıdır. Böyle olunca, tek kuyruklu bir test sıfır hipotezini daha kolay reddedecektir. Eğer hangisini kullanacağın konusunda bir kafa karışıklığı yaşarsan, bilmelisin ki çift kuyruklu test daima daha tutucu davranacaktır. Aynı zamanda, tek kuyruklu bir alternatif hipotez kurmak için de belli bir sebep olmalı.

Anlamlılık Düzeyi (α-riski)

Sıfır hipotezini reddetmek için, P-değeri ne kadar küçük olmalı? Bir yargıç bir sanık hakkında mahkumiyet kararı verirken, kararını "en az şüpheyle" kanıtlamak ister. Biz de H_0 hipotezini reddederken aynı şeyi isteriz. P-değeri, sıfır hipotezinin doğru varsayılması halinde, örnektekten hesaplanan istatistik kadar veya daha aşırı bir istatistik elde etme olasılığını ifade ettiğine göre, bu olasılık için bir eşik belirlemek gerekir. Eğer P-değerimiz bu eşiğin dışında kalırsa, sıfır hipotezini reddederiz. Belirlenen bu eşiğe α-riski veya anlamlılık düzeyi denir. Müştereken kullanılan anlamlılık düzeyleri (yani α-riskleri) 0.10, 0.05 ve 0.01 dir. Daha önce de söylendiği üzere, α-riski, aynı zamanda, doğru bir sıfır hipotezini yanılgıyla reddetme olasılığı, yani I. tip hatadır.

Anlamlılık düzeyinin seçimi, verinin geldiği öğrenme-çalışması bağlamında değerlendirilir. Mesela insan sağlığı ile doğrudan ilgili ilaç denemelerinde 0.01 düzeyini seçmek ya da tüketici tercihlerinin farklılığını test ederken 0.10 düzeyini seçmek gibi. Genellikle seçilen düzey ise 0.05 tir. Fakat şunu önemle belirtmek gerekir ki anlamlılık düzeyi veri elde edilmeden önce seçilmelidir. Aksi taktirde, α-riskini veriye uydurarak hile yapmakla suçlanabilirsiniz.

Eğer bir verinin sağladığı P-değeri seçilen α-riskinden daha küçük ise sıfır hipotezini reddederiz ve testin "söz konusu düzeyde anlamlı" olduğunu söyleriz. Mesela, "sıfır hipotezi %5 anlamlılık düzeyinde reddedildi" gibi.

Genel olarak, eğer $P > 0.10$ ise, H_0 ı reddetmek için *yeterli kanıt olmadığı* söylenir; eğer $P < 0.10$ veya $P < 0.05$ ise, bu anlamlılık düzeylerinde (veya α-riskleri ile) H_0 ı reddetmek için *yeterli kanıt olduğu* söylenir; eğer $P < 0.01$ ise, H_0 ı reddetmek için *güçlü kanıt olduğu* söylenir.

Bir anlamlılık testi prosedürünü yürütmeden önce, daima, varsayımlarını kontrol etmelisin.

Aritmetik ortalama için Anlamlılık Testi

Bu kısımda, yaklaşık normal dağılımlı bir popülasyondan alınmış basit rastgele örneğe ait ölçümlere dayanarak, popülasyonun aritmetik ortalaması hakkında kurulan hipotezi test edeceğiz. Anlamlılık testine girişmeden önce, elimizdeki örnek verinin normallik varsayımını sağlayıp sağlamadığını kontrol etmemiz gerekir.

Misal 1 Bir otomobil şirketi SUV modelinin 100 kilometrede 6.5 litre yakıt tükettiğini iddia etmektedir. Bir çevre ajansı bu rakamın daha yüksek olduğuna inanmaktadır ve 50 hacimli bir örnek üzerinde bir teste girişirler. Örnek aritmetik ortalamasını 7.0 litre/100km ve standart sapmasını 1.4 litre/100km olarak hesapladıklarına göre, şirketin iddiası %5 lik bir anlamlılık düzeyinde reddedilir mi? Ya %1?

Cevap: Örnek hacmi 40 ı geçtiği için normallik varsayımını çok önemsemeyerek, bir basit rastgele örneğe sahip olduğumuzu varsayıyoruz. Sıfır hipotezimiz, alternatif hipotezimiz ve anlamlılık düzeyimiz:

$$H_0: \mu = 6.5 \text{ litre/100km},$$
$$H_1: \mu > 6.5 \text{ litre/100km},$$
$$\alpha = 0.05.$$

Sıfır hipotezinin doğruluğunu, elimizdeki örneğin sağlamış olduğu enformasyonu kullanarak test etmek için, önce örnek aritmetik ortalamasının standart hatasını (yani, örnekleme dağılımının standart sapmasının bir kestirimini) hesaplayalım,

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{1.4}{\sqrt{50}} = 0.198$$

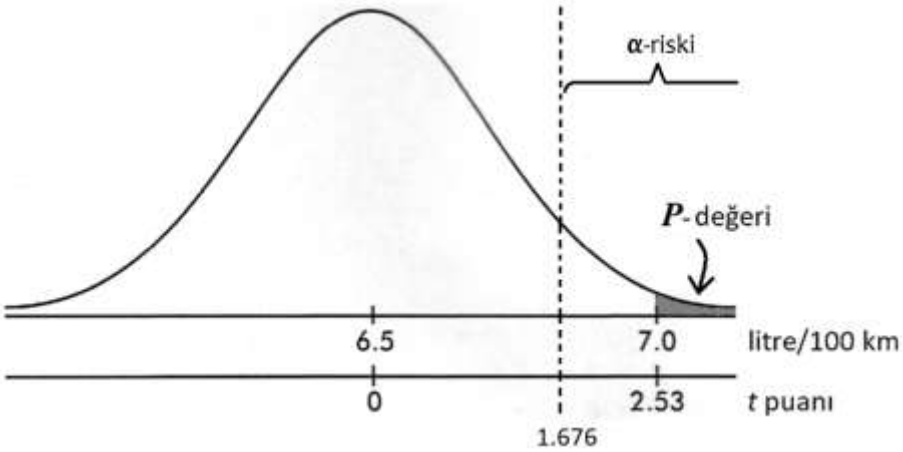
Sıfır hipotezini test etmek demek, örnek aritmetik ortalamasının H_0 da iddia edilen dağılımın bir elemanı olup olmadığına karar vermek demektir. Örnek aritmetik ortalamasının iddia edilen aritmetik ortalamaya uzaklığını ölçmek için P-değerini hesaplamamız gerekir. Ancak, popülasyon standart sapması yerine örnek standart sapmasını kullandığımızdan dolayı örnek aritmetik ortalamalarının dağılımını modellemek için normal dağılım yerine t-dağılımını kullanmak durumundayız. Böyle olunca, Tablo C deki Student t-dağılımı tablomuzun kısıtlılığından dolayı, örnek aritmetik ortalamamıza tekabül eden t-puanına ait P-değerini doğrudan okuyamayacağımız için, onun yerine, α-riskini (yani anlamlılık düzeyini) sınırlayan kritik t-puanını buluruz. Eğer örnek aritmetik ortalamasına tekabül eden t-puanı, anlamlılık düzeyini sınırlayan kritik t-puanından daha aşırı bir puan olursa, H_0 hipotezini reddederiz. Çünkü bu durumda örnek istatistiğimize ilişkin P-değeri seçmiş olduğumuz α-riskinin altında kalmış olur. Örnek-aritmetik ortalaması 7.0 litre/100km olduğuna göre, buna tekabül eden t-puanını,

$$\frac{7.0 - 6.5}{0.198} = 2.53$$

olarak hesaplarız. Alternatif hipotezimiz tek kuyruklu kurulmuş olduğu için, Tablo C de, %5 lik α-riskini sınırlayan kritik t-puanı 1.676 dır. Sonuç itibarıyla,

$$2.53 > 1.676$$

ifadesi bize örnek aritmetik ortalamasına tekabül eden t-puanının, anlamlılık düzeyini sınırlayan kritik t-puanından daha aşırı bir puan olduğunu göstermektedir, yani $P(t > 2.53) < 0.05$. Çevre ajansı, otomobil şirketinin iddiasını %5 anlamlılık düzeyinde reddedecektir.



Tablo C de, %1 lik α-riskini sınırlayan kritik t-puanı 2.403 tür. Buna göre,

$$2.53 > 2.403$$

ifadesi bize yine örnek aritmetik ortalamasına tekabül eden t-puanının, anlamlılık düzeyini sınırlayan kritik t-puanından daha aşırı bir puan olduğunu göstermektedir, yani $P(t > 2.53) < 0.01$. Çevre ajansı, otomobil şirketinin iddiasını %1 anlamlılık düzeyinde de reddedecektir. Bu durumda, $P < 0.01$ olduğuna göre, H_0 ı reddetmek için *güçlü* kanıt olduğunu söyleriz.

Anlamlılık testi sonucunda vardığın neticeyi, daima, öğrenme-çalışmasının kendi bağlamında ifade et.

Misal 2 Sigara üreticileri tarafından yayımlanan bir raporda, halihazırda pazara sunulan markalarda sigara başına katran miktarının 5 miligram olduğu iddia edilmiştir. Bir gazeteci bu iddiayı kontrol etmek maksadıyla pazardaki çeşitli markaların bir kesiti olan 15 sigarayı bir laboratuvara göndermiştir.

a. Örnekteki sigaraların katran miktarına ait aritmetik ortalama 5.63 miligram ve standart sapma 1.61

miligram olarak tespit edildiğine göre, hangi neticeye varılır? Anlamlılık düzeyini %10 varsayalım.

Cevap: Yaklaşık normal dağılımlı bir popülasyondan bir basit rastgele örnek alındığını varsayıyoruz,

$$H_0: \mu = 5 \text{ miligram},$$
$$H_1: \mu > 5 \text{ miligram},$$
$$\alpha = 0.10$$

ve

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{1.61}{\sqrt{15}} = 0.42.$$

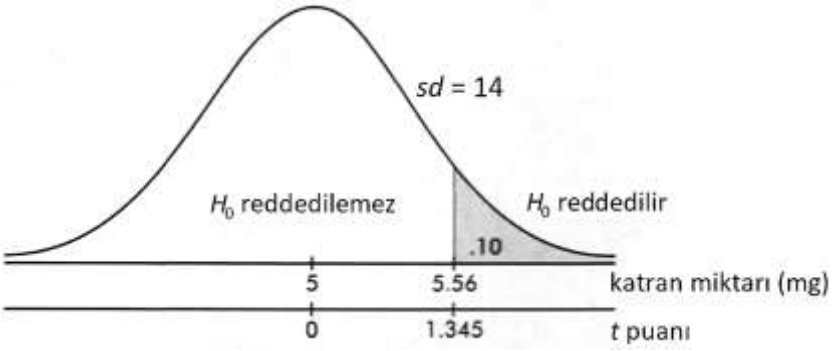
Alternatif hipotez tek kuyruklu, $sd = 15 - 1 = 14$ ve $\alpha = 0.10$ olduğuna göre kritik *t*-puanını Tablo C de 1.345 olarak okuruz. Buna göre, sigara başına düşen katran miktarı için kritik değer,

$$5 + 1.345(0.42) = 5.56 \text{ miligram}$$

olarak hesaplandıktan sonra,

$$5.63 > 5.56$$

olduğu için, yayımlanan raporda yer verilen iddia %10 anlamlılık düzeyinde reddedilir.



b. Anlamlılık düzeyi %5 olsaydı hangi neticeye varılırdı?

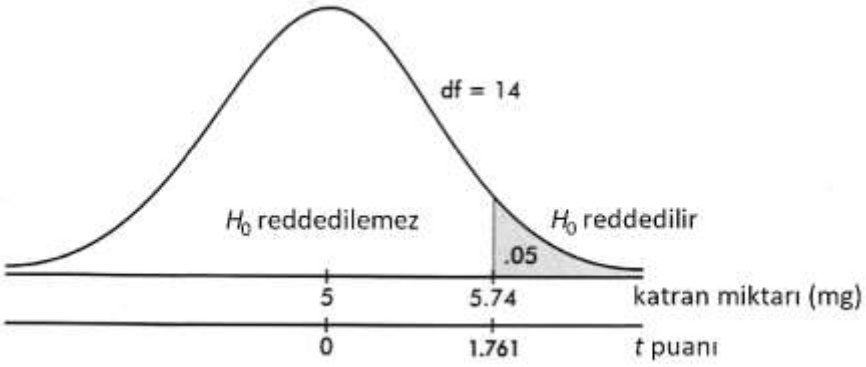
Cevap: Bu halde, kritik *t*-puanı Tablo C de 1.671 olarak okunacak ve buna göre sigara başına düşen katran miktarı için kritik değer,

$$5 + 1.671(0.42) = 5.74 \text{ miligram}$$

olarak hesaplandıktan sonra,

$$5.63 < 5.74$$

olduğu için, yayımlanan raporda yer verilen iddia %5 anlamlılık düzeyinde reddedilemeyecektir.



c. *P*-değeri nedir?

Cevap: Student *t*-dağılımına ait olasılık fonksiyonu ile integral işlemleri yapmadan veya kompüter yazılımı kullanmadan *P*-değerini tam olarak hesaplayamayız. Bununla beraber, örnek aritmetik ortalamasına tekabül eden *t*-puanı,

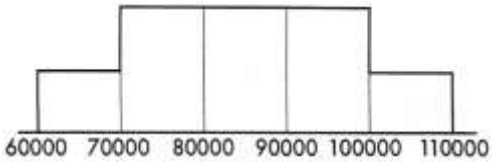
$$\frac{5.63 - 5}{0.42} = 1.5$$

olduğu için, 1.345 ve 1.671 arasında yer alacak ve böylece *P*-değerinin de 0.05 ve 0.10 arasında olduğu neticesine varılacaktır.

Bir sıfır hipotezini hiç bir zaman "kabul etmeyiz"; onu reddetmek için ya kanıtımız olur ya da olmaz.

Misal 3 Ticaret odası tarafından, şehirdeki evler için aritmetik ortalama satış fiyatı, 90000 TL olarak açıklanmıştır. Bir emlakçı ise, aracılık ettiği son satışları şöyle kaydetmiştir: 75000, 102000, 82000, 87000, 77000, 93000, 98000, 68000. Bu örneğin aritmetik-ortalamsı 85250 TL ve standart sapması 11877 TL olduğuna göre, ticaret odasının iddiasını reddetmek için yeterli kanıt varmı?

Cevap: Basit rastgele bir örneğe sahip olduğumuzu varsayıyoruz. Normalliği kontrol etmek için bir histogram çizebiliriz.



Çizdiğimiz histogram kabaca tek-tepeli ve yine kabaca simetrik bir frekans dağılımını göstermektedir. Sıfır hipotezimiz ve alternatif hipotezimiz:

$$H_0: \mu = 90000 \text{ TL},$$
$$H_1: \mu \neq 90000 \text{ TL}.$$

Örnek aritmetik ortalamasının standart hatası:

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{11877}{\sqrt{8}} = 4199.$$

Örnek aritmetik ortalamasına tekabül eden *t*-puanı:

$$\frac{85250 - 90000}{4199} = -1.13.$$

Tablo C de, $sd = 8 - 1 = 7$ satırında 1.13 lük *t*-puanı 1.119 ve 1.415 lik *t*-puanları arasına düşmektedir ve bu *t*-puanları sırasıyla 0.15 ve 0.10 luk *P*-değerlerine karşılık gelmektedir. Alternatif hipotezimiz çift kuyruklu olduğu için, bu *P*-değerlerinin iki katını almalıyız: sırasıyla 0.30 ve 0.20. Yani anlamlılık testimize ait *P*-değeri 0.20 ile 0.30 arasındadır. Bu büyüklükte bir *P*-değeri ile, ticaret odasının iddiasını reddetmek için yeterli kanıt yoktur.

İki Aritmetik ortalama Arasındaki Fark için Anlamlılık Testi

Bu kısımda, bir nitel değişken ve bir nicel değişken arasındaki ilişkiyle (yani, eğer varsa) alakadar olacağız. Dikkatini çekmek istiyorum ki söz konusu ilişkiyi yalnızca alacağımız örnekte değil, tüm popülasyonda arıyoruz. Bir önceki kısımda, *tek* bir nicel değişkenin aritmetik ortalaması için anlamlılık testini öğrendik. İki değişken arasındaki ilişkiyle ilgilendiğimiz zaman, orjinal popülasyonda bir ilişkinin mevcudiyetine dair istatistiksel kanıt olup olmadığına karar vermek için yine en iyi araç bir anlamlılık testidir.

Her ne zaman iki değişken arasında bir ilişkinin keşfine çıkılsa, sıfır hipotezi o iki değişken arasında ilişki olmadığını ifade ederken, alternatif hipotez ilişki olduğunu ifade eder. İster nicel isterse nitel değişkenler sözkonusu olsun, değişkenlerin tipine bakılmaksızın, bu hipotezler böylece kurulur. Ancak, sıfır hipotezinin ve alternatif hipotezin, *değişkenler* arasında ilişki olup olmadığı cinsinden formülasyonlarının yanısıra, uygulamada bu hipotezler daha ziyade münasebetli *parametreler* cinsinden formüle edilirler. İşte, değişkenlerin tiplerine bağlı olarak, burada büyük bir farklılık ortaya çıkar.

Bir nitel ve bir nicel değişken arasındaki ilişkiyi öğrenmeye-çalışırken, münasebetli parametreler bağımsız örnekler için popülasyon aritmetik ortalamaları (μ_1, μ_2) ve eşli veri için farkların

popülasyon aritmetik ortalaması (μ_d) olacaktır.

Misal 4 Araştırmacılar 1 yaşındaki ve 3 yaşındaki çocukların günlük TV seyretme sürelerine ait veri topladılar.

Değişkenler cinsinden bir sıfır hipotezini nasıl formüle edersin?

Cevap: "Yaş (1 yaş ve 3 yaş) ve TV seyretme süresi arasında bir ilişki yoktur."

Parametreler cinsinden bir sıfır hipotezini nasıl formüle edersin?

Cevap: "1 yaşındaki tüm çocukların popülasyonuna ait aritmetik ortalama TV seyretme süresi, μ_1, 3 yaşındaki tüm çocukların popülasyonuna ait aritmetik ortalama TV seyretme süresi, μ_2, ile aynıdır."

Demek ki, bir nitel değişken ve bir nicel değişken arasında ilişki olup olmadığı sorusunu, nitel değişkenin kategorilerilerine göre tanımlanan popülasyonlar bakımından nicel değişkenin aritmetik ortalamalarının aynı olup olmadığı şeklinde ifade edeceğiz. Yani, başka bir deyişle, sıfır hipotezimiz nitel değişkenin kategorilerine göre oluşan popülasyonların özdeşliğini iddia edecek.

Bağımsız Örnekleme

Elimizde iki bağımsız basit rastgele örnek olduğunu ve bunların her birinin yaklaşık normal dağılımlı bir popülasyondan geldiğini varsayıyoruz. Eğer örnek hacimleri yeterince geniş ise normallik varsayımı gevşetilebilir.

Bu halde, sıfır hipotezi, popülasyonların aritmetik ortalamalarının aynı olduğunu iddia eder, yani popülasyon aritmetik ortalamaları arasındaki farkın sıfır olduğunu:

$$H_0: \mu_1 - \mu_2 = 0$$

Alternatif hipotez tek kuyruklu veya çift kuyruklu olarak kurulabilir:

$$H_1: \mu_1 - \mu_2 < 0,$$

$$H_1: \mu_1 - \mu_2 > 0$$

veya

$$H_1: \mu_1 - \mu_2 \neq 0$$

Misal 5 Bir kompüter şirketi ürettiği yeni model sunuculardа haftalık atıl zaman aritmetik ortalamasını düşürmeyi hedeflemektedir. Eski modeli kullanan 40 firmanın bir basit rastgele örneğinde haftalık atıl zaman aritmetik ortalamasını 125 dakika ve standart sapmasını 37 dakika olarak belirler. Yeni modeli kullanmaya başlayan 35 firmada ise haftalık atıl zaman aritmetik-ortlaması 115 dakika ve standart sapması 43 dakika olarak ölçülür. Kompüter şirketi, %10 anlamlılık düzeyinde hangi neticeye varır?

Cevap: İki örnek biribirinden bağımsız ve birer basit rastgele örnek. Hacimleri yeterince geniş ve bu sebeple normallik varsayımını her ikisi için de gevşetebiliriz.

$$H_0: \mu_2 - \mu_1 = 0$$

$$H_1: \mu_2 - \mu_1 < 0$$

$$\alpha = 0.10$$

$$sd = (n_1 - 1) + (n_2 - 1) = (40 - 1) + (35 - 1) = 73$$

$$s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{43^2}{35} + \frac{37^2}{40}} = 9.93$$

Örnek aritmetik ortalamaları arasındaki fark,

$$115 - 125 = -10$$

ve bu farka tekabül eden *t*-puanı,

$$\frac{-10 - 0}{9.93} = -1.07.$$

Tablo C yi kullanmak suretiyle, bu *t*-puanının 0.10 ve 0.15 arasında bir olasılığı verdiğini buluruz. Buna göre, gözlenen fark %10 anlamlılık düzeyinde anlamlı değildir. Kompüter şirketi, yeni model sunucularda haftalık atıl zaman aritmetik ortalamasının düştüğünü gösteren yeterli kanıta sahip değildir. Böyle olmakla beraber, gözlenen fark, mesela %15 düzeyinde anlamlı bulunabilir. Bu halde yeni model sunucunun eski modele göre daha düşük haftalık atıl zaman aritmetik ortalamasına sahip olduğunu gösteren kanıtın zayıf olduğunu söyleriz.

Misal 6 Bir yumurta üreticisi kullandığı iki ayrı koli tipi arasında yumurtaların sağlam nakledilmesi bakımından anlamlı bir fark olup olmadığını belirlemek istiyor. Birinci koli tipi ile nakledilen 200 kolilik bir basit rastgele örnekte, koli başına kırık yumurtaların aritmetik ortalaması 0.7 ve standart sapması 0.31 olarak hesaplanıyor. İkinci koli tipi ile nakledilen 300 kolilik başka bir basit rastgele örnekte ise, koli başına kırık yumurtaların aritmetik ortalaması 0.775 ve standart sapması 0.42 olarak hesaplanıyor. Aritmetik ortalamalar arasındaki fark, %5 anlamlılık düzeyinde anlamlı mıdır?

Cevap: Elimizde biribirinden bağımsız basit rastgele örnekler var ve örnek hacimleri normallik varsayımını gevşetmemize yetecek kadar geniş. Böylece,

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

$$\alpha = 0.05$$

$$sd = (200 - 1) + (300 - 1) = 498$$

$$s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{0.31^2}{200} + \frac{0.42^2}{300}} = 0.0327$$

Koli tiplerinden birinin yumurtaları diğerinden daha iyi muahafaza ettiğine dair bir inanç belirtilmediği için, alternatif hipotezimizi çift kuyruklu kurduk. Dolayısıyla bu çift kuyruklu bir testtir. Aritmetik ortalamalar arasında gözlenen fark,

$$0.7 - 0.775 = -0.075$$

ve bu farka tekabül eden *t*-puanı:

$$\frac{-0.075 - 0}{0.0327} = -2.29.$$

Tablo C yi kullanmak suretiyle, bu *t*-puanının 0.01 ve 0.02 arasında bir olasılığı verdiğini buluruz. Test çift kuyruklu olduğu için, *P*-değeri 0.02 ve 0.04 arasındadır. Böylece, koli başına kırık yumurtaların aritmetik ortalamaları arasında gözlenen fark %5 düzeyinde istatistiksel olarak anlamlıdır fakat %1 düzeyinde anlamlı değildir. Eğer $0.01 < P < 0.05$ ise, genellikle H_0 ı reddetmek için güçlü kanıt olduğunu (veya aynı manada, aritmetik ortalamalar arasındaki farkın anlamlı olduğunu) söyleriz.

Misal 7 Büyük bir bankanın sendika temsilcisi kadın çalışanların terfi almak için erkek çalışanlara göre daha fazla süre beklemek zorunda kaldıklarını iddia etmektedir. Banka yönetimi ise, terfi için bekleme süreleri arasında çalışanların cinsiyeti bakımından bir fark olmadığını iddia etmektedir. Eğer bu bankanın çalışanı beş erkeğin terfi için bekleme süreleri 8, 7, 10, 5 ve 7 yıl ve dört kadının bekleme süreleri 9, 5, 12 ve 8 yıl ise, hangi neticeye varılır?

Cevap: İki basit rastgele örneğe sahip olduğumuzu varsayıyoruz. Örnekler biribirinden bağımsız ve yaklaşık normal dağılımlı popülasyonlardan çekildiklerini varsayıyoruz. (Örnek hacimleri çok dar

olduğu için histogramlarına bakarak dağılımları hakkında bir sonuca varamayız elbette, fakat geldikleri popülasyonun dağılımı hakkında halihazırda sahip olduğumuz enformasyonu kullanarak dağılımın normale yaklaştığına karar veriyoruz.)

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

$$\bar{y}_1 = 7.4 \text{ yıl}, s_1 = 1.82 \text{ yıl}; \quad \bar{y}_2 = 8.5 \text{ yıl}, s_2 = 2.89 \text{ yıl}$$

$$\bar{y}_1 - \bar{y}_2 = 7.4 - 8.5 = -1.1$$

İki örnek aritmetik ortalaması arasındaki farkın standart hatasını şöyle hesaplarız:

$$s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{1.82^2}{5} + \frac{2.89^2}{4}} = 1.66.$$

Gözlenen farka tekabül eden *t*-puanı,

$$\frac{-1.1 - 0}{1.66} = -0.66.$$

Tablo C de, $sd = 5 + 4 - 2 = 7$ ile görüyoruz ki $0.66 < 0.711$. Bu 0.711 lik *t*-puanı 0.25 lik bir kuyruk olasılığını vermektedir. Buna göre, *P*-değeri 0.25 ten daha büyüktür ve banka yönetiminin iddiasını çürütecek kanıt olmadığı neticesine varırız.

Eşli Veri

Yukarıda betimlenen analiz ve prosedür biribiriyle karşılaştırılacak iki örneğin biribirinden bağımsız olmasını gerektirir. Ancak, bazı deneylerde ve testlerde doğal olarak çiftler halinde veri üreten popülasyonları karşılaştırmak gerekir. Böylesi haller için münasip olan prosedür, eşli verinin farklarından meydana gelen değişkeni tek örnek testine tabii tutmaktır. Bu teste eşli *t*-testi adı verilir ve farklar kümesi üzerinde tek örnek anlamlılık testi icra edilir.

Misal 8 Piyasaya yeni sürülen bir ilacın, kullananların reaksiyon süresini uzattığına dair iddiayı test etmek için 30 kişilik bir basit rastgele örnek alınır. Örnekteki kişiler ilacı almadan önce ve aldıktan sonra teste tabii tutularak reaksiyon süreleri (saniye) ölçülür. Şu veri elde edilir:

Kişi	Önce	Sonra	Fark (*d*)
1	1.42	1.48	-0.06
2	1.87	1.75	0.12
3	1.34	1.31	0.03
4	0.98	1.22	-0.24
5	1.51	1.58	-0.07
6	1.43	1.57	-0.14
7	1.52	1.48	0.04
8	1.61	1.55	0.06
9	1.37	1.54	-0.17
10	1.49	1.37	0.12
11	0.95	1.07	-0.12
12	1.32	1.35	-0.03
13	1.68	1.77	-0.09
14	1.44	1.44	0.00
15	1.17	1.27	-0.10

Kişi	Önce	Sonra	Fark (*d*)
16	1.83	1.75	0.08
17	1.40	1.51	-0.11
18	1.75	1.67	0.08
19	1.56	1.72	-0.16
20	1.56	1.63	-0.07
21	2.03	1.81	0.22
22	1.38	1.48	-0.10
23	1.42	1.35	0.07
24	1.69	1.75	-0.06
25	1.50	1.39	0.11
26	1.12	1.24	-0.12
27	1.38	1.40	-0.02
28	1.71	1.65	0.06
29	0.91	1.11	-0.20
30	1.59	1.85	-0.26

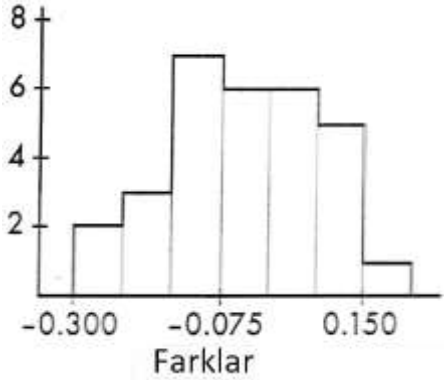
Veri çiftleri biribiriyle açıkça ilişkili olduğu için, bağımsız örnekler için icra ettiğimiz iki örnek testini

burada icra edemeyiz. Veri tablosunda her gözlem ünitesinin önceki ve sonraki ölçümleri arasındaki farklar hesaplanarak yazılmıştır. Bu veri kümesi için, münasip olan prosedür, farklar üzerinde yapılacak tek örnek anlamlılık testidir.

$$H_0: \mu_d = 0$$

$$H_1: \mu_d < 0$$

Veri eşlidir, çünkü ölçümler aynı bireyler üzerinde ilaç verilmeden önce ve ilaç verildikten sonra alınmıştır. Herhangi bir bireyin reaksiyon süresi diğer bireylerin reaksiyon sürelerinden bağımsızdır ve böylece farklar bağımsızdır. Kişilerin bir basit rastgele örneği teste tabii tutulmuştur. Farkların aşağıda çizilmiş histogramına bakınca, yaklaşık normal olduğu kanaatine varıyoruz. Tek tepeli ve hemen hemen simetrik.



Farklara ait sayısal betimleyiciler şöyle hesaplanmıştır:

$$n = 30, \bar{y} = -0.037667 \text{ ve } s = 0.11755.$$

Örneklenen farkların aritmetik ortalamasının örnekleme dağılımına ait standart hata,

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{0.11755}{\sqrt{30}} = 0.021462.$$

Örneklenen farkların aritmetik ortalamasına tekabül eden *t*-puanı,

$$\frac{\bar{y} - 0}{s_{\bar{y}}} = \frac{-0.037667}{0.021462} = -1.755.$$

Serbestlik derecesi,

$$sd = n - 1 = 30 - 1 = 29$$

olduğuna göre, Tablo C de *P*-değerini şu olasılık aralığında buluruz,

$$0.05 > P(t < -1.755) > 0.025.$$

P-değeri, içinde bulunduğu aralık itibarıyla, ilaç alındıktan sonra reaksiyon sürelerinde gözlenen artışın (%5 anlamlılık düzeyinde) anlamlı olduğu neticesine varmamızı sağlayacak kadar küçüktür.

Veri çiftler halinde geldiyse, iki örnek *t*-testini kullanabilmezsin.

Orantı için Anlamlılık Testi

Bu kısımda, başarı ve başarısızlık gibi yalnızca iki olanaklı çıkımı bulunan değişkenler için yapılacak belirli sayıdaki gözlemler içerisinde arzu edilen çıkımın orantısı hakkında kurulan hipotezleri test edeceğiz. Anlamlılık testine girişmeden önce, elimizdeki örnek verinin bir basit rastgele örnek olduğunu, hem $n\pi$ nin hem de $n(1-\pi)$ nin 10 dan büyük olduğunu ve örnek hacminin popülasyonun %10 unu geçmediğini varsayıyoruz.

Misal 9 Çok büyük bir otomobil fabrikasında, işçi sendikası temsilcisi kendi sendikalarına mensup işçilerin %75 inin, eğer talepleri karşılanmaz ise, greve gitme fikrini desteklediğini iddia etmektedir.

Fabrika yetkilisi gerçek orantının daha düşük olduğuna inanmaktadır ve %10 anlamlılık düzeyinde bir anlamlılık testi yürütür. Eğer sendika mensubu işçilerin 125 inden oluşan bir basit rastgele örnekte 87 işçi greve gitme fikrini desteklediğini ifade ederse, netice ne olur?

Cevap: Bakıyoruz, $n\pi = 125 \times 0.75 = 93.75$ ve $n(1-\pi) = 125 \times 0.25 = 31.25$, her ikisi de 10 dan büyük. Bir basit rastgele örneğimiz var ve 125 kişilik örnek hacminin toplam sendikalı mensubu işçilerin %10 unu geçmediğini varsayıyoruz.

$$H_0: \pi = 0.75$$

$$H_1: \pi < 0.75$$

İddia edilen orantıyı örnek orantısının ait olduğu örnekleme dağılımının standart sapmasını hesaplamak için kullanıyoruz:

$$\sigma_p = \sqrt{\frac{0.75 \times 0.25}{125}} = 0.03873.$$

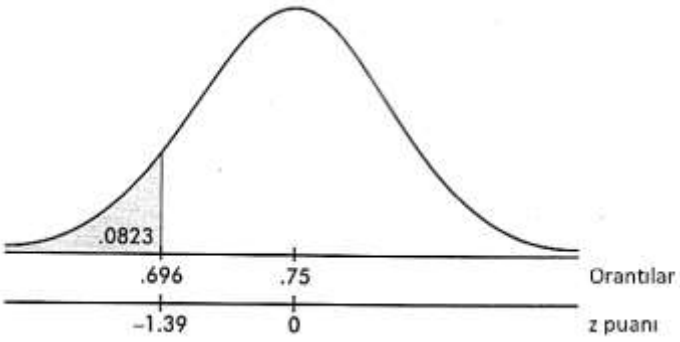
Gözlenen örnek orantısı,

$$p = \frac{87}{125} = 0.696.$$

Örnek orantısı ve iddia edilen popülasyon orantısı arasındaki uyuşmazlığın gücünü ölçmek için, P-değerini hesaplarız. Örnek orantısı için hesapladığımız 0.696 değeri ve ondan daha aşırı değerlerin olasılığını bulmak için, 0.696 ya tekabül eden z-puanını hesaplarız:

$$\frac{0.696 - 0.75}{0.03873} = -1.39.$$

Tablo B de, -1.39 luk z-puanının soluna düşen olasılığı 0.0823 olarak okuruz ve bu olasılık bizim P-değerimizdir. $0.0823 < 0.10$ olduğuna göre, H_0 ı %10 anlamlılık düzeyinde reddetmek için yeterli kanıt vardır. Fabrika yetkilisi sendika temsilcisine bu düzeyde meydan okuyabilir. Ancak, dikkatini şuna çekmek isterim ki H_0 ı %5 anlamlılık düzeyinde reddetmek için yeterli kanıt yoktur, hatta %8 de bile. Çünkü $0.0823 > 0.05$ ve $0.0823 > 0.08$.



Misal 10 Bir kanser bilimsel araştırma grubu 40 yaş üstü 500 kadını yoklayarak, bu yaş grubundaki kadınların %28 inin düzenli meme filmi çektirdiği hipotezini test etmek istiyor. Eğer kadınların 151 i olumlu yanıt verdiyse, hipotez %5 anlamlılık düzeyinde reddedilir mi?

Cevap: Bakıyoruz, $n\pi = 500(0.28) = 140$ ve $n(1-\pi) = 500(0.72) = 360$, her ikisi de 10 dan büyük. Bir basit rastgele örneğimiz olduğunu varsayıyoruz ve aşikar ki 500 lük örnek hacmi tüm 40 yaş üstü kadınların %10 unu geçmez. İddia edilen %28 lik orantının düşüklüğüne veya yüksekliğine dair herhangi bir şüphe seslendirilmediği için, orantı için çift kuyruklu bir z-testi yürüteceğiz:

$$H_0: \pi = 0.28$$

$$H_1: \pi \neq 0.28$$

İddia edilen orantıyı örnek orantısının ait olduğu örnekleme dağılımının standart sapmasını hesaplamak için kullanıyoruz:

$$\sigma_p = \sqrt{\frac{0.28(0.72)}{500}} = 0.0201.$$

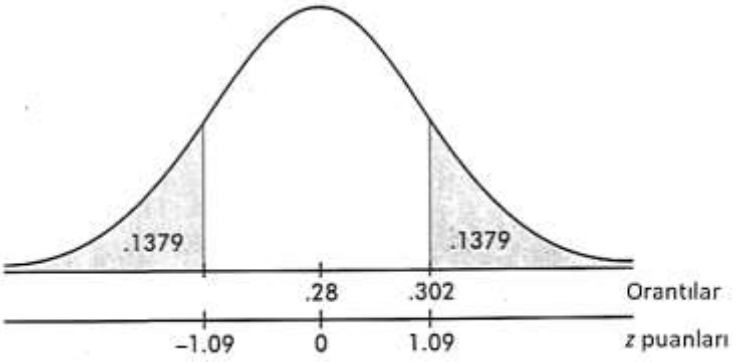
Gözlenen örnek orantısı,

$$p = \frac{151}{500} = 0.302.$$

Örnek orantısı ve iddia edilen popülasyon orantısı arasındaki uyuşmazlığın gücünü ölçmek için, P-değerini hesaplarız. Örnek orantısı için hesapladığımız 0.302 değeri ve ondan daha aşırı değerlerin olasılığını bulmak için, 0.302 ya tekabül eden z-puanını hesaplarız:

$$\frac{0.302 - 0.28}{0.0201} = 1.09.$$

Tablo B de, 1.09 luk z-puanının sağına düşen olasılığı $1 - 0.8621 = 0.1379$ olarak okuruz ve test çift kuyruklu olduğu için bizim P-değerimiz bu olasılığın iki katıdır. Yani, $2(0.1379) = 0.2758$ bizim P-değerimiz olur. $0.2758 > 0.05$ olduğuna göre, H_0 ı %5 anlamlılık düzeyinde reddetmek için yeterli kanıt yoktur. Yani, kanser bilimsel-araştırma grubu %28 lik iddianın yanlışlığını ifade edemez. Dikkat edersen, eğer α-riski mesela 0.25 kadar büyük olsa dahi H_0 reddedilemeyecektir.



İki Orantı Arasındaki Fark için Anlamlılık Testi

Bu kısımda, iki nitel değişken arasındaki ilişkiyle (yani, eğer varsa) alakadar olacağız. Yine dikkatini çekmek istiyorum ki söz konusu ilişkiyi yalnızca alacağımız örnekte değil, tüm popülasyonda arıyoruz. Daha önce (5. Bölüm'de) iki nitel değişken arasındaki ilişkiyi bir çapraz tablo vasıtasıyla nasıl betimleyebileceğimizi öğrenmiştik. Şimdi, amacımız, örneklenen bireyler için iki nitel değişken arasındaki ilişki hakkında bir ifadede bulunmanın ötesine geçerek, örnekten elde ettiğimiz enformasyonu kullanmak suretiyle, örneklenen popülasyon için o değişkenler arasındaki ilişki hakkında bir neticeye varmak.

Bir önceki kısımda, *tek* bir nitel değişkenin orantısı için anlamlılık testini öğrendik. İki nitel değişken arasındaki ilişkiyle ilgilendiğimiz zaman, örneklenen popülasyonda bir ilişkinin mevcudiyetine dair istatistiksel kanıt olup olmadığını belirlemek için yine anlamlılık testi yapacağız.

İki nitel değişken arasında herhangi bir ilişki olmadığını ifade etmek üzere, sıfır hipotezini, nitel değişkenlerden birinin kategorilerine göre tanımlanan popülasyonlar bakımından diğer nitel değişkenin orantıları (π_1, π_2) üzerine kuracağız. İki nitel değişkenin her biri için yalnızca iki olanaklı kategori olması halinde (yani, veri 2×2 lik bir çapraz tabloda özetlendiği zaman), orantıların eşitliğinin testi için z-dağılımını kullanacağız. Aralarındaki ilişkiyi test edeceğimiz nitel değişkenlerden herhangi birinin ikiden çok sayıda olanaklı kategoriye sahip olması halinde, z-testi artık kifayet etmeyeceği için, daha genel Ki-kare testini kullanacağız (15. Bölüm).

İki orantının farkı için güven aralığı kestiriminde olduğu gibi, n_1p_1, $n_1(1-p_1)$, n_2p_2 ve $n_2(1-p_2)$ değerlerinden her biri en az 10 olmalıdır; örneklerin hem *basit rastgele örnekler* olmaları hem de biribirinden *bağımsızca çekilmiş* olmaları önemlidir. Nihayetinde, örneklerin çekildiği popülasyonların da örnek hacimlerine kıyasla geniş olmaları gerekir.

Bir popülasyondan çekilen bir örneğin orantısının, ikinci bir popülasyondan çekilen bir örneğin orantısından daha büyük olması, popülasyonların kendi orantıları hakkında otomotik olarak benzer bir neticeye varılmasını haklı çıkarmaz. İki husus üzerinde durmamız gerekiyor. Birincisi, aynı popülasyondan çekilecek aynı hacimdeki örneklerin orantıları değişkendir. İkincisi, gerçekte karşılaştırdığımız şeyler, noktasal değerler değil, güven aralıklarıdır.

Pek çok problem için, sıfır hipotezi popülasyon orantılarının eşit olduğunu ifade eder, ya da aynı manada, iki orantının farkının sıfır olduğunu ifade eder:

$$H_0: \pi_1 - \pi_2 = 0$$

Buna karşılık, alternatif hipotez, üç farklı şekilde kurulabilmektedir,

$$H_1: \pi_1 - \pi_2 < 0, \qquad H_1: \pi_1 - \pi_2 > 0 \qquad \text{veya} \quad H_1: \pi_1 - \pi_2 \neq 0$$

ki burada ilk iki alternatif bizi tek kuyruklu testlere götürürken üçüncü alternatif çift kuyruklu bir teste götürür.

Sıfır hipotezimiz $\pi_1 = \pi_2$ olduğu için, müşterek bir π değeri farz edip, $\sigma_{p_1-p_2}$ nin değerini hesaplarken onu kullanacağız,

$$\sigma_{p_1-p_2} = \sqrt{\frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2}} = \sqrt{\pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Uygulamada, $\sigma_{p_1-p_2}$ nin değerini doğrudan hesaplayamayız, çünkü π bilinmiyor. Şunu yaparız, eğer $p_1 = \frac{y_1}{n_1}$ ve $p_2 = \frac{y_2}{n_2}$ ise, π nin bir kestirimi olarak $p = \frac{y_1+y_2}{n_1+n_2}$ yi kullanmak suretiyle, bir yaklaşım olarak, $s_{p_1-p_2}$ yi hesaplarız.

Misal 11 Kırsal bir bölgede yerel bir TV kanalının işbirliği ile yürütülen ve doğum kontrol yöntemlerinin ailelere benimsetilmesini amaçlayan bir eğitim çalışmasından önce yapılan yoklamada rastgele seçilen 1000 aileden 520 si bir doğum kontrolü yöntemini kullandığını ifade etmişken, eğitim çalışmasından sonra yapılan yoklamada yine rastgele seçilen 800 aileden 460 ı bir doğum kontrolü yöntemini kullandığını ifade etmiştir. Yüzde 10 anlamlılık düzeyinde, yeterli kanıt var mı ki ailelerin doğum kontrolü yöntemlerini benimsemelerinin arttığını göstersin.

Cevap: Görüyoruz ki $n_1p_1 = 520$, $n_1(1-p_1) = 480$, $n_2p_2 = 460$, $n_2(1-p_2) = 340$ ve her biri 10 dan büyük; örneklerin birer basit rastgele örnek ve bağımsız olduklarını varsayıyoruz; açık ki ailelerin popülasyonu geniş bir popülasyon.

$$H_0: \pi_1 - \pi_2 = 0$$

$$H_1: \pi_1 - \pi_2 < 0$$

$$\alpha = 0.10$$

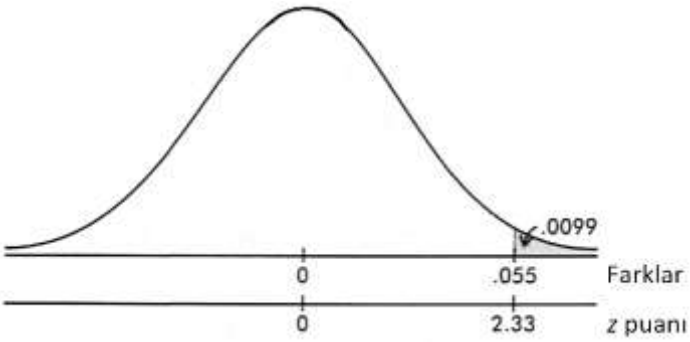
$$p_1 = \frac{520}{1000} = 0.52, \; p_2 = \frac{460}{800} = 0.575 \;\text{ ve }\; p = \frac{520+460}{1000+800} = 0.544$$

$$s_{p_1-p_2} = \sqrt{0.544(0.456)\left(\frac{1}{1000} + \frac{1}{800}\right)} = 0.236$$

Gözlenen fark $0.575 - 0.52 = 0.055$ ve bu farka tekabül eden z-puanı,

$$\frac{0.055 - 0}{0.0236} = 2.33$$

ve P-değeri $1 - 0.9901 = 0.0099$. Böylece, $0.0099 < 0.10$ olduğu için, %10 anlamlılık düzeyinde şu neticeye varıyoruz ki eğitim çalışması sonucunda ailelerin bilinçliliği artmıştır. Şunu da belirtelim ki $0.0099 < 0.01$ olduğundan dolayı, gözlenen fark hatta %1 düzeyinde bile istatistiksel olarak anlamlıdır.



Misal 12 Bir otomobil fabrikasında iki ayrı montaj preosedürü denenmektedir. A prosedürünün kullanıldığı üretim bandından çıkan 350 arabalık bir örneğin 28 i kusurluyken, B prosedürünün uygulandığı banttan çıkan 500 arabalık örneğin 32 si kusurlu. Yüzde 10 anlamlılık düzeyinde, fark anlamlı mı?

Cevap: Görüyoruz ki $n_1p_1 = 28$, $n_1(1-p_1) = 322$, $n_2p_2 = 32$, $n_2(1-p_2) = 468$ ve her biri 10 dan büyük; örneklerin birer basit rastgele örnek ve bağımsız olduklarını varsayıyoruz; ve varsayıyoruz ki bu fabrikada üretilen tüm otomobillerin popülasyonu geniş bir popülasyondur. Montaj prosedürlerinden herhangi birinin daha iyi ya da daha kötü olduğuna dair bir inanç belirtilmediği için, bu çift kuyruklu bir testtir.

$$H_0: \pi_1 - \pi_2 = 0$$

$$H_1: \pi_1 - \pi_2 \neq 0$$

$$\alpha = 0.10$$

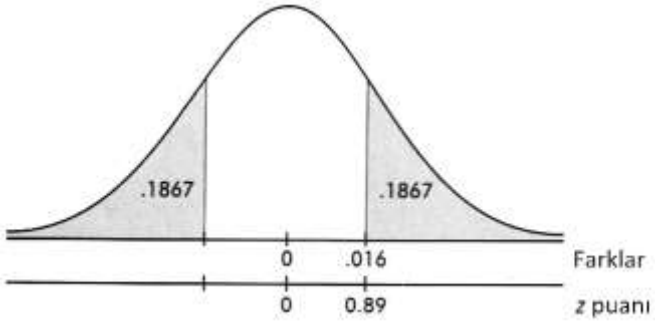
$$p_1 = \frac{28}{350} = 0.08,\ p_2 = \frac{32}{500} = 0.064 \text{ ve } p = \frac{28+32}{350+500} = 0.0706$$

$$s_{p_1-p_2} = \sqrt{0.0706(0.9294)\left(\frac{1}{350} + \frac{1}{500}\right)} = 0.0179$$

Gözlenen fark $0.08 - 0.064 = 0.016$ ve bu farka tekabül eden z-puanı,

$$\frac{0.016 - 0}{0.0179} = 0.89$$

ve bu z-puanına ait kuyruk olasılığı $1 - 0.8133 = 0.1867$. Test çift kuyruklu olduğu için, bu değerin iki katını almak suretiyle, P-değerimiz $2(0.1867) = 0.3734$ olur. Böylece, $0.3734 > 0.10$ olduğundan dolayı, %10 anlamlılık düzeyinde şu neticeye varıyoruz ki gözlenen fark, istatistiksel olarak anlamlı bir fark değildir. Şunu da belirtelim ki gözlenen farkın anlamlı olabildiği en küçük anlamlılık düzeyi, %37.34 ten daha yukarıdadır. P-değeri bu denli yüksek olduğu zaman, güvenle söyleyebiliyoruz ki H_0 a karşı hiç bir kanıt yoktur. Yani, iki ayrı montaj prosedürüyle üretilen kusurlu arabaların orantıları arasında gözlenen fark anlamlı değildir.

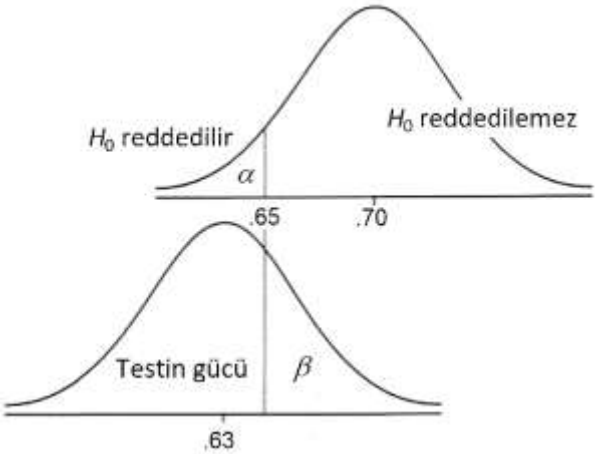


Testin Gücü

Popülasyon parametresinin her olanaklı doğru değeri için β nın farklı bir değeri mevcuttur. Her β için, $1-\beta$, mevcut doğru değere karşı testin gücü olarak adlandırılır. Bir anlamlılık testinin *gücü* II. tip hatanın işlenmeme olasılığıdır. Yani, doğru bir alternatif hipotez kurulduğu zaman, testin gücü yanlış sıfır hipotezini reddetme olasılığıdır. Hem örnek hacminin artırılması hem anlamlılık düzeyinin artırılması, kendi başına, testin gücünü artırır.

Şimdilik, bir testin gücünü veya II. tip hata olasılığını hesaplamaya girişmiyoruz. Zaten, referans alacağımız bir spesifik alternatif hipotez elimizde olmadığı sürece bunu yapamayız. Fakat, bu kavramları ve bu kavramlar arasındaki interaksiyonları anlamak gerekir.

Misal 13 Yaşadığın şehrin belediye başkan adaylarından biri, arkasındaki seçmen desteğinin %70 olduğunu iddia ediyor, fakat sen de bu desteğin daha düşük olduğuna inanıyorsun. Bir basit rastgele örnek alarak, eğer örnekte bu adaya gösterilen destek %65 veya daha az çıkarsa, %70 lik iddiayı reddetmeyi planlıyorsun. Peki, eğer seçmenlerin gerçekte yalnızca %63 ü bu adayı destekliyor ise?



Grafiklerden üstte olanı, $\pi = 0.70$ iddiası ile sıfır hipotezi modelini ve eğer $p < 0.65$ ise H_0 ı reddetme planını gösteriyor. Alttaki grafik, $\pi = 0.63$ ile gerçek modeli gösteriyor. Sıfır hipotezinin doğru olmadığını söylemekte ne zaman başarısız oluruz? Başka bir deyişle, H_0 ı ne zaman reddedemeyiz?

Cevap: Tam olarak, örnek orantısı 0.65 ten daha büyük olduğu zaman. Bu, β olasığına sahip II. Tip hatadır.

Sıfır hipotezinin doğru olmadığı neticesine tam olarak ne zaman varırız?

Cevap: Örnek orantısı 0.65 ten daha küçük olduğu zaman. Bu, testin gücüdür ve $1-\beta$ olasılığına sahiptir.

Testin gücüne dair, şu hususlar önem arzeder:

- Testin gücü, II. Tip hatadan kaçınma olasılığını verir.
- Testin gücü, popülasyon parametrelerinin farklı olanaklı doğru değerleri için farklı değerlere sahiptir.
- Daha küçük bir α seçmek (yani, H_0 ı reddedebilmeyi daha zorlaştırmak), daha yüksek II. Tip hata riski ve testin gücünün azalmasıyla sonuçlanır. Yukarıdaki grafiğe bakarsan, α küçülürken (yani, kritik değer noktasının sola doğru kayması durumunda), testin gücünün azaldığını ve β nın büyüdüğünü görebilirsin.
- Sıfır hipotezinin popülasyon parametresi için ileri sürdüğü değer ve popülasyon parametesinin hakiki değeri arasındaki fark arttıkça, II. Tip hata riski azalır ve testin gücü artar. Yukarıdaki grafiklerin altta olanını sola doğru kaydırdıkça, testin gücünün arttığını ve β nın küçüldüğünü görebilirsin. (Sıfır hipotezinin popülasyon parametresi için ileri sürdüğü değer ve popülasyon parametresinin hakiki değeri arasındaki fark, çoğu zaman *etki* olarak adlandırılır; ve böylece

etki ne kadar büyük olursa, o etkiyi ortaya koymakta, anlamlılık testi o kadar güçlü olur.

- Geniş bir n örnek hacmi, standart sapmaları azaltacağından dolayı, her iki grafiği de daha daraltacak ve bunun sonucunda α ve β küçülürken testin gücü artacaktır.

Anlamlılık testi problemlerini çözerken;
(1) hipotezleri, tüm değişkenleri açıkça tanımlamak suretiyle, popülasyon parametreleri cinsinden ifade etmelisin;
(2) icra edeceğin testi ismiyle ya da formülü ile belirtmelisin ve test varsayımlarının yerine getirilip getirilmediğini göstermelisin;
(3) test istatistiğini ve ilişkin P-değerini hesaplamalısın;
(4) P-değerini bir karara bağlamalısın; ve
(5) verilen problem bağlamında vardığın neticeyi ifade etmelisin.

Mini Proje 14

Duyu-ötesi algı için bir deney dizayn edeceksin ve yürüteceksin. Şunu öğrenmek istiyoruz: Kadınlar ve adamlar arasında duyu-ötesi algı bakımından istatistikselce anlamlı bir fark var mı? Bir oyun destesine ait kartları kullanarak, göstermeden seçtiğin kartların türünü veya rengini adamlardan ve kadınlardan tahmin etmelerini iste. Oyun kartları yerine, kendi yarattığın hedef resimleri de kullanabilirsin.

a. Deneyi nasıl dizayn ettiğini ve yürüttüğünü açıkla.
b. Deneyin için sıfır hipotezini ve alternatif hipotezi ifade et.
c. Sonuçlarını rapor et. Sonuçların interpretasyonunu yap.
d. Hangi neticeye vardın?
e. Testin gücü hakkında ne söyleyebilirsin?

BÖLÜM 15 | ANLAMLILIK TESTLERİ: FREKANSKLAR VE EN KÜÇÜK KARELER DOĞRUSUNUN EĞİMİ

Giriş

Bu bölümde, veri analizinde kullandığımız araçları geliştirmeye devam edeceğiz. Ki-kare olasılık modellerini kullanarak, nitel değişkenlerin frekans dağılımları hakkında çıkarımda bulunmayı tartışacağız. Bu frekanslar tek bir nitel değişkenin kategorilerine ait olabileceği gibi, iki ayrı nitel değişkenin ortak frekansları da olabilecek. Bu, sana, şu türden gerçek dünya problemlerini çözmende yardımcı olacak: "Test sonuçları, Mendel'in genetik prensiplerini destekliyor mu?" (Uyum iyiliği testi); "Bin dokuz yüz on iki yılında Atlantik'te buz dağına çarparak batan RMS Titanik adlı gemide seyahat eden yolcuların akıbetleri seyahat ettikleri mevkiden bağımsız mı?" (Bağımsızlık testi); ve "Türklerin, Yunanlıların ve Bulgarların kan gruplarına göre dağılımları aynı mı?" (Homojenlik testi).

Bir başka deyişle, bu bölümde, önce, bir nitel değişkenin gözlenen frekans dağılımının anlamlılık testini (uyum iyiliği testi), iki nitel değişken arasında gözlenen bir ilişkinin anlamlılık testini (bağımsızlık testi) ve iki nicel değişken arasında gözlenen bir ilişkinin anlamlılık testini nasıl yapacağını öğreneceksin.

Daha sonra, iki nicel değişken arasındaki doğrusal çağrışım hakkında çıkarımda bulunmayı tartışacağız. Bu da, sana, şu türden gerçek dünya problemlerini çözmende yardımcı olacak: "Sınavdan alınan başarı puanı ve öğrencinin sınava çalışma süresi arasında doğrusal bir ilişki var mı?"

Ki-kare ile Uyum İyiliği Testi

Kategorik veride gözlenen bir dağılım biçiminin, belli bir dağılıma uyup uymadığı sorusu, her zaman kritik bir sorudur. Pratikte mükemmel bir uyum beklemek elbette gerçekçi değil. Fakat, gözlenen frekanslar ve beklenen frekanslar arasındaki farklara bakmak ve *uyum iyiliğine* dair bir hüküm vermek gerekir.

Daha önce geliştirdiğimize benzer bir yaklaşımda bulunuruz. Şöyle ki: Uyumun iyiliğini ifade eden bir sıfır hipotezimiz vardır. Yani, belli bir teorik dağılımın, ilgilendiğimiz verinin dağılımını doğru bir şekilde betimlediğini iddia eden hipotez. Elimizdeki gözlem verisi tüm olanaklı örnekler uzayından çekilmiş örneklerden biridir. Eğer sıfır hipotezi gerçekte doğru ise, kategorilerin gözlenen ve beklenen frekansları arasındaki farklara sahip bu örneği çekme şansı nedir? Nihayetinde, bu şans çok küçük ise, sıfır hipotezini reddederiz ve uyumun iyi olmadığını söyleriz.

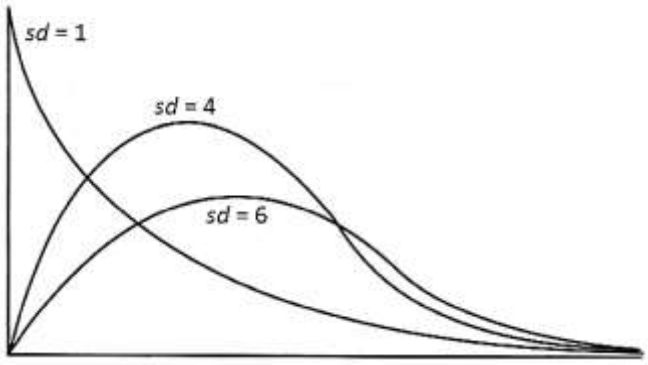
Gözlenen frekanslar ve beklenen frekanslar arasındaki farkların anlamlılığına nasıl karar vereceğiz? En iyi enformasyonun bizzat fark değerlerince temin edildiğini söylersem, belki şaşırtıcı gelebilir. Fakat, daha evvel, bir değişkenin varyansını öğrenmeye çalışırken yaptığımız gibi, burada da farkların karelerini alacağız. Ancak, mesela, 23 olarak gözlenen bir fark eğer 133 ve 110 un farkı ise, 10630 ve 10607 nin farkına göre daha anlamlı olacağı için, her bir farkı mütenasipçe ağırlıklandırmak gerekir. Bu ağırlıklandırmayı, her bir farkı kendi beklenen değerine bölmek suretiyle başarırız. Bu ağırlıklandırılmış kareli farkların toplamı, *ki-kare* olarak adlandırılır ve χ^2 (Yunan alfabesinin küçük chi harfi) ile simgelenir:

$$\chi^2 = \sum \frac{(\text{Gözlenen} - \text{Beklenen})^2}{\text{Beklenen}}$$

Örnekten hesaplanan χ^2 değeri ne kadar küçük ise, uyum o kadar iyidir. Sıfır hipotezinin doğru olduğu varsayımı altında, elde edilen kadar aşırı bir χ^2 değerinin elde edilmesi olasılığı, uyum iyiliği testi için kullanacağımız *P*-değeridir. Eğer χ^2 değeri yeterince büyük ise, yani, *P*-değeri yeterince küçük ise, sıfır hipotezini reddetmek için yeterli kanıt olduğunu söyleriz ve uyumun zayıf olduğunu iddia ederiz.

Hesaplanan bir χ^2 değerinin, anlamlı olabilmesi için ne kadar büyük olması gerektiğine karar vermek

için, yani, bir kritik değer seçebilmek için, χ^2 değerlerinin nasıl dağıldığını anlamak gerekir. Bir χ^2 dağılımı simetrik değildir ve daima sağa çarpıktır. Her biri bir serbestlik derecesine (sd) ilişkin, biribirinden ayrı χ^2 dağılımları mevcuttur. Serbestlik derecesinin değeri ne kadar büyük olursa, χ^2 dağılımı da normal dağılıma o kadar yaklaşır.



Bu bağlamda, büyük bir χ^2 değerinin anlamlı olup olmayacağı, hangi χ^2 dağılımını kullandığımıza bağlı olacaktır. Çok müştereken kullanılan yüzdeler veya olasılıklar için kritik χ^2 değerlerinin bir tablosu Tablo D de verilmiştir. Uyum iyiliği problemlerinde χ^2 dağılımını kullanabilmek için, beklenen değerlerin çok küçük olmamaları gerekir. Genel bir kural olarak, hiç bir beklenen değer 5 ten küçük olmamalıdır. Nihayetinde, konu ettiğimiz tüm anlamlılık testlerinde olduğu üzere, elimizdeki örnek belli bir popülasyondan rastgele seçilmiş olmalıdır.

Misal 1 Geçen yıl, prime-time'da ATV, BTV, CTV ve DTV televizyon kanalları tüm seyircinin sırasıyla %30 u, %25 i, %20 si ve %25 i tarafından izlendi. Bu yılın ilk haftası sonunda 500 izleyici ile bir yoklama yapıldı.
a. Eğer izleyici tercihleri değişmediyse, her bir kanalı kaç kişinin izlemesi beklenir?
Cevap: $0.30(500) = 150$, $0.25(500) = 125$, $0.20(500) = 100$ ve $0.25(500) = 125$; böylece,

	Kanal			
	ATV	BTV	CTV	DTV
Beklenen izleyici sayısı	150	125	100	125

b. Fiilen gözlenen izleyici sayıları şöyledir:

	Kanal			
	ATV	BTV	CTV	DTV
Gözlenen izleyici sayısı	139	138	112	111

Bu izleyici sayıları bir değişikliği gösterir mi? Farklar anlamlı mı?
Cevap: Varsayımları kontrol edelim, önce. Beş yüz izleyicinin temsilci bir örnek olduğunu varsayıyoruz. Beklenen değerlerin (150, 125, 100, 125) hepsi 5 ten büyük.

H_0: Televizyon izleyicileri ATV, BTV, CTV ve DTV kanallarına, sırasıyla, %30, %25, %20 ve %25 lik yüzdelerle dağılmaktadır.
H_1: Televizyon izleyicileri ATV, BTV, CTV ve DTV kanallarına, sırasıyla, %30, %25, %20 ve %25 lik yüzdelerle dağılmamaktadır.

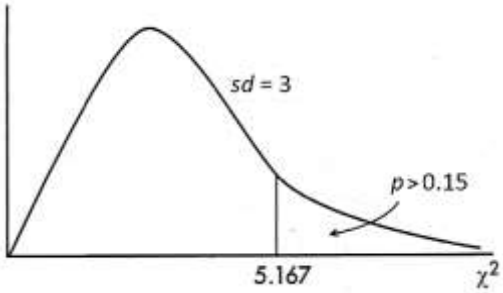
Ki-kare değerini hesaplıyoruz:

$$\chi^2 = \sum \frac{(\text{Gözlenen} - \text{Beklenen})^2}{\text{Beklenen}}$$
$$= \frac{(139-150)^2}{150} + \frac{(138-125)^2}{125} + \frac{(112-100)^2}{100} + \frac{(111-125)^2}{125}$$
$$= 5.167$$

Buna göre, *P-değeri*ni, Tablo D de, $sd = n - 1 = 3$ ile, şu olasılık aralığında buluruz,

$$0.15 < P(\chi^2 > 5.167) < 0.20.$$

Bu kadar büyük bir P-değeri $(P > 0.15)$ ile H_0 ı reddetmek için yeterli kanıt yoktur. Yani, izleyici tercihlerinin değiştiğini gösteren yeterli kanıt mevcut değil.



Misal 2 Kapıda satış yapan bir firmanın yöyeticisi, beş ayrı satış temsilcisinin eşit sayıda ürün satıp satmadıklarını belirlemek istiyor. Yöneticinin rastgele seçtiği bir ayda satış temsilcilerinin sattıkları ürün sayıları şöyledir:

	Satış temsilcisi				
	I	II	III	IV	V
Gözlenen satış sayısı	43	29	52	34	48

Satış temsilcilerinin sattıkları ürün sayılarının eşit olmadığına dair yeterli kanıt var mı? Hem %5 hem de %10 anlamlılık düzeylerinde test et.

Cevap:

H_0: Satılan ürün sayıları, satış temsilcileri arasında üniform dağılmaktadır.
H_1: Satılan ürün sayıları, satış temsilcileri arasında üniform dağılmamaktadır.

Satılan ürünlerin toplam sayısı, $43 + 29 + 52 + 34 + 48 = 206$ adettir. Eğer satış temsilcileri eşit sayıda ürün satmış olsalardı, her bir satış temsilcisinin,

$$\frac{206}{5} = 41.2$$

adet ürün satması beklenirdi (üniform dağılım).

	Satış temsilcisi				
	I	II	III	IV	V
Beklenen satış sayısı	41.2	41.2	41.2	41.2	41.2

Giriştiğimiz anlamlılık testinin varsayımları karşılanıyor mu? *Rastgelelik*: Satılan 206 adet ürünün temsilci bir örnek olduğunu varsayıyoruz. Beklenen değerlerin hepsi 5 ten büyüktür. Böylece,

$$\chi^2 = \frac{(43-41.2)^2}{41.2} + \frac{(29-41.2)^2}{41.2} + \frac{(52-41.2)^2}{41.2} + \frac{(34-41.2)^2}{41.2} + \frac{(48-41.2)^2}{41.2}$$
$$= 8.903$$

Serbestlik derecesi, sınıfların sayısının 1 eksiği olduğuna göre,

$$sd = 5 - 1 = 4.$$

P-değeri için Tablo D ye baktığımız zaman, görüyoruz ki

$$0.10 > P(\chi^2 > 8.903) > 0.05.$$

$P < 0.10$ olduğuna göre, H_0 ı %10 anlamlılık düzeyinde reddetmek için yeterli kanıt vardır. Fakat, H_0 ı %5 anlamlılık düzeyinde reddetmek için yeterli kanıt yoktur, çünkü $P > 0.05$.

Uyum iyiliği testi tek örnek orantısının z-testini genişletir. Eğer çıkımın yalnızca iki kategorisi varsa, tek örnek orantısının z-testini kullanırız. Kategorilerin sayısı ikiden fazlaysa, uyum iyiliği

testini kullanırız.

Ki-kare ile Bağımsızlık Testi

Yukarıda ele aldığımız uyum iyiliği problemlerinde, bir beklenen değerler kümesi, dağılımın nasıl olması gerektiğine dair bir varsayıma dayandırıldı. Ondan sonra, gözlenen bir örnek dağılımının varsayılan bir dağılıma sahip daha geniş bir kümeden gelip gelmediğini test ettik.

Pek çok gerçek dünya problemlerinde, çekilen bir örnekte gözlenen iki nitel değişkenin ortak dağılımını, beklenen dağılım hakkında herhangi bir varsayımda bulunmaksızın, karşılaştırmak isteriz. *Bağımsızlık testi* adı verilen prosedür ile, iki değişken bakımından sınıflandırılan tek bir örneği ele alacağız ve aralarında hakiki bir ilişki olup olmadığını öğrenmeye çalışacağız. Mesela, tütün içmeyenlerde, hafif tütün içenlerde ve ağır tütün içenlerde nihayetinde kanser, kalp hastalığı veya amfizem teşhis edilmesi aynı olabilirlikte midir? Tütün içme durumu ile bu hastalıklardan birinin teşhis edilmesi arasında bir ilişki (çağrışım) var mıdır? Örnekte gözlenen çağrışımı örneğin çekildiği popülasyona genelleştirmek anlamlı mıdır?

Gözlemlerimizi iki boyutlu sınıflandırırız ve ondan sonra bu iki boyutun biribirinden bağımsız olup olmadığını bulmaya çalışırız. Mesela, çeşitli yaş gruplarını nazar-ı dikkate alabiliriz ve bu yaş gruplarındaki kişilerin çeşitli hayattan memnuniyet düzeylerine nasıl dağıldığını öğrenmek isteyebiliriz. Bu durumda, sıfır hipotezimiz yaşın ve hayattan memnuniyetin biribirinden bağımsız olduğunu iddia eder. Yani, herhangi bir kişinin ait olduğu yaş grubu ve memnuniyet düzeyi arasında bir ilişki olmadığını iddia eder.

Analiz, iki nitel değişkenin bağımsız olduğunu iddia eden sıfır hipotezinin doğruluğu varsayımı altında, çapraz tablodaki her hücre için *beklenen değerlerin* hesaplanmasını gerektirir. Ondan sonra, beklenen değerleri gözlenen değerlerle karşılaştırırız ve meydana gelen farkların, eğer H_0 doğru olsaydı makul olup olmayacağını sorarız. Farkların anlamlılığı, uyum iyiliği testinde yapıldığı üzere, ağırlıklandırılmış kareli farkların toplamı olan χ^2 değeri ile kritize edilecektir. Hesaplanan χ^2 değeri ne kadar küçük ise, bağımsızlığı ifade eden sıfır hipotezi o kadar makuldür. Eğer χ^2 değeri yeterince büyük ise, yani, eğer P-değeri yeterince küçük ise, sıfır hipotezini reddetmek için yeterli kanıt olduğunu söyleyebiliriz ve iki değişken, metot veya sınıflandırma arasında bir *ilişki olduğunu* iddia edebiliriz.

Bu tipteki çapraz tablo problemlerinde, adet olduğu üzere, satır sayısını r ile ve sütun sayısını c ile simgeledikten sonra, serbestlik derecesi şöyle hesaplanır:

$$sd = (r-1)(c-1).$$

Belirtmeden geçmememiz gereken bir husus şudur ki değişkenlerin bağımsızlığını ifade eden sıfır hipotezini reddetmek için yeterli kanıt olsa bile, iki değişken arasında doğrudan bir *nedensel* ilişki olduğunu iddia edebilmeyiz. Başka bir deyişle, iki değişken arasında bir çağrışıma, bağlantıya veya ilişkiye dair bir ifadede bulanabilmekle beraber, bu değişkenlerden birinin diğerinin nedeni olduğu iddiasını gerekçelendirebilmeyiz. Mesela, bir yoklama verisinin analizi sonucunda, ücret düzeyi ve iş memnuniyeti arasında bir ilişki ortaya çıkarılabilir, fakat bu veri toplama metoduyla yüksek ücret düzeyinin yüksek iş memnuniyetine neden olduğunu gösteremeyiz. Belki, bir çalışanın yüksek iş memnuniyetinden, üstleri etkilenmiş ve bu çalışanın ücretini artırmış olabilirler. Ya da, belki, öğrenim düzeyi, eğitim düzeyi, veya kişilik gibi bir üçüncü değişken vardır ki hem ücret düzeyi hem de iş memnuniyeti ile doğrudan bir nedensel ilişkiye sahiptir.

Misal 3 Ülke genelinde 1000 yetişkin ile yapılan bir telefon yoklamasında deneklere öğrenim düzeyleri ve televizyonda izlemeyi tercih ettikleri program türü soruldu. Yanıtlar aşağıdaki çapraz tabloda özetlendi:

Program tercihi

		Film	Müzik	Magazin
Öğrenim düzeyi	İlk	175	220	55
	Orta	150	165	35
	Yüksek	75	105	20

Yetişkinlerin televizyon programı tercihlerinin öğrenim düzeylerinden bağımsız olduğunu ifade eden sıfır hipotezini test et. Yüzde 10 luk anlamlılık düzeyini kullan.

Cevap:

H_0: Televizyon programı tercihi ve öğrenim düzeyi bağımsızdır.
H_1: Televizyon programı tercihi ve öğrenim düzeyi bağımsız değildir.

Yukarıdaki tablo bize gözlenen sonuçları veriyor. beklenen değerleri belirlemek için, önce satır ve sütun toplamlarını belirlememiz gerekiyor:

Gözlenen		Program tercihi			Toplam
		Film	Müzik	Magazin	
Öğrenim düzeyi	İlk	175	220	55	450
	Orta	150	165	35	350
	Yüksek	75	105	20	200
	Toplam	400	490	110	1000

Çapraz tablonun her bir hücresi satır ve sütun değişkenlerinin her bir kombinasyona tekabül etmektedir. Mesela sol üst hücre için beklenen değeri hesaplamak için şu yaklaşımlarda bulunabiliriz. Önce şöyle yapalım, ilk öğrenimlillerin orantısını bulalım,

$$\frac{450}{1000} = 0.45;$$

şimdi buna göre, eğer bağımsız ise, ilk öğrenimliler içerisinde film izlemeyi tercih edenlerin beklenen sayısı,

$$0.45(400) = 180.$$

Böyle yapmak yerine, film izlemeyi tercih edenlerin orantısını bulalım,

$$\frac{400}{1000} = 0.40;$$

şimdi buna göre, eğer bağımsız ise, film izlemeyi tercih edenler içerisinde ilk öğrenimlilerin beklenen sayısı,

$$0.40(450) = 180.$$

Nihayetinde, görüyoruz ki bu hesaplamaların her ikisi de basitçe şunu gerektiriyor:

$$\frac{450(400)}{1000} = 180.$$

Başka bir deyişle, *herhangi bir hücrenin beklenen değeri, mütekabil satır toplamının ve sütun toplamının çarpımının genel toplama bölümüdür*. Böylece, mesela, ikinci satır ve ikinci sütunun kesiştiği hücrenin beklenen değeri, ki orta öğrenimlilerin müzik programı tercihine tekabül eder,

$$\frac{350(490)}{1000} = 171.5.$$

Bu yolla devam ederek, tabloyu hücrelerin beklenen değerleriyle şöyle doldururuz:

Beklenen		Program tercihi			Toplam
		Film	Müzik	Magazin	
Öğrenim düzeyi	İlk	180	220.5	49.5	450
	Orta	140	171.5	38.5	350

	Yüksek	80	98	22	200
	Toplam	400	490	110	1000

Şimdi, ki-kare değerini hesaplayabiliriz:

$$\chi^2 = \frac{(175-180)^2}{180} + \frac{(220-220.5)^2}{220.5} + \frac{(55-49.5)^2}{49.5}$$
$$= \frac{(150-140)^2}{140} + \frac{(165-171.5)^2}{171.5} + \frac{(35-38.5)^2}{38.5}$$
$$= \frac{(75-80)^2}{80} + \frac{(100-98)^2}{98} + \frac{(20-22)^2}{22}$$
$$= 3.024$$

Şunu da görebiliyoruz ki 180, 220.5, 140 ve 171.5 luk beklenen değerlerler hesaplandıktan sonra, geriye kalan hücrelerin beklenen değerleri satır veya sütun toplamları kullanılarak bulunabilir. Böyle olunca, serbestlik derecesi, yani bağımsızca değer alabilen hücre sayısı,burada 4 tür. Ya da yine diğer yoldan hesaplayalım,

$$sd = (r-1)(c-1) = (3-1)(3-1) = 4.$$

Bu serbestlik derecesi ile, *P*-değeri için Tablo D ye baktığımız zaman, görüyoruz ki

$$P(\chi^2 > 3.024) > 0.25.$$

Bu kadar büyük bir *P*-değeri ile, H_0 ı %10 anlamlılık düzeyinde reddedemeyiz. Yani, topladığımız veri, öğrenim düzeyi ve televizyon programı tercihi arasında herhangi bir ilişki olduğuna dair bir kanıt sağlamıyor.

Ki-kare ile bağımsızlık testlerinin varsayımlarına gelince, örneğin rastgele seçilmiş olması ve her hücrenin beklenen değerinin 5 ten büyük olması gerekmektedir. Eğer bir hücrenin beklenen değeri, 5 ten küçük bulunursa, ilgili kategorilerin birleştirilmesi yoluna, mantıklı olmak koşuluyla, gidilebilir. Mesela, "aynı fikirde değilim" ve "hiç aynı fikirde değilim" diye iki kategori var ise ve bu kategorilere ait hücrelerde 5 ten az beklenen değer hesaplanmış ise, bu iki kategori birleştirilebilir. Böyle bir olanaktan mahrum olunduğu durumlarda ise, 5 ten küçük beklenen değere sahip kategoriler "diğerleri" başlığı altında birleştirilebilirler.

Misal 4 Çocuklukta tacize uğramış ve belli bir kusurlu geni taşıyan erkeklerin şiddet suçu işlemeye daha meyilli olup olmadığını belirlemek için, 25 ve 35 yaşları arasındaki 575 erkeğin bir basit rastgele örneği üzerinde bir öğrenme-çalışması yürütülmüştür. Veri aşağıdaki tabloda özetlenmiştir:

	Tacize uğramamış, normal gen	Tacize uğramış, normal gen	Tacize uğramamış, kusurlu gen	Tacize uğramış, kusurlu gen
Suçlu davranış	48	21	32	26
Normal davranış	201	79	118	50

a. Kişilik durumunun dört kategorisi ve davranış biçimi arasında bir ilişkiyi gösteren kanıt mevcut mu? Açıkla.

Cevap: Ki-kare ile bağımsızlık testi sorumuzu cevaplamamızı sağlayacaktır.

H_0: Kişilik durumunun dört kategorisi ve davranış biçimi bağımsızdır.
H_1: Kişilik durumunun dört kategorisi ve davranış biçimi bağımsız değildir.

Hücreler için beklenen değerler şöyledir:

55.0	22.1	33.1	16.8
194.0	77.9	116.9	59.2

Hücrelerin hepsinde beklenen değerler 5 ten büyüktür. Dolayısıyla varsayım karşılanır. Zaten elimizdekinin bir basit rastgele örnek olduğu enformasyonuna sahibiz.

Hücrelerin gözlenen ve beklenen değerlerini kullanarak, ki-kare değeri, $\chi^2 = 7.752$ olarak hesaplanır. Bu ki-kare değerine ilişkin serbestlik derecesi, $sd = (2-1)(4-1) = 3$ tür. Buna göre, Tablo D de P-değerini,

$$0.10 > P(\chi^2 > 7.752) > 0.05$$

olarak okuruz. Sıfır hipotezini %10 anlamlılık düzeyinde reddetmek için kanıt mevcut. Yani, kişilik durumunun dört kategorisi ve davranış biçimi arasında bir ilişki olduğu yargısında bulunabilmemizi sağlayacak kanıtımız var.

b. Normal gene sahip erkekler içerisinde, şiddet suçu işleyen tacize uğramış erkeklerin orantısı, şiddet suçu işleyen tacize uğramamış erkeklerin orantısından büyük olduğunu gösteren kanıt mevcut mu? Açıkla.

Cevap: İki orantının *z* testi sorumuzu cevaplamamızı sağlayacaktır. Önce *n* in yeterince büyük olup olmadığını kontrol etmeliyiz: $n_1p_1 = 21 > 10$, $n_1(1-p_1) = 79 > 10$, $n_2p_2 = 48 > 10$, $n_2(1-p_2) = 201 > 10$. Verilmemekle beraber, örneklerin birer basit rastgele örnek ve bağımsız olduklarını varsayıyoruz. Açık ki popülasyon geniş.

$H_0: \pi_1 - \pi_2 = 0$ (ki burada π_1 şiddet suçu işleyen normal gene sahip tacize uğramış erkeklerin orantısı ve π_2 şiddet suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısıdır)

$H_1: \pi_1 - \pi_2 > 0$ (şiddet suçu işleyen normal gene sahip tacize uğramış erkeklerin orantısı, şiddet suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısından daha büyüktür)

$$p_1 = \frac{21}{100} = 0.21,\ p_2 = \frac{48}{249} = 0.193 \text{ ve } p = \frac{21+48}{100+249} = 0.198$$

$$s_{p_1-p_2} = \sqrt{0.198\,(0.802)\left(\frac{1}{100} + \frac{1}{249}\right)} = 0.0472$$

Gözlenen fark $0.21 - 0.193 = 0.017$ ve bu farka tekabül eden *z* puanı,

$$\frac{0.017 - 0}{0.0472} = 0.36$$

ve P-değeri $1 - 0.641 = 0.359$. Böylece, $0.359 > 0.01$, $0.359 > 0.05$, $0.359 > 0.10$ ve hatta $0.359 > 0.20$ olduğu için, şu neticeye varıyoruz ki şiddet suçu işleyen normal gene sahip tacize uğramış erkeklerin orantısının, şiddet suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısından daha büyük olduğunu gösteren kanıt yoktur.

c. Çocukluğunda tacize uğramamış erkekler içerisinde, şiddet suçu işleyen kusurlu gene sahip erkeklerin orantısı, şiddet suçu işleyen normal gene sahip erkeklerin orantısından büyük olduğunu gösteren kanıt mevcut mu? Açıkla.

Cevap: İki orantının *z* testi sorumuzu cevaplamamızı sağlayacaktır. Önce *n* in yeterince büyük olup olmadığını kontrol etmeliyiz: $n_1p_1 = 32 > 10$, $n_1(1-p_1) = 118 > 10$, $n_2p_2 = 48 > 10$, $n_2(1-p_2) = 201 > 10$. Verilmemekle beraber, örneklerin birer basit rastgele örnek ve bağımsız olduklarını varsayıyoruz. Açık ki popülasyon geniş.

$H_0: \pi_1 - \pi_2 = 0$ (ki burada π_1 şiddet suçu işleyen kusurlu gene sahip tacize uğramamış erkeklerin orantısı ve π_2 şiddet suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısıdır)

$H_1: \pi_1 - \pi_2 > 0$ (şiddet suçu işleyen normal gene sahip tacize uğramış erkeklerin orantısı, şiddet

suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısından daha büyüktür)

$$p_1 = \frac{32}{150} = 0.213,\ p_2 = \frac{48}{249} = 0.193 \text{ ve } p = \frac{32+48}{150+249} = 0.2005$$

$$s_{p_1-p_2} = \sqrt{0.2005\,(0.7995)\left(\frac{1}{150}+\frac{1}{249}\right)} = 0.0414$$

Gözlenen fark $0.213 - 0.193 = 0.02$ ve bu farka tekabül eden z puanı,

$$\frac{0.02-0}{0.0414} = 0.483$$

ve P-değeri $1 - 0.686 = 0.314$. Böylesine büyük bir P-değeri ile H_0 reddedemiyoruz. Şu neticeye varıyoruz ki şiddet suçu işleyen kusurlu gene sahip tacize uğramamış erkeklerin orantısının, şiddet suçu işleyen normal gene sahip tacize uğramamış erkeklerin orantısından daha büyük olduğunu gösteren kanıt yoktur.

d. Yukarıdaki sonuçlar arasında bir çelişki yok mu? Açıkla.

Cevap: Herhangi bir çelişki yok. Şu olanaklı ki mevcudiyeti kanıtlanan genel bir ilişki için, kategorilerin bir alt kümesinde anlamlı bir kanıt bulunamayabilir (Simpson paradoksu).

Bağımsızlık testinde, her bir gözlem ünitesi üzerinde iki değişkenin gözlendiğine dikkat et.

Ki-kare ile Orantıların Homojenliği Testi

Ki-kare ile uyum iyiliği testlerinde tek bir değişkeni ele aldık ve bu değişkenin bir örneğinin dağılımını bir popülasyon modeli ile karşılaştırdık. Ki-kare ile bağımsızlık testlerinde iki değişken bakımından sınıflandırılan tek bir örneği ele aldık. Ki-kare prosedürleri aynı zamanda tek bir değişkenin iki veya daha çok popülasyondan çekilmiş örneklerini karşılaştırmakta da kullanılabilir. Önemli olan, örneklerin *basit rastgele örnekler* olmaları, biribirlerinden *bağımsızca* çekilmeleri, popülasyonların hacimlerinin örnek hacimlerine kıyasla çok daha geniş olmaları ve çapraz tablonun her hücresi için beklenen değerlerin en az 5 olmasıdır.

Orantıların homojenlik testi için kullanacağımız çapraz tabloda, her örnek için bir satır ihdas edileceği için, serbestlik derecesi de satır sayısının 1 eksiğine eşit olacaktır.

Misal 5 Atatürk Üniversitesi ülkenin en büyük üniversitelerinden birisidir. Bir bilimsel-araştırma projesi olarak, istatistik dersi öğrencileri, üniversite mensupları arasında tütün içiciliği orantılarının meslek gruplarına göre farklılık gösterip göstermediğini belirlemek istediler. Öğrenme-çalışması kapsamında, 100 öğretim üyesiden, 45 araştırma görevlisinden, 55 memurdan ve 60 işçiden oluşan bağımsız basit rastgele örnekler aldılar. Bu örneklerde tütün içicilerinin sayıları, sırasıyla şöyle tecelli etti: 18, 22, 11 ve 19. Üniversite mensupları arasında tütün içiciliği orantılarının meslek gruplarına göre farklılık gösterdiğine dair kanıt var mı?

Cevap:

H_0: Meslek gruplarında tütün içicilerinin orantısı aynıdır.
H_1: En az iki meslek grubunda tütün içicilerinin orantısı farklıdır.

Gözlenen sayılar şöyledir:

	Tütün içen	Tütün içmeyen
Öğretim üyeleri	18	82
Araştırma görevlileri	11	34
Memurlar	19	36
İşçiler	22	38

Tam olarak bir önceki kısımda yaptığımız gibi, çapraz tablonun her bir hücresi için beklenen değeri mütekabil satır ve sütun toplamlarının çarpımını genel toplama bölerek hesaplarız. Böyle yapınca, şu beklenen değerlere ulaşırız:

	Tütün içen	Tütün içmeyen	Toplam
Öğretim üyeleri	26.9	73.1	100
Araştırma görevlileri	12.1	32.9	45
Memurlar	14.8	40.2	55
İşçiler	16.2	43.8	60
Toplam	190	70	260

Tüm beklenen değerlerin 5 ten büyük olduğunu görüyoruz. Öyleyse, ki-kare değerini hesaplayabiliriz,

$$\chi^2 = \sum \frac{(\text{Gözlenen} - \text{Beklenen})^2}{\text{Beklenen}} = \frac{(18 - 26.9)^2}{26.9} + \cdots + \frac{(38 - 43.8)^2}{43.8} = 8.64$$

Serbestlik derecesi, $4 - 1 = 3$ olduğuna göre, P-değerini, Tblo C de,

$$0.05 > P(\chi^2 > 8.64) > 0.025$$

olarak buluruz. Bu kadar küçük bir P-değeri ile H_0 ı %5 ten küçük bir α-riski ile reddedebiliriz. Yani, tütün içenlerin orantılarının meslek gruplarında aynı olmadığını gösteren kanıtın mevcut olduğu neticesine varırız.

Bağımsızlık testi ve homojenlik testi arasındaki ayrım kafa karıştırıcı olabilir. Önümüze sadece bir çapraz tablo koyulduğu zaman, hangi testin icra edileceğini bilemeyiz. Ayrımı yaratan öğrenme-çalışmasının *dizaynıdır*. Bir değişkenin, farklı popülasyonlardaki dağılımlarını karşılaştırmak için, iki veya daha çok sayıdaki popülasyonların her birinden örnekler mi çektik? Eğer böyle ise, bir homojenlik testi yapıyoruz demektir. İki değişken arasındaki çağrışımı görmek için, tek bir popülasyondan çektiğimiz bir örneği, iki değişkene göre çapraz-sınıflandırmaya mı tabii tuttuk. Eğer böyle ise, bir bağımsızlık testi yapıyoruz demektir.

Mesela, eğer kürtaja dair kanaatlerini (lehinde, aleyhinde veya kararsız) belirlemek üzere, Müslümanları, Hıristiyanları ve Yahudileri ayrı ayrı örneklersek, o zaman, kürtaja dair kanaatlerin dağılımının Müslümanlarda, Hıristiyanlarda ve Yahudilerde aynı olduğunu ifade eden bir sıfır hipotezi ile bir homojenlik testi yaparız. Mamafih, eğer popülasyonun genelinden bir örnek çektikten sonra, deneklerin dinsel tercihlerini ve kürtaja dair kanaatlerini kaydedersek (bir örnek, iki değişken), o zaman, kürtaja dair kanaatin dinsel tercihten bağımsız olduğunu ifade eden bir sıfır hipotezi ile bir bağımsızlık testi yaparız.

Homojenlik testi iki örnek orantısı arasındaki farkın z-testini genişletir. Eğer çıkımın yalnızca iki kategorisi varsa, iki örnek orantısı arasındaki farkın z-testini kullanırız. Kategorilerin sayısı ikiden fazlaysa, homojenlik testini kullanırız.

Nihayetinde, şunu bilmende yarar var ki kategorik veri için χ^2 dağılımını kullanıyoruz, ama, χ^2 dağılımı *sürekli* bir dağılımdır ve χ^2 dağılımını sayma sayılarına uygulamak ancak bir yaklaşımdır.

En Küçük Kareler Doğrusunun Eğimi için Anlamlılık Testi

Regresyon doğrusunun hakiki eğimi için bir güven aralığı bulmak dışında, eğimin belli bir değeri için bir anlamlılık testi de icra edilebilir. Bu durumda, sıkça kullanılan sıfır hipotezi,

$$H_0: \beta = 0;$$

yani, iki değişken arasında doğrusal bir ilişkinin mevcut olmadığıdır.

En küçük kareler doğrusunun eğimi için çıkarımda bulunmanın varsayımları şunlardır: (1) Örnek rastgele seçilmiş olmalıdır; (2) Saçılım grafiği yaklaşık doğrusal olmalıdır; (3) Kalıntıların grafiği

belirgin bir biçim göstermemelidir; (4) Kalıntıların dağılımı yaklaşık normal olmalıdır. (Dördüncü varsayımı kontrol etmek için kalıntıların bir histogramı, nokta grafiği, dal-yaprak grafiği veya normal olasılık grafiği çizdirilebilir.)

Eğim için güven aralığını kestirirken yaptığımız gibi, $(n-2)$ serbestlik dereceli t-dağılımını kullanacağız. Örnek regresyon doğrusunun eğimi b nin standart hatasını da yine aynı şekilde hesaplayacağız,

$$s_b = \frac{\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}}}{\sqrt{\sum(x_i - \bar{x})^2}} = \frac{s_e}{s_x\sqrt{n-1}}$$

Küçük bir P-değeri, bize, eğer iki değişken arasında bir doğrusal ilişki olmasaydı, böyle bir rastgele örneği çekmenin hayli olabilirliksiz olacağını söyler. Mamafih, bir çağrışımın varlığına dair kuvvetli kanıt, çağrışımın güçlü olduğu manasına gelmez.

Misal 6 Aşağıdaki tabloda, bir kavak ağacından rastgele örneklenen 10 yaprağın kuru ve taze genişliklerine ait ölçümler mm cinsinden verilmektedir.

Kuru genişlik (X):	125	133	108	128	115	135	125	117	130	121
Taze genişlik (Y):	133	134	112	139	123	142	140	129	139	126

Şu istatistikler hesaplanmıştır: $\bar{x} = 123.7$, $s_x = 8.499$; $\bar{y} = 131.7$, $s_y = 9.381$; $r = 0.9$.

a. Kavak ağacı yapraklarının kuru genişlikleri ve taze genişlikleri arasında pozitif eğimli bir doğrusal ilişkinin mevcudiyetine dair kanıt var mı?

Cevap:

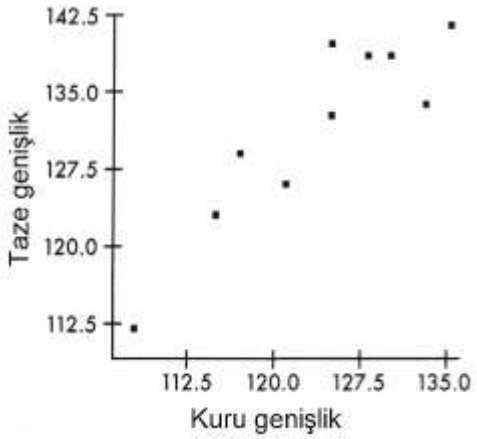
$$b = r\frac{s_y}{s_x} = 0.9\frac{9.381}{8.499} = 0.993$$

$$a = \bar{y} - b\bar{x} = 131.7 - 0.993(123.7) = 8.76$$

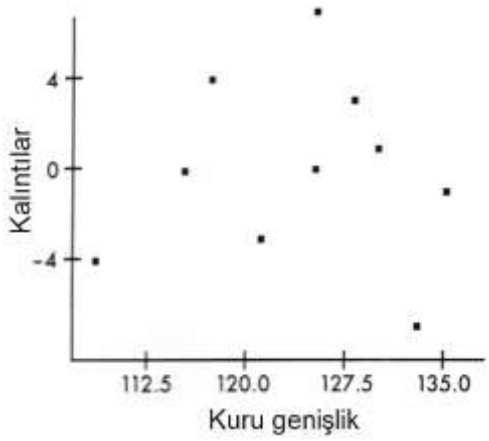
$$\hat{y} = 8.76 + 0.993x$$

x_i:	125	133	108	128	115	135	125	117	130	121
y_i:	133	134	112	139	123	142	140	129	139	126
$\hat{y}_i$:	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9
e_i:	0.0	-6.9	-4.1	3.0	-0.1	-0.9	7.0	4.0	1.0	-3.0

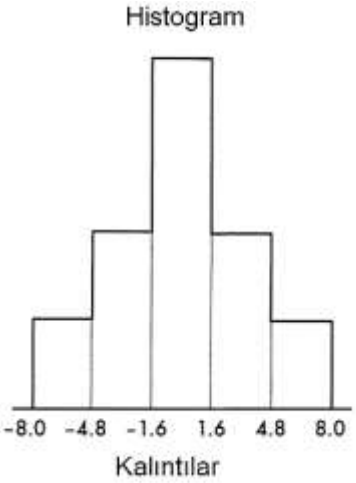
Öncelikle, kullanacağımız istatistiksel prosedürün varsayımlarını kontrol etmeliyiz: (1) Verinin yaprakların bir rastgele örneğinden elde edildiği bize söyleniyor. (2) Saçılım grafiği yaklaşık doğrusal.

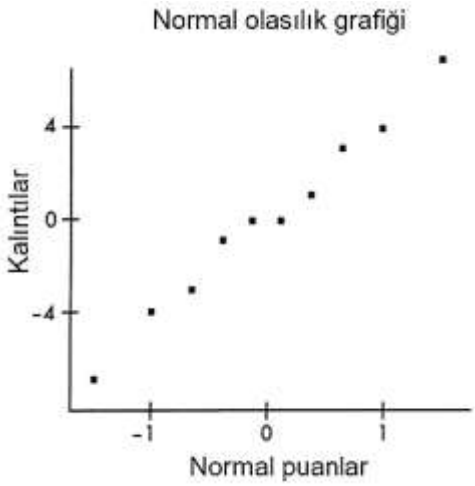


(3) Kalıntı grafiği belirgin bir biçim göstermiyor.



(4) Kalıntıların dağılımı yaklaşık normal.





Şimdi, regresyon doğrusunun eğimi için anlamlılık testine geçebiliriz:

$$H_0: \beta = 0 \qquad \text{ve} \qquad H_1: \beta > 0$$

Kalıntıların standart sapması,

$$s_e = \sqrt{\frac{\sum e_i^2}{n-2}} = 4.33$$

En küçük kareler doğrusunun eğimi b nin standart hatası,

$$s_b = \frac{s_e}{s_x\sqrt{n-1}} = \frac{4.33}{8.499\sqrt{10-1}} = 0.1698$$

Regresyon doğrusunun örnekten hesaplanan eğimine tekabül eden t-puanı,

$$\frac{0.993 - 0}{0.1698} = 5.85$$

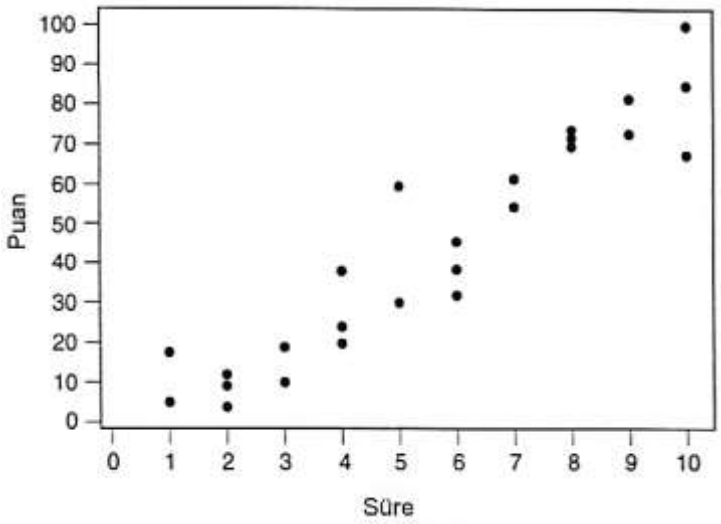
Alternatif hipotez tek kuyruklu olduğu için, $10 - 2 = 8$ lik bir serbestlik derecesi ile P-değerini Tablo C de $P(t > 5.85) < 0.0005$ olarak okuruz. Böylesine küçük bir P-değeri ile, kuru ve taze yaprak genişlikleri arasındaki pozitif eğime sahip doğrusal ilişkinin mevcudiyetine dair çok kuvvetli kanıt olduğu neticesine varırız.

b. Kestirilen en küçük kareler doğrusunun problem bağlamında interpretasyonunu yapabilir misin?

Cevap: Eğimi yaklaşık 1 ve y-kesimi 8.76 olan regresyon doğrusu şunu belirtiyor ki taze yaprak genişliği, kuru yaprak genişliğinden aritmetik ortalama 8.76 mm daha fazladır. Yani, ister daha küçük

olsun, ister daha büyük olsun kuru yaprakların taze genişlikleri aritmetik ortalamada aynı miktarda artmaktadır.

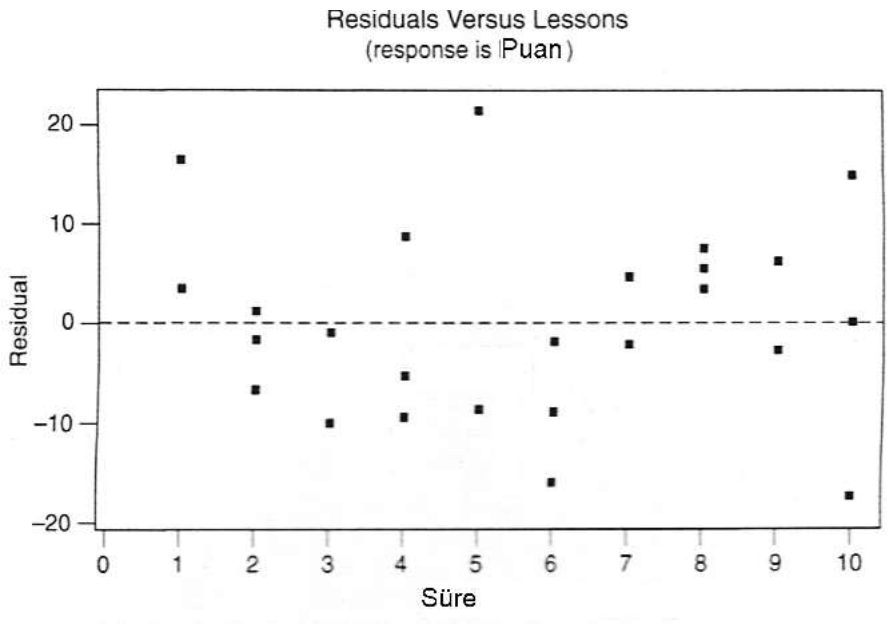
Misal 7 İstatistik dersi vize sınavından sonra, rastgele örneklenen 25 öğrenciye sınava sınıf dışında toplamda kaç saat ders çalıştıkları soruluyor. Öğrencilerin ders çalışma süreleri ve sınav puanlarının saçılım grafiği şudur:

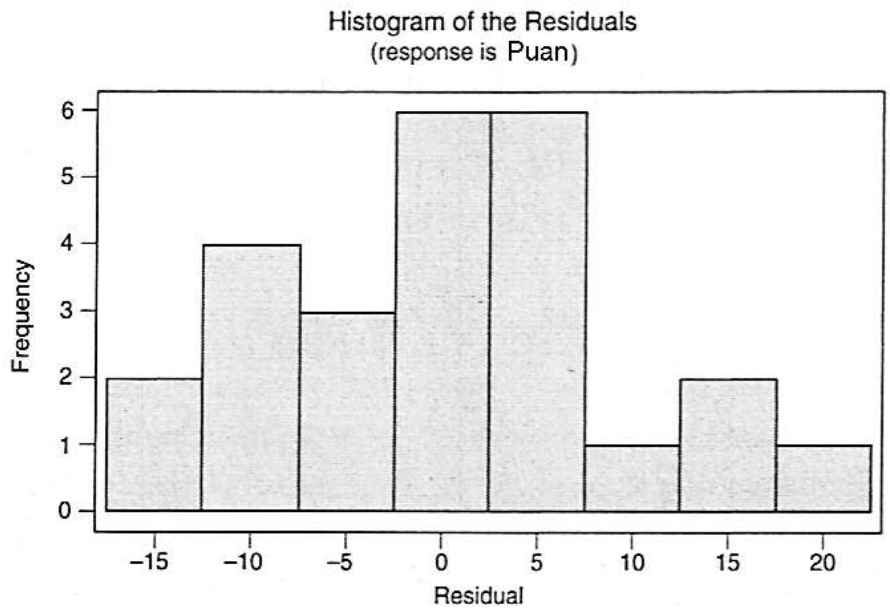


Bir istatistiksel yazılım paketi vasıtasıyla icra edilen regresyon analizine ait bilgisayar çıktısı kısmen şöyledir:

Dependent variable: Puan				
Predictor	Coef	SE Coef	T	P
Constant	-7.718	4.272	-1.81	0.084
Süre	9.2854	0.6789	13.68	0.000
S = 9.744	R-Sq = 89.1%			

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	17761	17761	187.06	0.000
Residual Error	23	2184	95		
Total	24	19945			





a. Ders çalışma süresinin (saat) bir fonksiyonu olarak sınav puanını tahmin eden regresyon denklemi nedir?

Cevap:

$$\widehat{\text{Puan}} = -7.718 + 9.2854(\text{Süre})$$

b. Problem bağlamında, regresyon doğrusunun eğimini ne ifade ediyor?

Cevap: Regresyon doğrusunun eğimi 9.2854, şu manaya geliyor ki her ilave 1 saatlik ders çalışma, öğrencinin sınav puanını aritmetik ortalama 9.2854 puan artırmaktadır.

c. Bu öğrenme-çalışması bağlamında, R-Sq (yani R^2) nin manası nedir?

Cevap: R-Sq = 89.1% şu manaya gelir ki sınav puanındaki değişimin %89.1 i ders çalışma süresinin değişimi ile açıklanabilmektedir.

d. Korelasyon katsayısının değeri nedir?

Cevap:

$$+r = +\sqrt{R^2} = +\sqrt{0.891} = 0.949$$

ki burada işaret pozitif, çünkü eğim pozitiftir.

e. Regresyon doğrusunun eğimi için bir anlamlılık testi icra et.

Cevap: Regresyon doğrusunun eğiminin anlamlılık testi için,

$$H_0: \beta = 0 \qquad \text{ve} \qquad H_1: \beta \neq 0$$

Varsayımlar: Verinin rastgele bir örnekten geldiği, bize söyleniyor, saçılım grafiği yaklaşık doğrusal görünüyor, kalıntıların grafiğinde belirgin bir biçim görünmüyor ve kalıntıların histogramı yaklaşık normal görünüyor.

$$t = \frac{9.2854 - 0}{0.6789} = 13.68$$

ve $sd = 25 - 2 = 23$ olduğuna göre, Tablo C de, kuyruk olasılığı $P = 0.000$ olarak okunur. (Ya da, P-değerini, basitçe, bilgisayar çıktısından okuyabiliriz.) Böylesine küçük bir P-değeri ile, H_0 ı reddetmek için çok kuvvetli kanıt mevcut; ve şu neticeye varılır ki sınav puanı ve ders çalışma süresi arasında betimlenen ilişki istatistiksel olarak anlamlıdır.

Mini Proje 15

Bir madeni parayı ve bir zarı 10 kere, 20 kere ve 30 kere beraber at ve her atışta paranın ve zarın çıkımını kaydet. Sonuçları aşağıdaki çapraz tablolara sırasıyla yerleştir. Beraber atılan parada ve zarda gelen sayıların bağımsızlığını test et. Örnek hacminin test sonuçlarına etkisini tartış.

H_0: .....
H_1: .....

10 ATIŞ		ZAR		
		TEK	ÇİFT	Toplam
PARA	YAZI			
	TURA			
	Toplam			

$\chi^2 =$ .....
P-değeri = ... Netice: .....

20 ATIŞ		ZAR		
		TEK	ÇİFT	Toplam
PARA	YAZI			
	TURA			
	Toplam			

$\chi^2 =$ .....
P-değeri = ... Netice: .....

30 ATIŞ		ZAR		
		TEK	ÇİFT	Toplam
PARA	YAZI			
	TURA			
	Toplam			

$\chi^2 =$ .....
P-değeri = ... Netice: .....

Tartışma:.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

BÖLÜM 16 | İKİDEN ÇOK ARİTMETİK ORTALAMAYI KARŞILAŞTIRMAK

Giriş

Bu bölümde, yine, bir nicel değişken ve bir nitel değişken arasındaki ilişkiyi öğrenmeye-çalışıyoruz. Bir başka deyişle, iki veya daha çok sayıdaki popülasyon aritmetik ortalamasının karşılaştırılması durumunu tartışacağız. Cevabını aradığımız soru şu: İlgilendiğimiz nicel değişkenin ilgilendiğimiz nitel değişkenin kategorileri bakımından oluşturduğu popülasyonlar biribirinden farklı mı? Mevcut popülasyonların her birinden birer bağımsız rastgele örnek seçeceğiz ve her birinin aritmetik ortalamasını hesaplayacağız. Örnek aritmetik ortalamaları arasında gözlenen farkların anlamlı olduğunu ne zaman söyleyebiliriz? Hakiki popülasyon aritmetik ortalamalarının farklı olduğuna dair yetersiz kanıt olduğunu ne zaman söylemeliyiz?

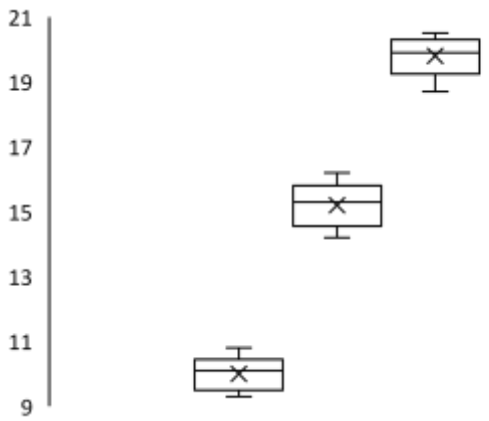
Bunu söylemek için bir metot, *t*-dağılımını kullanmak suretiyle, tüm örneklerin aritmetik ortalamalarını ikişerli-ikişerli biribirleriyle karşılaştırmak olabilirdi. Mamafih, $H_0: \mu_1 = \mu_2 = \mu_3 = \cdots$ hakkında bir karar verebilmek için bu ikişerli karşılaştırma sonuçlarını nasıl biraraya getirebileceğimizi kimse bize söyleyemiyor.

Örneklerin çekildiği popülasyonların aritmetik ortalamaları arasında en az ikisinin biribirinden farklı olup olmadığını (yani, aritmetik ortalamaların tümünün aynı olmadığını) tespit etmek için kullanılan bir metot, verinin değişkenliğinin örneklerin *arasındaki* ve örneklerin *içindeki* davranışına bakmaktır.

Örneklerin Arasındaki ve İçindeki Varyans

Üç ayrı popülasyondan şu örneklerin çekildiğini farzedelim:

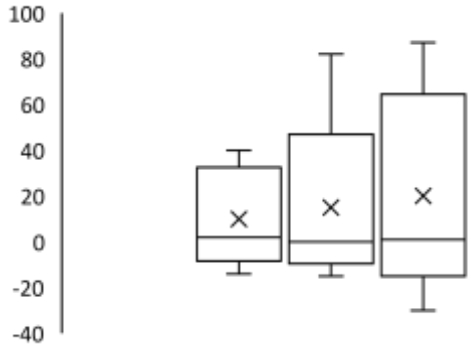
I. {10.1, 10.8, 9.7, 9.3, 10.1}; ve aritmetik ortalaması 10
II. {14.2, 14.9, 15.3, 15.4, 16.2} ve aritmetik ortalaması 15
III. {20.5, 19.8, 19.9, 20.1, 18,7} ve aritmetik ortalaması 20



Sezgisel olarak, üç orjinal popülasyonun aritmetik ortalamalarının aynı olmadığı neticesine varmak için yeterli kanıta sahip olduğumuzu söyleyebiliriz.

Peki, çekilen örnekler şunlar olsaydı:

I. {2, 25, -14, -3, 40} ve aritmetik ortalaması 10
II. {-4, 82, 0, -15, 12} ve aritmetik ortalaması 15
III. {42, -30, 87, 0, 1} ve aritmetik ortalaması 20



Şimdi, üç orjinal popülasyonun aritmetik ortalamalarının aynı olmadığı neticesine varmak için yeterli kanıta sahip olduğumuzu söyleyebilir miyiz? Hayır, söyleyebilmeyiz. Çünkü, önceki durumda, örnek aritmetik ortalamalarının *aralarındaki* farklar, her bir örneğin kendi *içindeki* değişim (veya varyasyon) ile karşılaştırıldığı zaman, daha büyüktü. Oysa, şimdiki durumda, örnek aritmetik ortalamalarının *aralarındaki* farklar, her bir örneğin kendi *içindeki* değişim (veya varyasyon) ile karşılaştırıldığı zaman, çok daha küçüktür.

Örnek aritmetik ortalamaları *arasında* gözlenen farklar, her bir örneğin kendi içindeki farklara kıyasla daha büyük olduğu zaman, daha anlamlıdır. Bunu bir oran formunda şöyle ifade ederiz:

$$\frac{\text{arasındaki farklar}}{\text{içindeki farklar}}$$

Bu oran ne kadar büyükse, örneklerin içinden çekildikleri popülasyonların aritmetik ortalamalarının farklı olduğuna (yani, aynı olmadıklarına) dair o kadar fazla kanıt vardır. Bu oranın payına ve paydasına neleri koymalıyız? Farkları ölçmek için mantıklı bir seçim *varyanslardır*. Bu sebeple, bir "*arasındaki varyans*" ve bir "*içindeki varyans*" tanımlayarak, hesaplamak ve daha sonra bu iki varyansın bir oranını teşkil etmek gerekir. Bu oran *F-oranı* olarak adlandırılır ve örneklerin çekildiği popülasyonların aynı olduğunu iddia eden sıfır hipotezinin doğru varsayılması halinde, bu oranın 1 e eşit olması beklenir. Ancak, bu derste şimdiye kadar öğrendiklerimize istinaden biliyoruz ki rastgele örneklerden hesaplanacak bu oran (ki bir örnek istatistiğidir kendisi), tüm örnekler aynı popülasyondan (veya özdeş popülasyonlardan) çekilmiş olsalar dahi, bir değişkenliğe sahip olacaktır. Yani, bir örnekleme dağılımına sahip olacaktır.

Varyansların oranının istatistiksel anlamlılığını test etmek için *F*-dağılımını kullanırız (*F*-oranının örnekleme dağılımı olarak). *F*-dağılımı bir popülasyonun varyansına ait bağımsız iki kestirimin oranına ait bir olasılık dağılımıdır. Hesaplayacağımız *F*-oranının *pay ve payda serbestlik dereceleri* için bu *F*-dağılımında, seçtiğimiz α-riskinden daha küçük bir *P*-değerine sahip olduğumuz zaman, örneklerin tümünün aynı popülasyondan çekildiğini iddia eden sıfır hipotezini reddetmek için yeterli kanıtın mevcut olduğu neticesine varabiliriz.

Misal 1 Bir emlak ofisinde üç farklı tipte mukavele ile çalışan 15 satış temsilcinin aylık satış rakamları şunlardır:

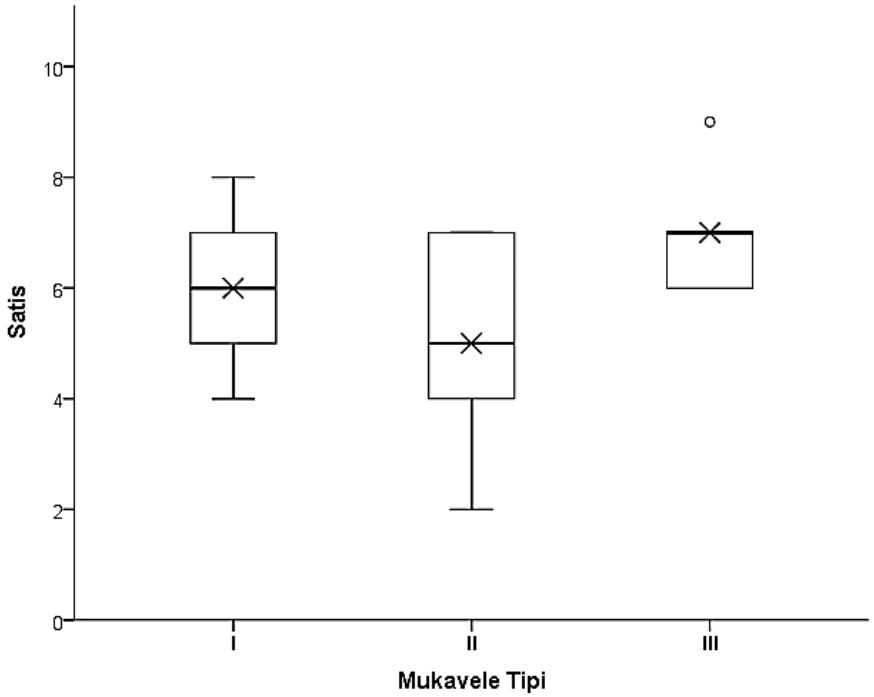
Sabit ücret (I):	4, 7, 8, 5, 6
Sabit komisyon (II):	7, 2, 4, 5, 7
Ücret + komisyon (III):	7, 9, 6, 6, 7

Üç grubun aritmetik ortalamalarını hesaplayalım,

$$\bar{y}_1 = \frac{30}{5} = 6, \qquad \bar{y}_2 = \frac{25}{5} = 5, \qquad \bar{y}_3 = \frac{35}{5} = 7.$$

Aritmetik ortalamalar arasında gözlenen bu farklar anlamlı mı? Her üç örnekteki gözlem

değerlerinin aynı popülasyondan (veya özdeş popülasyonlardan) geldiğini iddia eden sıfır hipotezini reddetmek için yeterli kanıt mevcut mu?



Cevap:

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1: Popülasyonların aritmetik ortalamalarınnın hepsi eşit değil

Sıfır hipotezinin doğru olduğu varsayımı altında, popülasyon varyansının (σ^2) bir kestirimi olarak örneklerin *içindeki* varyansı ($s^2_{içindeki}$) hesaplamak için her bir örneğin varyansını hesapladıktan sonra onların aritmetik ortalamasını alacağız:

$$s_1^2 = \frac{(4-6)^2 + (7-6)^2 + (8-6)^2 + (5-6)^2 + (6-6)^2}{5-1} = \frac{10}{4} = 2.5$$

$$s_2^2 = \frac{(7-6)^2 + (2-6)^2 + (4-6)^2 + (5-6)^2 + (7-6)^2}{5-1} = \frac{18}{4} = 4.5$$

$$s_3^2 = \frac{(7-6)^2 + (9-6)^2 + (6-6)^2 + (6-6)^2 + (7-6)^2}{5-1} = \frac{6}{4} = 1.5$$

$$s^2_{\text{içindeki}} = \frac{2.5 + 4.5 + 1.5}{3} = 2.83$$

Yine, sıfır hipotezinin doğru olduğu varsayımı altında, popülasyon varyansının (σ^2) bir kestirimi olarak örneklerin *arasındaki* varyansı ($s^2_{\text{arasındaki}}$) hesaplamak için, önce örnek aritmetik ortalamalarının varyansını ($s^2_{\bar{y}}$) hesaplarız, daha sonra, merkezi limit teoreminin bize sunduğu

$$\sigma^2_{\bar{y}} = \frac{\sigma^2}{n}$$

ifadesi gereğince

$$s^2_{\text{arasındaki}} = n s^2_{\bar{y}}$$

işlemini yapabiliriz:

$$\bar{\bar{y}} = \frac{\bar{y}_1 + \bar{y}_2 + \bar{y}_3}{3} = \frac{6+5+7}{3} = 6$$

$$s_{\bar{y}}^2 = \frac{(6-6)^2 + (5-6)^2 + (7-6)^2}{3-1} = \frac{2}{2} = 1$$

$$s_{\text{arasındaki}}^2 = n s_{\bar{y}}^2 = 5(1) = 5$$

Buna göre, F-oranını, yukarıda yapmış olduğumuz tanım gereğince, sıfır hipotezinin doğru olduğu varsayımı altında, popülasyon varyansının (σ^2) iki ayrı kestiriminin oranı olarak hesaplayabiliriz,

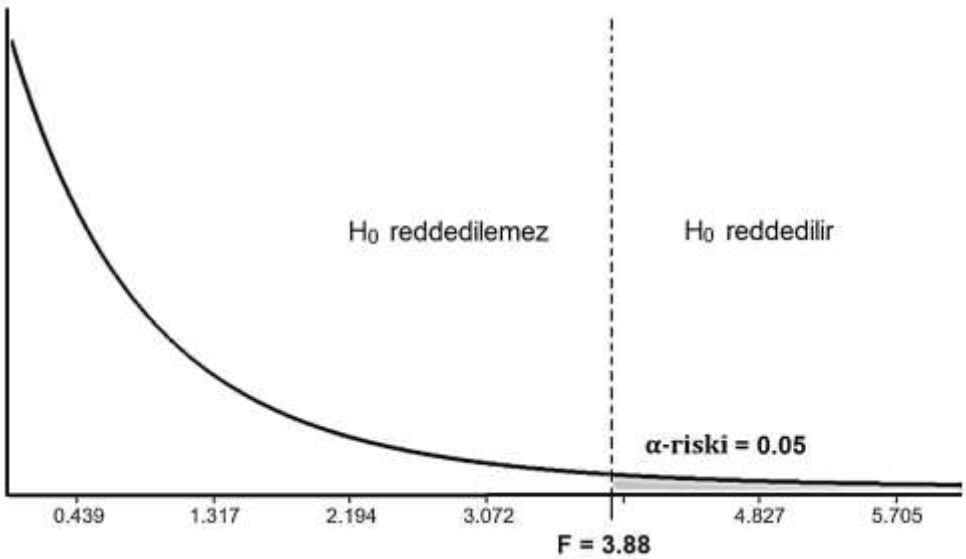
$$F = \frac{s_{\text{arasındaki}}^2}{s_{\text{içindeki}}^2} = \frac{5}{2.83} = 1.76.$$

Hesapladığımız bu F-oranı H_0 ı red edebilmemize yetecek kadar büyük mü? Bu soruyu cevaplayabilmek için, H_0 ın doğru olduğu varsayımı altında, 1.76 veya daha büyük F-oranlarının olasılığını yani P-değerini bilmemiz gerekir. Ancak, Tablo E deki F-dağılımı tablolarından, bir F-oranına ait P-değerini doğrudan okumak olanağına sahip değiliz. Onun yerine, Student t-dağılımını kullanırken yaptığımız gibi, belirli bir α-riski için, oranladığımız varyansların serbestlik derecelerine ilişkin kritik F-puanını buluruz. Eğer hesapladığımız F-oranı, anlamlılık düzeyini sınırlayan kritik F-puanından daha aşırı bir puan olursa, H_0 hipotezini reddederiz. Çünkü bu durumda örnek istatistiğimize ilişkin P-değeri seçmiş olduğumuz α-riskinin altında kalmış olur.

Şimdi, bir serbestlik derecesi paydaki varyans için ve bir serbestlik derecesi de paydadaki varyans için hesaplamalıyız.

$$\text{Payın serbestlik derecesi} = \text{örneklerin sayısı} - 1$$
$$\text{Paydanın serbestlik derecesi} = (\text{örneklerin sayısı})(\text{örnek hacmi} - 1)$$

Buna göre, payın serbestlik derecesini $3 - 1 = 2$ ve paydanın serbestlik derecesini $3(5-1) = 12$ olarak hesaplarız. Eğer anlamlılık düzeyini (yani, α-riskini) 0.05 olarak seçersek, Tablo E de kritik F puanını 3.88 olarak okuruz. Hesapladığımız F-oranı (yani, 1.76) seçtiğimiz α-riskini (yani, 0.05) sınırlayan F-puanından (yani, 3.88) daha küçük olduğu için, H_0 ı ret edebilmemiz için yeterli kanıt yok.



Formal Prosedür

- İkiden daha çok sayıdaki örnek aritmetik ortalamalarını karşılaştırmak için F-dağılımını kullanırken, örneklerin içlerinden çekildiği tüm orjinal popülasyonların **normal dağılımlı** ve **eşit varyanslı** olduğu (bir başka deyişle, normal dağılıma sahip tek bir popülasyondan çekildikleri) varsayılır.
- Örneklerin aritmetik ortalamaları hesaplanır:

$$\bar{y}_1 = \frac{\sum y}{n}, \qquad \bar{y}_2 = \frac{\sum y}{n}, \qquad \cdots, \qquad \bar{y}_k = \frac{\sum y}{n}$$

- Örneklerin varyansları hesaplanır:

$$s_1^2 = \frac{\sum(y-\bar{y}_1)^2}{n-1}, \qquad s_2^2 = \frac{\sum(y-\bar{y}_1)^2}{n-1}, \qquad \cdots, \qquad s_k^2 = \frac{\sum(y-\bar{y}_1)^2}{n-1}$$

- Genel aritmetik ortalama hesaplanır:

$$\bar{\bar{y}} = \frac{\bar{y}_1 + \bar{y}_2 + \cdots + \bar{y}_k}{k}$$

- Örnek aritmetik ortalamalarının varyansı hesaplanır:

$$s_{\bar{y}}^2 = \frac{(\bar{y}_1 - \bar{\bar{y}})^2 + (\bar{y}_2 - \bar{\bar{y}})^2 + \cdots + (\bar{y}_k - \bar{\bar{y}})^2}{k-1}$$

- Örneklerin içindeki varyans hesaplanır:

$$s_{\text{içindeki}}^2 = \frac{s_1^2 + s_2^2 + \cdots + s_k^2}{k}$$

- Örneklerin arasındaki varyans hesaplanır:

$$s_{\text{arasındaki}}^2 = n s_{\bar{y}}^2$$

- F-oranı hesaplanır:

$$\frac{s_{\text{arasındaki}}^2}{s_{\text{içindeki}}^2}$$

- F-oranının payına ve paydasına ait serbestlik dereceleri belirlenir:

$$sd_{\text{pay}} = k - 1 \qquad \text{ve} \qquad sd_{\text{payda}} = k(n-1)$$

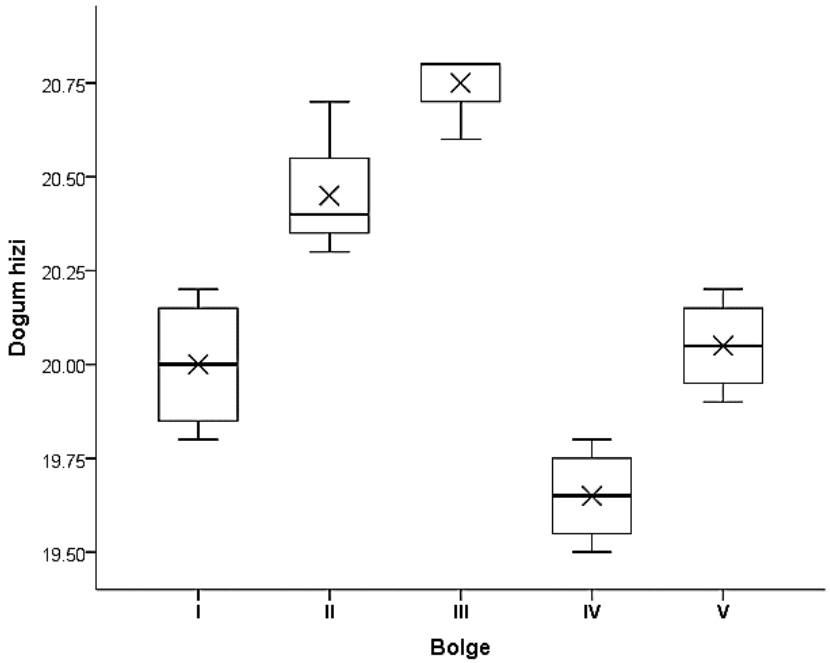
- Anlamlılık düzeyi (α-riski) belirlenir.
- Tablo E den, anlamlılık düzeyine ve serbestlik derecelerine tekabül eden kritik F-puanı bulunur.
- Eğer F-oranı > kritik F-puanı ise, H_0 ı reddetmek için yeterli kanıt vardır. Aksi halde, H_0 reddedilebilmez.

Yukarıda tanımladığımız prosedür *tek-faktörlü* veya *tek-boyutlu varyans analizi* olarak adlandırılır ki 1920'li yıllarda Sir Ronald Aylmer Fisher (1890-1962) tarafından geliştirilmiştir ve F-oranındaki, F-dağılımındaki ve kritik F-puanındaki F'ler onun adına atfen kullanılagelmiştir. Varyans analizi kısaca ANOVA ("Analysis of Variances" ibaresinin baş harfleri) olarak ta yazılır. Farklı veri toplama koşulları için, daha farklı ANOVA teknikleri de mevcuttur.

Misal 2 Doğum hızının, Nepal'in beş ekonomik kalkınma bölgesi arasında farklı olduğu iddiasını test etmek üzere, her bölgeden ratgele seçilen dörder ilçeye ait doğum hızları şöyledir:

Bölge 1: {19.8, 20.1, 20.2, 19.9}
Bölge 2: {20.4, 20.3, 20.4, 20.7}
Bölge 3: {20.6, 20.8, 20.8, 20.8}
Bölge 4: {19.5, 19.7, 19.6, 19.8}
Bölge 5: {20.0, 19.9, 20.2 20.1}

Sıfır hipotezini, $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$, %1 anlamlılık düzeyinde reddetmek için yeterli kanıt var mı?



Cevap: Örnek aritmetik ortalamaları:

$$\bar{y}_1 = \frac{19.8 + 20.1 + 20.2 + 19.9}{4} = 20.00$$

$$\bar{y}_2 = \frac{20.4 + 20.3 + 20.4 + 20.7}{4} = 20.45$$

$$\bar{y}_3 = \frac{20.6 + 20.8 + 20.8 + 20.8}{4} = 20.75$$

$$\bar{y}_4 = \frac{19.5 + 19.7 + 19.6 + 19.8}{4} = 19.65$$

$$\bar{y}_5 = \frac{20.0 + 19.9 + 20.2 + 20.1}{4} = 20.05$$

Örnek varyansları:

$$s_1^2 = \frac{(19.8 - 20.00)^2 + (20.1 - 20.00)^2 + (20.2 - 20.00)^2 + (19.9 - 20.00)^2}{4 - 1} = 0.03333$$

$$s_2^2 = \frac{(20.4 - 20.45)^2 + (20.3 - 20.45)^2 + (20.4 - 20.45)^2 + (20.7 - 20.45)^2}{4 - 1} = 0.03$$

$$s_3^2 = \frac{(20.6 - 20.75)^2 + (20.8 - 20.75)^2 + (20.8 - 20.75)^2 + (20.8 - 20.75)^2}{4 - 1} = 0.01$$

$$s_4^2 = \frac{(19.5-19.65)^2+(19.7-19.65)^2+(19.6-19.65)^2+(19.8-19.65)^2}{4-1} = 0.01667$$

$$s_5^2 = \frac{(20.0-20.05)^2+(19.9-20.05)^2+(20.2-20.05)^2+(20.1-20.05)^2}{4-1} = 0.01667$$

Genel aritmetik ortalama:

$$\bar{\bar{y}} = \frac{20.00+20.45+20.75+19.65+20.05}{5} = 20.18$$

Örnek aritmetik ortalamalarının varyansı:

$$s_{\bar{y}}^2 = \frac{(20.00-20.18)^2+(20.45-20.18)^2+(20.75-20.18)^2+(19.65-20.18)^2+(20.05-20.18)^2}{5-1}$$

$$= \frac{0.728}{4} = 0.182$$

Örneklerin içindeki varyans:

$$s_{\text{içindeki}}^2 = \frac{0.03333+0.03+0.01+0.01667+0.01667}{5} = 0.021333$$

Örneklerin arasındaki varyans:

$$s_{\text{arasındaki}}^2 = 4(0.182) = 0.728$$

F-oranı:

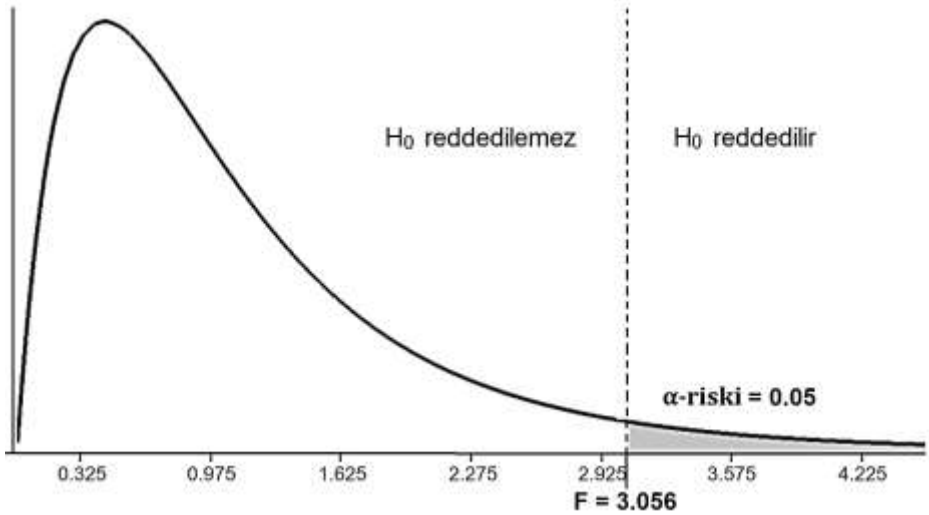
$$F = \frac{0.728}{0.021333} = 34.13$$

F-oranının payına ve paydasına ait serbestlik dereceleri:

$$sd_{\text{pay}} = 5 - 1 = 4$$

$$sd_{\text{payda}} = 5(4 - 1) = 15$$

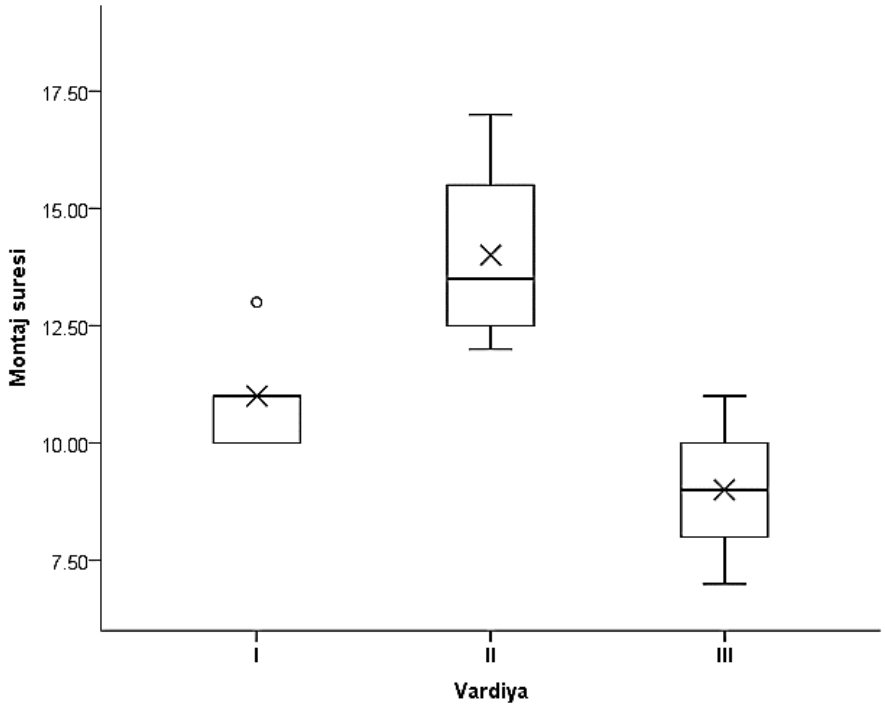
Serbestlik dereceleri 4, 15 ve $\alpha = 0.01$ için kritik F-puanı Tablo E de 4.90 olarak bulunur. Böylece, $34.13 > 4.90$ olduğuna göre, H_0 ı reddetmek için yeterli kanıt var. Doğum hızı bölgelere göre istatistiksel olarak anlamlıca değişmektedir.



Misal 3 Bir makina fabrikası üç vardiyalı üretim yapmaktadır. Üç vardiyanın montaj sürelerini karşılaştırmak üzere aşağıdaki veri toplandı.

Vardiya 1: {13, 10, 10, 11, 11}
Vardiya 2: {13, 14, 17, 12}
Vardiya 3: {10, 11, 7, 9, 8, 9}

Rakamlar gözlenen montaj süreleridir (dakika cinsinden). Üç vardiyanın aritmetik-ortlama montaj sürelerinin farklı olduğu neticesine varmak için yeterli kanıt var mı?



Cevap: Görülüyor ki örnek hacimleri farklı. Bu halde, şimdiye kadar uyguladığımız prosedürü hafifçe modifiye etmek zorundayız. Örneklerin arasındaki ve içindeki varyansları hesaplarken, örnek hacimlerinin farklılığını hesaba katan bir ağırlıklandırma yapacağız.

Örnek aritmetik ortalamalarını önceki gibi hesaplıyoruz:

$$\bar{y}_1 = \frac{13 + 10 + 10 + 11 + 11}{5} = 11 \text{ dakika}$$

$$\bar{y}_2 = \frac{13 + 14 + 17 + 12}{4} = 14 \text{ dakika}$$

$$\bar{y}_3 = \frac{10 + 11 + 7 + 9 + 8 + 9}{6} = 9 \text{ dakika}$$

Örnek varyanslarını önceki gibi hesaplıyoruz:

$$s_1^2 = \frac{(13-11)^2 + (10-11)^2 + (10-11)^2 + (11-11)^2 + (11-11)^2}{5-1} = 1.5 \text{ dakika}^2$$

$$s_2^2 = \frac{(13-14)^2 + (14-14)^2 + (17-14)^2 + (12-14)^2}{4-1} = 4.67 \text{ dakika}^2$$

$$s_3^2 = \frac{(10-9)^2 + (11-9)^2 + (7-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2}{6-1} = 2 \text{ dakika}^2$$

Genel aritmetik ortalamayı örneklerdeki tüm verinin aritmetik ortalamasını alarak hesaplıyoruz, önceki gibi örnek-aritmetik ortalamalarının aritmetik ortalamasını alarak değil:

$$\bar{\bar{y}} = \frac{13 + 10 + 10 + 11 + 11 + 13 + 14 + 17 + 12 + +10 + 11 + 7 + 9 + 8 + 9}{15}$$

$$= 11 \text{ dakika}$$

Örneklerin arasındaki varyansı hesaplamak için, örneklerin her birinin eleman sayılarına (n_1, n_2, n_3) dayanan bir ağırlıklandırma kullanacağız:

$$s^2_{\text{arasındaki}} = \frac{n_1(\bar{y}_1 - \bar{\bar{y}})^2 + n_2(\bar{y}_2 - \bar{\bar{y}})^2 + n_3(\bar{y}_3 - \bar{\bar{y}})^2}{k - 1}$$

$$= \frac{5(11 - 11)^2 + 4(14 - 11)^2 + 6(9 - 11)^2}{3 - 1}$$

$$= \frac{60}{2} = 30 \text{ dakika}^2$$

Örneklerin içindeki varyansı hesaplamak için, yine, örneklerin her birinin eleman sayılarına (n_1, n_2, n_3) dayanan bir ağırlıklandırma kullanacağız:

$$s^2_{\text{içindeki}} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2}{(n_1 + n_2 + n_3) - k}$$

$$= \frac{(5 - 1)1.5 + (4 - 1)4.67 + (6 - 1)2}{(5 + 4 + 6) - 3}$$

$$= \frac{30.01}{12} = 2.5 \text{ dakika}^2$$

F-oranı, yine, örneklerin arasındaki varyansın örneklerin içindeki varyansa oranıdır:

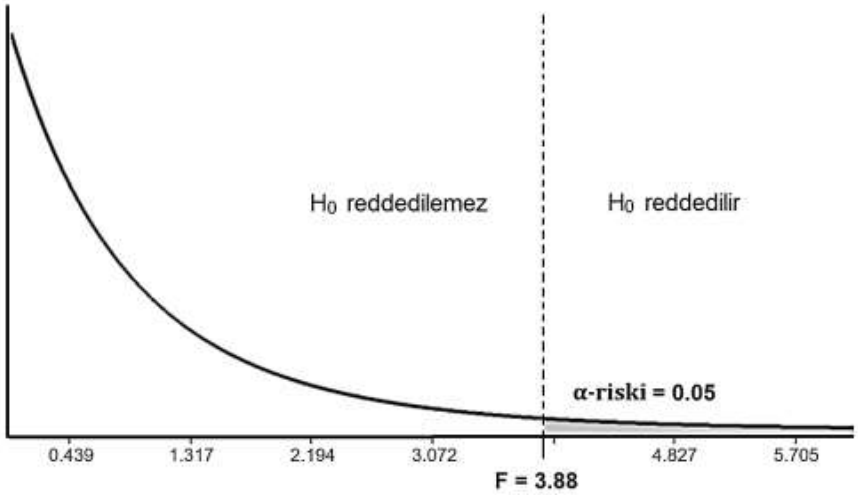
$$F = \frac{30}{2.5} = 12$$

F-oranının payına ve paydasına ait serbestlik dereceleri:

$$sd_{\text{pay}} = 3 - 1 = 2$$

$$sd_{\text{payda}} = (5 - 1) + (4 - 1) + (6 - 1) = 12$$

Serbestlik dereceleri 2, 12 ve $\alpha = 0.05$ için kritik F-puanı Tablo E de 3.88 olarak bulunur. Böylece, $12 > 3.88$ olduğuna göre, H_0 ı reddetmek için yeterli kanıt var. Yani, vardiyaların aritmetik ortalama montaj süreleri arasındaki fark istatistiksel olarak anlamlı.



Alıştırma 1 Eğer örnek hacimleri eşit olsaydı, yani eğer $n_1 = n_2 = n_3$ olsaydı, Misal 3 teki hesaplamaların Formal Prosedür başlığı altındakilere indirgeneceğini göster.

Tamamen Rastgeleleştirilmiş Dizayn

Her deney ünitesinin muamelelerden herhangi birine maruz kalma şansının aynı olduğu dizayna *tamamen rastgeleleştirilmiş deney dizaynı* adı verilir. Bu kurulumda, deney üniteleri muamele gruplarına tamamen rastgele atanır. Bu bölümde, şimdiye kadar, bir nicel değişken ve bir nitel değişken arasındaki ilişkiyi öğrenme-çalışmamızda, tamamen rastgeleleştirilmiş dizayn çerçevesinde toplanan verinin analizini söz konusu ettik. Fakat şunu peşinen söylemek icabediyor ki tamamen rastgeleleştirilmiş dizayn, ancak deney mataeryalinin homojen olduğu hallerde mütenasiptir. Mesela, laboratuar deneylerinde ve sera deneylerinde deney ünitelerinin homojenliğini sağlamak kolaydır ve tamamen rastgeleleştirilmiş dizayn böyle deneyler için çok kullanışlıdır. Öte taraftan, genel olarak, saha deneylerinde, pek çok faktörden dolayı, deney üniteleri arasında büyük varyason (değişim) meydana gelmektedir ki tamamen rastgeleleştirilmiş dizayn tercih edilmemektedir.

Rastgeleleştirilmiş Blok Dizaynı

Deney üniteleri başka bir faktör bakımından heterojen olduğu zaman, *blok* adı verilen kendi içerisinde daha homojen alt-gruplara ayrılırlar. Her bir blok içerisinde muameleler eksiksiz denenir. Böylece, her blokta muamele sayısınca deney ünitesi yer alır. Her blok içerisindeki deney üniteleri muamelelere rastgele atanır. Blokların sayısı muamelelerin tekerrür sayısına eşittir.

Rastgeleleştirilmiş blok dizaynında, deney hatası azaltıldığı için, kesinlik daha yüksektir. Tamamen rastgeleştirilmiş dizayna göre elde edilen enformasyon miktarı daha çoktur. Fakat muamelelerin sayısı arttığı zaman, blok genişliği de artacağı için, blokların içerisinde homojenliğin muhafazası zorlaşır ve böylece bu dizaynın avantajı yok olabilir.

Rastgeleleştirilmiş bloklar dizaynına göre yürütülen bir deneyden elde edilen verinin analizi için istihdam edeceğimiz prosedür *iki-boyutlu varyans analizi* olarak adlandırılır. Bu prosedürün *tek-boyutlu varyans analizi* prosedüründen farkı, muamele aritmetik ortalamalarının varyansına ilaveten, blok aritmetik ortalamalarının varyansının da hesaplanmasıdır. Aynı şekilde bir F-oranı hesaplanarak, anlamlılığı test edilir. Eğer blok aritmetik ortalamaları arasındaki fark anlamlı bulunursa, deney üniteleri arasındaki heterojenliğin etkisi giderimiş olur ve muamele aritmetik ortalamaları arasındaki fark daha kesinlikli olarak gözlenmiş olur. Eğer anlamsız bulunursa, dikkate alınan blok faktörünün verinin değişimine anlamlı bir katkısı olmadığı ve deney ünitelerinin bu faktör bakımından homojen olduğu neticesine varılır. Bu halde, rastgeleleştirilmiş blok dizaynı ve tamamen rastgeleleştirilmiş dizayn aynı sonucu verecektir.

Misal 4. Altı farklı dozda nitrojen uygulamasının yonca bitkisinin verimine etkisini öğrenmeye-çalışmak için bir araştırma dizayn ediliyor. Deney materyali olarak bir nehrin kenarında yer alan 1 dönümlük arazi kullanılıyor. Bu arazi, deney ünitelerini oluşturacak şekilde eşit genişlikte parsellere bölünüyor şöyle ki: Toprak nehre yaklaştıkça daha nemli ve daha verimli olacağından, nehre paralel 5 blok oluşturuluyor. Her blok içerisindeki 6'şar parsel muamelelere rastgele atanıyor. Deneyin bitiminde müteakip veri elde ediliyor. Blokların sayısı 5 ve nitrojen muameleleri A, B, C, D, E ve F ile temsil ediliyor. Parsellerden elde edilen verim (kg) o parsele uygulanan muameleyi temsil eden harf ile beraber veriliyor.

Blok 1	D (17 kg)	C (12 kg)	F (70 kg)	B (6 kg)	A (20 kg)	E (28 kg)
Blok 2	B (12 kg)	C (15 kg)	E (26 kg)	A (26 kg)	D (10 kg)	F (62 kg)

Blok 3	E (23 kg)	A (30 kg)	C (16 kg)	D (20 kg)	F (56 kg)	B (10 kg)
Blok 4	A (28 kg)	F (64 kg)	B (9 kg)	D (23 kg)	E (33 kg)	C (14 kg)
Blok 5	F (75 kg)	C (14 kg)	D (20 kg)	B (7 kg)	E (30 kg)	A (23 kg)

Veri analizi: Veri analizi için yukarıdaki tabloyu aşağıdaki şekilde yeniden tertipliyoruz.

	Muameleler					
	A	B	C	D	E	F
Blok 1	20	9	12	17	28	70
Blok 2	26	12	15	10	26	62
Blok 3	30	10	16	20	23	56
Blok 4	28	9	14	23	35	64
Blok 5	23	7	14	20	30	75

Muamele aritmetik ortalamaları:

$$\bar{y}_{\mathrm{A}} = \frac{20 + 26 + 30 + 28 + 23}{5} = 25.4 \text{ kg}$$

$$\bar{y}_{\mathrm{B}} = \frac{9 + 12 + 10 + 9 + 7}{5} = 9.4 \text{ kg}$$

$$\bar{y}_{\mathrm{C}} = \frac{12 + 15 + 16 + 14 + 14}{5} = 14.2 \text{ kg}$$

$$\bar{y}_{\mathrm{D}} = \frac{17 + 10 + 20 + 23 + 20}{5} = 18.0 \text{ kg}$$

$$\bar{y}_{\mathrm{E}} = \frac{28 + 26 + 23 + 35 + 30}{5} = 28.4 \text{ kg}$$

$$\bar{y}_{\mathrm{F}} = \frac{70 + 62 + 56 + 64 + 75}{5} = 65.4 \text{ kg}$$

Blok aritmetik ortalamaları:

$$\bar{y}_1 = \frac{20 + 9 + 12 + 17 + 28 + 70}{6} = 26.0 \text{ kg}$$

$$\bar{y}_2 = \frac{26 + 12 + 15 + 10 + 26 + 62}{6} = 25.2 \text{ kg}$$

$$\bar{y}_3 = \frac{30 + 10 + 16 + 20 + 23 + 56}{6} = 25.8 \text{ kg}$$

$$\bar{y}_4 = \frac{28 + 9 + 14 + 23 + 35 + 64}{6} = 28.8 \text{ kg}$$

$$\bar{y}_5 = \frac{23 + 7 + 14 + 20 + 30 + 75}{6} = 28.2 \text{ kg}$$

Genel aritmetik ortalama:

$$\bar{\bar{y}} = \frac{25.4 + 9.4 + 14.2 + 18 + 28.4 + 65.4}{6} = 26.8 \text{ kg}$$

Muameleler arasındaki varyans:

$$s^2_{\text{muamelelerarası}}$$
$$= 5\left[\frac{(25.4-26.8)^2+(9.4-26.8)^2+(14.2-26.8)^2+(18-26.8)^2+(28.4-26.8)^2+(65.4-26.8)^2}{6-1}\right]$$
$$= 5\left[\frac{2033.4}{5}\right] = 2033.4\ \text{kg}^2$$

Bloklar arasındaki varyans:

$$s^2_{\text{bloklararası}}$$
$$= 6\left[\frac{(26-26.8)^2+(25.2-26.8)^2+(25.8-26.8)^2+(28.8-26.8)^2+(28.2-26.8)^2}{5-1}\right]$$
$$= 6\left[\frac{10.2}{4}\right] = 15.4\ \text{kg}^2$$

Deney dizaynımıza, muamele faktörüne ilaveten, bir blok faktörü eklediğimizden ötürü, örneklerin içindeki varyansı doğrudan hesaplamak yerine, şimdi önce genel varyansı hesaplayacağız:

$$s^2_{\text{genel}}$$
$$= \frac{(20-26.8)^2+\cdots+(9-26.8)^2+\cdots+(12-26.8)^2+\cdots+(17-26.8)^2+\cdots+(28-26.8)^2+\cdots+(70-26.8)^2+\cdots}{6(5)-1}$$
$$= \frac{10646.8}{29} = 367.1\ \text{kg}^2$$

Daha sonra, genel varyansı, muameleler arasındaki varyansı ve bloklar arasındaki varyansı kullanmak suretiyle, örneklerin içindeki varyansı dolaylı olarak hesaplayacağız:

$$s^2_{içindeki} = \frac{29(367.1)-5(2033.4)-4(15.4)}{29-5-4}$$
$$= \frac{10646.8-10167.2-61.5}{20}$$
$$= 20.9\ \text{kg}^2$$

Burada paydada yer alan serbestlik derecesini, yine doğrudan hesaplamak yerine, genel varyansın serbestlik derecesini, muameleler arasındaki varyansın serbestlik derecesini ve bloklar arasındaki varyansın serbestlik derecesini kullanmak suretiyle, dolaylı olarak hesapladık.

Muameleler arasındaki varyansın örneklerin içindeki varyansa oranı:

$$F_{\text{muamelelerarası}} = \frac{2033.4}{20.9} = 97.3$$

Serbestlik dereceleri 5, 20 ve $\alpha = 0.05$ için kritik F-puanı Tablo E de 2.71 olarak bulunur. Böylece, $97.3 > 2.71$ olduğuna göre, H_0 ı reddetmek için yeterli kanıt var. Yani, nitrojen muamelelerin aritmetik ortalama verimleri arasındaki fark istatistiksel olarak anlamlı.

Bloklar arasındaki varyansın örneklerin içindeki varyansa oranı:

$$F_{\text{bloklararası}} = \frac{15.4}{20.9} = 0.74$$

Serbestlik dereceleri 4, 20 ve $\alpha = 0.05$ için kritik F-puanı Tablo E de 2.87 olarak bulunur. Böylece, $0.74 < 2.71$ olduğuna göre, H_0 ı reddetmek için yeterli kanıt yok. Yani, blokların aritmetik ortalama verimleri arasındaki fark istatistiksel olarak anlamlı değil. (Hesaplanan F-oranı 1 den küçük iken, kritik F-puanına tablodan bakmaya lüzum var mıydı?) Blok aritmetik ortalamaları arasındaki fark anlamlı bulunmadığına göre, dikkate alınan blok faktörünün parsellerin veriminin değişimine anlamlı bir katkısı olmadığı ve parsellerin bu faktör bakımından homojen olduğu neticesine varılır. Demek ki parsellerin nehre uzaklığının bitkinin verimiyle bir ilişkisi gözlenememiştir. Bu halde, tamamen rastgeleleştirilmiş dizayn yerine rastgeleleştirilmiş blok dizaynını kullanmak yararlı olmamıştır.

Mini Proje 16

İkiden çok popülasyon aritmetik ortalamasının eşitliğini test ederken (bir başka deyişle, ikiden çok örnek aritmetik ortalamasını karşılaştırırken), örnek hacimleri P-değerini nasıl etkiler? Bunun için, bir senaryoyu canlandırmanı istiyorum. İki aşamalı olacak. Birinci aşamada, dar örnek hacimleri için bir F-testi yapacaksın. İkinci aşamada, daha geniş örnek hacimleri için bir F-testi yapacaksın.

I. Hepsi eşit hacimli $n_i = 5$, aritmetik ortalamaları sırasıyla 3, 4 ve 5 olan ve hepsinin standart sapmaları eşit, $s_i = 1.58$ olan üç örneğe ait suni veriyi simülasyonla üret. Grupların yan-yana kutu grafiklerini çiz. Grup aritmetik ortalamaları arasındaki farkın F-testini yap.

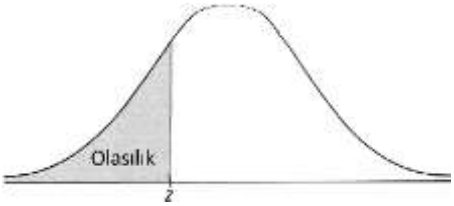
II. Hepsi eşit hacimli $n_i = 12$, aritmetik ortalamaları sırasıyla 3, 4 ve 5 olan ve hepsinin standart sapmaları eşit, $s_i = 1.58$ olan üç örneğe ait suni veriyi simülasyonla üret. Grupların yan-yana kutu grafiklerini çiz. Grup aritmetik ortalamaları arasındaki farkın F-testini yap.

Örnek hacimleri, yaptığın testlerin P-değerlerinin büyüklüklerini nasıl etkiledi? Testlerin sonuçlarının farklı olmasının sebebi nedir? Açıkla.

Tablo A Rastgele sayılar tablosu.

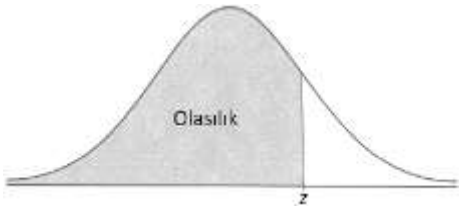
84177	06757	17613	15582	51506	81435	41050	92031	06449	05059
59884	31180	53115	84469	94868	57967	05811	84514	75011	13006
63395	55041	15866	06589	13119	71020	85940	91932	06488	74987
54355	52704	90359	02649	47496	71567	94268	08844	26294	64759
08989	57024	97284	00637	89283	03514	59195	07635	03309	72605
29357	23737	67881	03668	33876	35841	52869	23114	15864	38942
17264	57327	38224	29301	31381	38109	34976	65692	98566	29550
95639	99754	31199	92558	68368	04985	51092	37780	40261	14479
61555	76404	86210	11808	12841	45147	97438	60022	12645	62000
78137	98768	04689	87130	79225	08153	84967	64539	79493	74917
25638	70268	93605	87101	24685	98768	43202	03165	49815	02484
21215	91791	76831	58678	87054	31687	93205	43685	19732	08468
10438	44482	66558	37649	08882	90870	12462	41810	01806	02977
36792	26236	33266	66583	60881	97395	20461	36742	02852	50564
73944	04773	12032	51414	82384	38370	00249	80709	72605	67497
62490	99215	84987	28759	19177	14733	24550	28067	68894	38490
24216	63444	21283	07044	92729	37284	13211	37485	10415	36457
16975	95428	33226	55903	31605	43817	22250	03918	46999	98501
59138	39542	71168	57609	91510	77904	74244	50940	31553	62562
29478	59652	50414	31966	87912	87154	12944	49862	96566	48825
96155	95009	27429	72918	08457	78134	48407	26061	58754	05326
29621	66583	62966	12468	20245	14015	04014	35713	03980	03024
12639	75291	71020	17265	41598	64074	64629	63293	53307	48766
14544	37134	54714	02401	63228	26831	19386	15457	17999	18306
83403	88827	09834	11333	68431	31706	26652	04711	34593	22561
67642	05204	30697	44806	96989	68403	85621	45556	35434	09532
64041	99011	14610	40273	09482	62864	01573	82274	81446	32477
17048	94523	97444	59904	16936	39384	97551	09620	63932	03091
93039	89416	52795	10631	09728	68202	20963	02477	55494	39563
82244	34392	96607	17220	51984	10753	76272	50985	97593	34320
96990	55244	70693	25255	40029	23289	48819	07159	60172	81697
09119	74803	97303	88701	51380	73143	98251	78635	27556	20712
57666	41204	47589	78364	38266	94393	70713	53388	79865	92069
46492	61594	26729	58272	81754	14648	77210	12923	53712	87771
08433	19172	08320	20839	13715	10597	17234	39355	74816	03363
10011	75004	86054	41190	10061	19660	03500	68412	57812	57929
92420	65431	16530	05547	10683	88102	30176	84750	10115	69220
35542	55865	07304	47010	43233	57022	52161	82976	47981	46588
86595	26247	18552	29491	33712	32285	64844	69395	41387	87195
72115	34985	58036	99137	47482	06204	24138	24272	16196	04393

Tablo B Standart Normal Olasılıklar ($z \leq 0$)



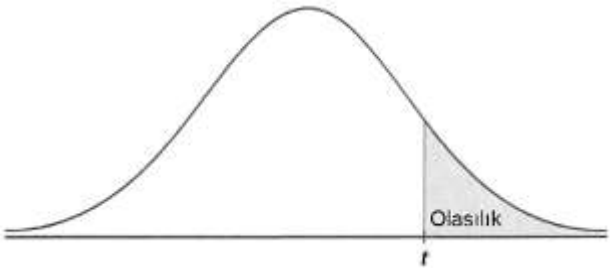
z	**.00**	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Tablo B Standart Normal Olasılıklar ($z \geq 0$)



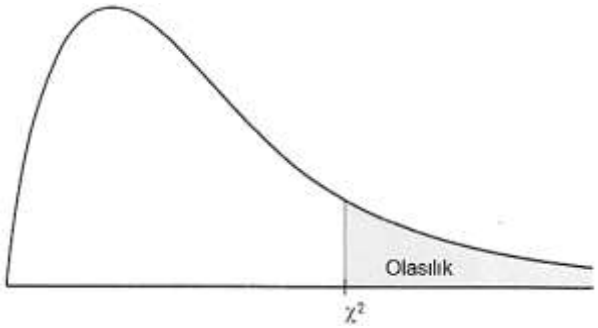
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Tablo C Student *t*-dağılımına ait kritik değerler.



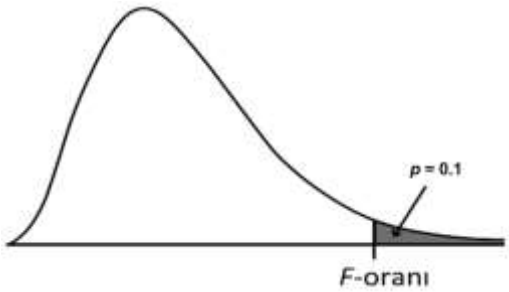
	Kuyruk olasılığı, *p*											
sd	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
•	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Güven düzeyi											

Tablo D Ki-kare dağılımına ait kritik değerler



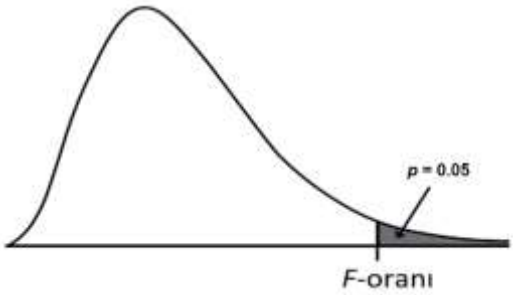
	Kuyruk olasılığı, *p*										
sd	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4

Tablo E F-dağılımına ait kritik değerler ($p = 0.1$)



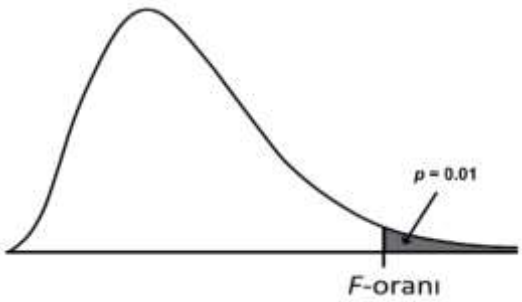
	Pay *s.d.*												
Payda *s.d.*	1	2	3	4	5	6	7	8	12	24	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	60.71	62.00	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.41	9.45	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.22	5.18	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.90	3.83	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.27	3.19	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.90	2.82	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.67	2.58	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.50	2.40	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.38	2.28	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.28	2.18	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.21	2.10	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.15	2.04	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.10	1.98	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.05	1.94	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.02	1.90	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	1.99	1.87	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	1.96	1.84	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	1.93	1.81	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.91	1.79	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.89	1.77	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.87	1.75	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.86	1.73	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.84	1.72	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.83	1.70	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.82	1.69	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.81	1.68	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.80	1.67	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.79	1.66	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.78	1.65	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.77	1.64	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.71	1.57	1.47	1.42	1.38
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.68	1.54	1.42	1.38	1.33
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.66	1.51	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.60	1.45	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.55	1.38	1.24	1.17	1.00

Tablo E F-dağılımına ait kritik değerler ($p = 0.05$)



	Pay *s.d.*												
Payda *s.d.*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**12**	**24**	**60**	**120**	**∞**
1	**161.4**	**199.5**	**215.7**	**224.6**	**230.2**	**234.0**	**236.8**	**238.9**	**243.9**	**249.1**	**252.2**	**253.2**	**254.3**
2	**18.51**	**19.00**	**19.16**	**19.25**	**19.30**	**19.33**	**19.35**	**19.37**	**19.41**	**19.45**	**19.48**	**19.49**	**19.50**
3	**10.13**	**9.55**	**9.28**	**9.12**	**9.01**	**8.94**	**8.89**	**8.85**	**8.74**	**8.64**	**8.57**	**8.55**	**8.53**
4	**7.71**	**6.94**	**6.59**	**6.39**	**6.26**	**6.16**	**6.09**	**6.04**	**5.91**	**5.77**	**5.69**	**5.66**	**5.63**
5	**6.61**	**5.79**	**5.41**	**5.19**	**5.05**	**4.95**	**4.88**	**4.82**	**4.68**	**4.53**	**4.43**	**4.40**	**4.36**
6	**5.99**	**5.14**	**4.76**	**4.53**	**4.39**	**4.28**	**4.21**	**4.15**	**4.00**	**3.84**	**3.74**	**3.70**	**3.67**
7	**5.59**	**4.74**	**4.35**	**4.12**	**3.97**	**3.87**	**3.79**	**3.73**	**3.57**	**3.41**	**3.30**	**3.27**	**3.23**
8	**5.32**	**4.46**	**4.07**	**3.84**	**3.69**	**3.58**	**3.50**	**3.44**	**3.28**	**3.12**	**3.01**	**2.97**	**2.93**
9	**5.12**	**4.26**	**3.86**	**3.63**	**3.48**	**3.37**	**3.29**	**3.23**	**3.07**	**2.90**	**2.79**	**2.75**	**2.71**
10	**4.96**	**4.10**	**3.71**	**3.48**	**3.33**	**3.22**	**3.14**	**3.07**	**2.91**	**2.74**	**2.62**	**2.58**	**2.54**
11	**4.84**	**3.98**	**3.59**	**3.36**	**3.20**	**3.09**	**3.01**	**2.95**	**2.79**	**2.61**	**2.49**	**2.45**	**2.40**
12	**4.75**	**3.89**	**3.49**	**3.26**	**3.11**	**3.00**	**2.91**	**2.85**	**2.69**	**2.51**	**2.38**	**2.34**	**2.30**
13	**4.67**	**3.81**	**3.41**	**3.18**	**3.03**	**2.92**	**2.83**	**2.77**	**2.60**	**2.42**	**2.30**	**2.25**	**2.21**
14	**4.60**	**3.74**	**3.34**	**3.11**	**2.96**	**2.85**	**2.76**	**2.70**	**2.53**	**2.35**	**2.22**	**2.18**	**2.13**
15	**4.54**	**3.68**	**3.29**	**3.06**	**2.90**	**2.79**	**2.71**	**2.64**	**2.48**	**2.29**	**2.16**	**2.11**	**2.07**
16	**4.49**	**3.63**	**3.24**	**3.01**	**2.85**	**2.74**	**2.66**	**2.59**	**2.42**	**2.24**	**2.11**	**2.06**	**2.01**
17	**4.45**	**3.59**	**3.20**	**2.96**	**2.81**	**2.70**	**2.61**	**2.55**	**2.38**	**2.19**	**2.06**	**2.01**	**1.96**
18	**4.41**	**3.55**	**3.16**	**2.93**	**2.77**	**2.66**	**2.58**	**2.51**	**2.34**	**2.15**	**2.02**	**1.97**	**1.92**
19	**4.38**	**3.52**	**3.13**	**2.90**	**2.74**	**2.63**	**2.54**	**2.48**	**2.31**	**2.11**	**1.98**	**1.93**	**1.88**
20	**4.35**	**3.49**	**3.10**	**2.87**	**2.71**	**2.60**	**2.51**	**2.45**	**2.28**	**2.08**	**1.95**	**1.90**	**1.84**
21	**4.32**	**3.47**	**3.07**	**2.84**	**2.68**	**2.57**	**2.49**	**2.42**	**2.25**	**2.05**	**1.92**	**1.87**	**1.81**
22	**4.30**	**3.44**	**3.05**	**2.82**	**2.66**	**2.55**	**2.46**	**2.40**	**2.23**	**2.03**	**1.89**	**1.84**	**1.78**
23	**4.28**	**3.42**	**3.03**	**2.80**	**2.64**	**2.53**	**2.44**	**2.37**	**2.20**	**2.01**	**1.86**	**1.81**	**1.76**
24	**4.26**	**3.40**	**3.01**	**2.78**	**2.62**	**2.51**	**2.42**	**2.36**	**2.18**	**1.98**	**1.84**	**1.79**	**1.73**
25	**4.24**	**3.39**	**2.99**	**2.76**	**2.60**	**2.49**	**2.40**	**2.34**	**2.16**	**1.96**	**1.82**	**1.77**	**1.71**
26	**4.23**	**3.37**	**2.98**	**2.74**	**2.59**	**2.47**	**2.39**	**2.32**	**2.15**	**1.95**	**1.80**	**1.75**	**1.69**
27	**4.21**	**3.35**	**2.96**	**2.73**	**2.57**	**2.46**	**2.37**	**2.31**	**2.13**	**1.93**	**1.79**	**1.73**	**1.67**
28	**4.20**	**3.34**	**2.95**	**2.71**	**2.56**	**2.45**	**2.36**	**2.29**	**2.12**	**1.91**	**1.77**	**1.71**	**1.65**
29	**4.18**	**3.33**	**2.93**	**2.70**	**2.55**	**2.43**	**2.35**	**2.28**	**2.10**	**1.90**	**1.75**	**1.70**	**1.64**
30	**4.17**	**3.32**	**2.92**	**2.69**	**2.53**	**2.42**	**2.33**	**2.27**	**2.09**	**1.89**	**1.74**	**1.68**	**1.62**
40	**4.08**	**3.23**	**2.84**	**2.61**	**2.45**	**2.34**	**2.25**	**2.18**	**2.00**	**1.79**	**1.64**	**1.58**	**1.51**
50	**4.03**	**3.18**	**2.79**	**2.56**	**2.40**	**2.29**	**2.20**	**2.13**	**1.95**	**1.74**	**1.58**	**1.51**	**1.44**
60	**4.00**	**3.15**	**2.76**	**2.53**	**2.37**	**2.25**	**2.17**	**2.10**	**1.92**	**1.70**	**1.53**	**1.47**	**1.39**
120	**3.92**	**3.07**	**2.68**	**2.45**	**2.29**	**2.18**	**2.09**	**2.02**	**1.83**	**1.61**	**1.43**	**1.35**	**1.25**
∞	**3.84**	**3.00**	**2.60**	**2.37**	**2.21**	**2.10**	**2.01**	**1.94**	**1.75**	**1.52**	**1.32**	**1.22**	**1.00**

Tablo E F-dağılımına ait kritik değerler ($p = 0.01$)



	Pay *s.d.*												
Payda *s.d.*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**12**	**24**	**60**	**120**	**∞**
1	4052	4999	5403	5625	5764	5859	5929	5981	6106	6235	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.42	99.46	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.05	26.60	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.37	13.93	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	9.89	9.47	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.72	7.31	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.47	6.07	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.67	5.28	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.11	4.73	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.71	4.33	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.40	4.02	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.16	3.78	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	3.96	3.59	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	3.80	3.43	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.67	3.29	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.55	3.18	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.46	3.08	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.37	3.00	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.30	2.92	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.23	2.86	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.17	2.80	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.12	2.75	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.07	2.70	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.03	2.66	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	2.99	2.62	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	2.96	2.58	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	2.93	2.55	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	2.90	2.52	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	2.87	2.49	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	2.84	2.47	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.66	2.29	2.02	1.92	1.80
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.56	2.18	1.91	1.80	1.68
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.50	2.12	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.34	1.95	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.18	1.79	1.47	1.32	1.00