

## Varyans Analizinde Sabit ve Rastgele Faktörler

Prof.Dr. Ömer Cevdet Bilgin

ANOVA da en zor kararlardan biri hangi faktörün sabit ve hangisinin rastgele olduğuna karar vermektir. Bu zor olduğu kadar da çok önemlidir. Modelin sabit ve rastgele faktörlerini doğruca spesifikleştirmek isabetli bir analize ulaşmak için elzemdir.

Ne yazık ki pek çok kitaptaki tanımlar faktörleri sabit veya rastgele olarak spesifikleştirmeye karar vermede pek yardımcı olmaz. Kitap misalleri çoğunlukla yapaydır ve araştırmacının üzerinde çalıştığı daha kafa karıştırıcı gerçek veriye uygulanmaları zordur.

Esasında, nükte şurada ki araştırmacının objektifine bağlı olarak, aynı faktör çoğu zaman sabit veya rastgele olabilir.

Bundan ötürü, her zaman doğru veya yanlış bir karar yoktur. Vereceğin kararın doğruluğu veya yanlışlığı o faktör hakkında yapmak istediğin çıkarımlara bağlıdır.

### Bir Misal

Şöyle bir deneyi mülhaza edelim ki kerevizlerin verimini inceliyor olsun. Araştırmacı deneyi altı çiftlikte ve her çiftlikte dört tarlada tekrarlıyor.

Kerevizin iki çeşidi var ve araştırmacı mevsim sonunda her tarlada 30 bitkinin her birinin verimini ölçüyor. Araştırmacı çeşitlerin verimlerindeki farkları tespit etmek istiyor.

Buna göre, deney şu faktörlere sahip: ÇEŞİT, ÇİFTLİK ve TARLA

### Sabit faktörler farklar cinsinden tanımlanır.

Bir sabit faktörün test edilen düzeylerini araştırmacı kendi keyfince belirler.

Bir kategorik sabit faktörün etkisi genel aritmetik-ortalamadan farkları ile ölçülür.

Bir sürekli sabit faktörün (çoğu zaman *kovaryete* olarak adlandırılır) etkisi onun eğimi ile tanımlanır. Yani, çıkım değişkeninin aritmetik ortalaması tahminci değişkenin farklı değerleri ile ne kadar farklılaşıyor.

Sabit faktörlerin analiz çıktısı, buna göre, aritmetik ortalama farkların ve eğimlerin kestirimlerini verir.

Sabit faktörlere dair neticeler bu faktörlerin deneye dahil edilen düzeylerine münhasırdır. Mesela, kerevizin bir çeşidi diğerine göre anlamlıca daha çok verime sahipse, yalnızca bu iki çeşit hakkında çıkarımda bulunabiliriz. Test etmediğimiz kereviz çeşitleri hakkında bu sonuç hiç bir şey söylemez.

**Rastgele faktörler bir dağılım ile tanımlanır.**

Bir rastgele faktörün test edilen düzeylerinin bir düzeyler popülasyonundan rastgele seçildiği varsayılır.

Bir rastgele faktörün çıktısı varyansın bir kestirimidir. Aritmetik ortalamadan farklar değildir.

Rastgele faktörlere dair neticeler bu faktörlerin deneye dahil edilmeyen düzeylerine genelleştirilebilir. Mesela, bulabiliriz ki tarlalar arasındaki varyans verimin genel varyansının belli bir yüzdesini teşkil ediyor. Fakat verimin bir tarlada bir başkasına göre ne kadar daha çok olduğunu ölçmüyoruz. Yalnızca değişimleriyle ilgileniyoruz, ne kadar farklı olduklarıyla değil.

**Hangi durumda bir faktör sabit olarak spesifikleştirilmelidir?**

1. Faktör karşılaştırmak istediğin birincil muameledir. Misalimizde, ÇEŞİT şüphesiz sabittir çünkü araştırmacı çeşitlerin aritmetik ortalama verimlerini karşılaştırmak istiyor.
2. Faktör ikincil kontrol değişkenidir ve bu faktörün spesifik düzeyleri arasındaki farkları kontrol altına almak istiyorsun. Diyelim ki araştırmacı mesela toprak tipleri veya topoğrafyaları gibi verimi etkileyebilecek herhangi bazı vasıflardan dolayı spesifik olarak bu çiftlikleri seçti. Eğer araştırmacı o toprak tiplerinin veya topoğrafyaların temsilcileri olarak bu çiftlikleri karşılaştırmak istiyorsa, o zaman ÇİFTLİK sabit olmalıdır.
3. Faktörün yalnızca iki düzeyi var. Diğer her şey bir faktörün rastgele olması gerektiğini söylüyor olsa bile, eğer yalnızca iki düzeyi varsa, sabittir.

**Hangi durumda bir faktör rastgele olarak spesifikleştirilmelidir?**

1. Genel varyansın ne kadarının bu faktöre atfedileceğiyle ilgileniyorsun. Eğer verimin varyansının ne kadarını verimin alındığı çiftliğe atfedebileceğini bilmek istiyorsan, ÇİFTLİK rastgele olacaktır.
2. Hangi aritmetik ortalamaların farklı olduğuyla değil de, bu faktördeki varyasyonu hesaba katmakla ilgileniyorsun. Eğer çiftlikler rastgele seçildilerse, ÇİFTLİK rastgele olmalıdır. Öğrenme-çalışmasına dahil edilen spesifik çiftliklerin seçimi kilitlidir. Eğer farklı spesifik çiftlikleri – çiftlik faktörünün farklı düzeylerini – dahil etmek suretiyle, öğrenme-çalışmasını yeniden yürütebiliyorsan ve aynı neticelere varabiliyorsan, o zaman ÇİFTLİK rastgele olmalıdır.
3. Bu faktör hakkındaki neticeleri popülasyonun tamamına genelleştirmek istiyorsun. Araştırmacı spesifik tarlaların verimlerini karşılaştırmakla ilgilenmiyor. Ondan ziyade, araştırmacı bu deneyin sonuçlarını tüm tarlalara genelleştirmek istiyor, bundan ötürü TARLA rastgeledir.
4. Bir rastgele faktörün herhangi bir interaksyonu da rastgeledir. Bir modelin faktörlerinin nasıl spesifikleştirildiğinin yapılacak analiz ve çıkarılacak neticeler üzerinde büyük tesiri vardır.