

# Çoklu Regresyon

# Çoklu Regresyon

- Çoklu regresyon birçok alanda popüler bir tekniktir.

# Çoklu Regresyon

- Regresyon analizleri bir bağımlı değişken ve birkaç bağımsız değişken arasındaki ilişkiyi değerlendirmeye imkân veren bir takım istatistiksel tekniklerdir.
- Örneğin, ilköğretim düzeyindeki okuma yeteneği (bağımlı değişken) algısal gelişim, motor gelişim ve yaş gibi farklı bağımsız değişkenler ile ilişkili midir?
- *Regresyon* terimi genellikle analizin amacı yordama olduğunda ve *korelasyon* terimi ise amaç basitçe bağımlı değişken ve bağımsız değişkenler arasındaki ilişkiyi değerlendirmek olduğunda kullanılmasına rağmen; *regresyon* ve *korelasyon* terimleri az ya da çok bu süreçleri adlandırmak için birbirinin yerine kullanılmaktadırlar.

# Çoklu Regresyon

Çoklu regresyon tek bağımsız değişken yerine her bir bireye ait bir bağımlı değişken değerini yordamak için birden çok bağımsız değişkenin bir araya getirildiği iki değişkenli regresyonun bir uzantısıdır.

Regresyonun sonucu, birkaç sürekli (ya da ikili) bağımsız değişkenden bağımlı değişkeni en iyi yordayan denklem elde edilmeye çalışılır:

$$Y' = A + B_1X_1 + B_2X_2 + \dots + B_kX_k$$

# Çoklu Regresyon

$$Y' = A + B_1X_1 + B_2X_2 + \dots + B_kX_k$$

Denklemden  $Y'$  bağımlı değişkenin yordanan değerini,  $a$   $Y'$ 'nin kesişimini (tüm  $X$  değeri 0 olduğunda  $Y'$ 'nin değeri),  $X_s$  bağımsız değişkenleri (burada  $k$  tane) temsil eder ve  $B_s$  regresyon sırasında her bir bağımsız değişkene atanan katsayılarıdır. Örneklerdeki tüm denekler için bağımlı değişken değerini yordanmakta aynı kesişim ve katsayılar kullanılmasına rağmen, her bireyin kendi  $X$  değerlerinin denkleme girilmesinin sonucu olarak her bir birey için  $Y'$  değeri farklı yordandır.

# Çoklu Regresyon

- Regresyonun amacı, denklemden yordanan  $Y$  değerlerini ölçme sonucu elde edilen  $Y$  değerlerine mümkün olduğu kadar yaklaştıran ve *regresyon katsayıları* olarak adlandırılan bağımsız değişkenlere ait  $B$  değerleri setine ulaşmaktır.
- Amaç: regresyon katsayıları elde edilen ve yordanan  $Y$  değerleri arasındaki sapmaları (kareler toplamı) en aza indirir ve veri seti için yordanan ve elde edilen  $Y$  değerleri arasındaki korelasyonu en uygun hale getirir.

# Çoklu Regresyon

## Regresyon teknikleri

- standart çoklu regresyon,
- sıralı (hiyerarşik) regresyon
- istatistiksel (adımsal) regresyon

Bu teknikler arasındaki farklılıklar değişkenlerin denkleme giriş şeklinden kaynaklanır:

Değişkenler tarafından paylaşılan varyansa ne oluyor  
Değişkenlerin denkleme giriş sırasına kim karar veriyor?

# Çoklu Regresyon

- Regresyon analizin temel amacı genellikle bir bağımlı değişken ve birkaç bağımsız değişken arasındaki ilişkiyi incelemektir. İlk aşama olarak, araştırmacı bağımlı değişken ve bağımsız değişkenler arasındaki ilişkinin ne kadar güçlü olduğunu belirler; daha sonra bazı belirsizliklerle her bir bağımsız değişkenin ilişkideki önemi değerlendirilir.



- Biraz daha karmaşık bir amaç, bir bağımlı değişken ve bazı bağımsız değişkenler arasındaki ilişkiyi diğer bağımsız değişkenlerin etkisi istatistiksel olarak ortadan kaldırarak incelemek olabilir.
- Örneğin, matematiksel eğitimin kapsamı ve zorluğu istatistiksel olarak düzeltildikten sonra cinsiyet matematik performansının yordanmasına katkı sağlar mı?

# Çoklu Regresyon

- Regresyon analizleri sürekli ve ikili bağımsız değişkenlerle kullanılabilir. Başlangıçta kesikli olan bir değişken ilk olarak 1'ler ve 0'larla yapay kodlanarak ikili değişkenlere (kesikli değişken kategori sayısının bir eksiği) dönüştürülürse kullanılabilir.

# Çoklu Regresyon

## İlişkinin Derecesi

- Regresyon denklemi ne kadar iyidir?
- Regresyon denklemi gerçekten şansla yordamadan daha iyi sonuç sağlar mı?
- Örneğin, algısal gelişim, motor gelişim ve yaş hakkında bilgi verildiğinde okuma yeteneği güvenilir olarak yordanabilir mi?

# Çoklu Regresyon

## Bağımsız Değişkenlerin Önemi

- Çoklu korelasyonlar sıfırdan farklıysa, hangi bağımsız değişkenlerin denklem için önemli olduğunu ve hangilerinin önemli olmadığını sormak isteyebilirsiniz. Örneğin, motor gelişim bilgisi okuma yeteneğini yordamada yararlı mıdır ya da sadece yaş ve algısal gelişim bilgisi ile aynısını yapabilir miyiz?

# Çoklu Regresyon

## Bağımsız Değişkenleri Ekleme

Bir regresyon denklemi hesapladığınızı düşününün ve denkleme bir ya da daha fazla bağımsız değişken eklemenin bağımlı değişkeni yordamayı geliştirip geliştiremeyeceğinizi bilmek istiyorsunuz. Örneğin, denklemde var olan 3 bağımsız değişkene ebeveynlerin okuma ilgisini yansıtan bir değişken eklenmesi ile çocuğun okuma yeteneğinin yordanması geliştirilir mi?

# Çoklu Regresyon

## Bağımsız Değişkenlerin Setlerinin Karşılaştırılması

- Bir bağımlı değişkeni bir bağımsız değişken setinden kestirmek diğer bağımsız değişken setinden kestirmekten daha iyi midir? Örneğin, okuma yeteneğini algısal gelişim, motor gelişim ve yaşa dayanan yordama aile geliri ve aile eğitim erişimine dayanan yordama kadar iyi midir?

# Çoklu Regresyon

## Parametre Kestirimleri

- Çoklu regresyonda parametre kestirimleri standartlaştırılmamış regresyon katsayılarıdır ( $B$  ağırlıkları).
- Örneğin, GPA puanlarından GRE puanlarını yordamak istediğimizi ve bizim analizimizin aşağıdaki denklemi ürettiğini düşünelim:

$$(GRE)' = 200 + 100 (GPA)$$

- Denklemden  $B = 100$ , GPA'daki her bir birimlik artışta (örneğin, GPA'nın 2'den 3'e artması) GRE puanlarında 100 birimlik artış beklediğimizi gösterir. Bazen bu, bağımlı değişkendeki kazanç yüzdesi cinsinden yararlı şekilde ifade edilir. Örneğin, ortalama GRE'nin 500 olduğunu varsayalım, bir puanlık artış GRE'de ortalama %20 artışı temsil eder.

# Çoklu Regresyon

## Kuramsal Konular

Bununla birlikte, bağımsız değişkeni seçmede dikkate alınacak bazı genel hususlar vardır. Her bir bağımsız değişken bağımlı değişkenle güçlü ilişkili olduğunda ve diğer bağımsız değişkenlerle ilişkisiz olduğunda regresyon en iyi halini alır. Öyleyse, regresyonun genel amacı, bağımlı değişkeni yordamak için gerekli olan en az sayıdaki bağımsız değişkenleri belirlemektir, burada her bir bağımsız değişken bağımlı değişkendeki değişkenliğin önemli ve bağımsız bir bölümünü yordar.



# Çoklu Regresyon

- **Deneklerin Bağımsız Değişkenlere Oranı**

Deneklerin bağımsız değişkenlere oranlarının mantıklı olması gerekir yoksa sonuç mükemmel ama anlamsız çıkar. Deneklerden daha fazla bağımsız değişken olduğunda sadece deneklerin bağımsız değişkenlere oranının yapay olarak ayarlanması ile araştırmacı her denek için bağımlı değişkenin tamamen yordandığı bir regresyon çözümü bulunabilir.

# Çoklu Regresyon

- **Deneklerin Bağımsız Değişkenlere Oranı**

Pratik bazı basit temel kurallar çoklu korelasyon testi için  $N \geq 50 + 8m$  ( $m$  bağımsız değişken sayısı) ve bireysel yordayıcıların testi için  $N \geq 104 + 8m$  biçimindedir. Bu temel kurallar  $\alpha = .05$  ve  $\beta = .20$  olduğunda bağımlı değişken ve bağımsız değişkenler arasında ortalama bir ilişkiyi varsayar.

# Çoklu Regresyon

## Deneklerin Bağımsız Değişkenlere Oranı

İstatistiksel (adımsal) regresyon kullanılacak ise, daha fazla deneğe ihtiyaç vardır. Deneklerin bağımsız değişkenlere oranının 40:1 olması uygundur; çünkü istatistiksel regresyon örneklem büyük olmadığında örneklemin ötesine genelleştirilemeyen bir çözüm üretebilir. Çözümün genellenebilirliğini test etmek için çapraz geçerlik kullanılırsa (deneklerin bir kısmından çözüm elde edilir, kalan kısmında çözüm test edilir), istatistiksel regresyonda daha büyük örneklemelere ihtiyaç duyulur.

# Çoklu Regresyon

**Bağımsız Değişkenler Arasında ve Bağımlı Değişkende Uç Değerlerin Olmaması**

Uç deneklerin regresyon çözümü üzerinde çok fazla etkisi vardır ve ayrıca uç denekler regresyon ağırlıklarının kestiriminin doğruluğunu etkiler (Fox, 1991).

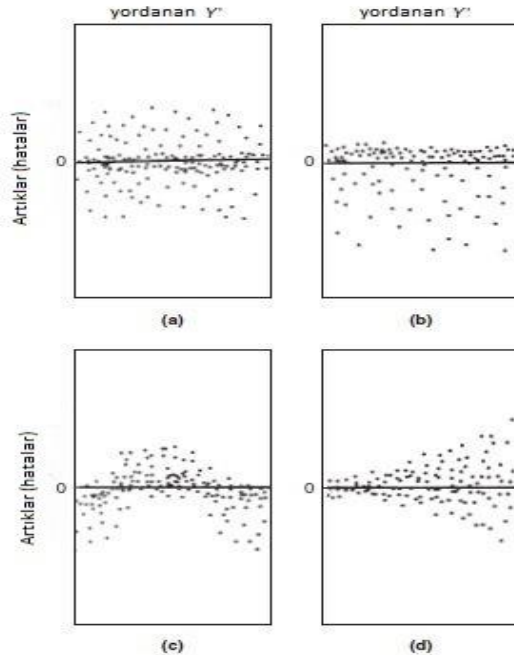
# Çoklu Regresyon

## Çoklu Bağlantı ve Tekillliğin Yokluğu

Çoklu regresyon programlarının çoğunda kullanıcıyı çoklu bağlantılı bağımsız değişkenlerin dahil edilmesine karşı koruyan varsayılan tolerans değerleri ( $1 - \text{SMC}$ ) vardır. Eğer programların varsayılan değerleri uygunsa, denklemde var olan bağımsız değişkenlerle yüksek korelasyon gösteren bağımsız değişkenler denkleme eklenmez. Bu durum, bu bağımsız değişkenler regresyon katsayılarını şişirerek analizi tehdit ettikleri için ve diğer bağımsız değişkenlerle korelasyonlarından dolayı bu bağımsız değişkenlere ihtiyaç duyulmadığı için hem istatistiksel hem de mantıksal olarak anlamlıdır.

# Çoklu Regresyon

## Artıkların Normallığı, Doğrusallığı ve Eşvaryanslılığı



(a) Varsayımlar karşılandığında, (b) normallik sağlanmadığında, (c) doğrusal olmama durumunda ve (d) eşvaryanslı olmama durumunda bağımlı değişkenin yordanan değerine ( $\hat{Y}$ ) karşı artıkların grafikleri

# Çoklu Regresyon

- artıklar normallik, doğrusallık ve eşvarsyanslılık varsayımlarını sağlıyorsa,
- herhangi bir uç değer yoksa,
- denek sayısı yeterli ise
- çoklu bağlantı ve teklik için hiçbir kanıt yoksa – regresyon için sadece bir analiz gereklidir .

