

# Korelasyon Analizi

## 1. Korelasyon Analizi Nedir?

Korelasyon analizi, iki veya daha fazla deęişken arasındaki ilişkiyi ölçmek ve bu ilişkinin yönü ile gücünü belirlemek için kullanılan bir istatistiksel yöntemdir. Korelasyon, bir deęişkendeki deęişimin dięer bir deęişkende nasıl bir deęişikliğe yol açtığını anlamamıza yardımcı olur. Korelasyon katsayısı, bu ilişkinin sayısal göstergesidir ve genellikle "-1" ile "1" arasında bir deęer alır.

- **Pozitif Korelasyon:** Bir deęişken arttığında dięeri de artıyorsa pozitif korelasyon vardır.
- **Negatif Korelasyon:** Bir deęişken arttığında dięeri azalıyorsa negatif korelasyon vardır.
- **Korelasyon Yok:** Deęişkenler arasında belirgin bir ilişki yoksa korelasyon sıfırdır.

---

## 2. Korelasyon Türleri

- **Pearson Korelasyon Katsayısı:** Her iki deęişkenin de sürekli ve en az interval (aralık) ölçeğinde olduğu durumlarda kullanılır.
  - **Spearman Sıra Korelasyonu:** Deęişkenlerin sıralı (ordinal) ölçekte olduğu veya normal dağılım göstermediği durumlarda kullanılır.
  - **Kendall Tau Korelasyonu:** Spearman korelasyonuna benzer ama küçük örneklem büyüklükleri veya çok sayıda bağımsız ölçüm varken daha tutarlı sonuçlar verir.
  - **Point-Biserial Korelasyonu:** Bir deęişkenin iki kategorili (dichotomous), dięerinin sürekli olduğu durumlarda kullanılır.
  - **Phi Korelasyonu:** Her iki deęişkenin de iki kategorili olduğu durumlarda kullanılır.
-

### 3. Korelasyon Testleri

Korelasyon analizinde kullanılan bazı testler şunlardır:

- **Pearson Korelasyon Testi:** İki sürekli değişken arasındaki doğrusal ilişkiyi test eder.
- **Spearman Sıra Korelasyon Testi:** İki sıralı veya sürekli değişken arasındaki ilişkiyi test eder.
- **Kendall Tau Testi:** İki sıralı değişken arasındaki ilişkiyi test eder.
- **Chi-Kare Testi:** İki kategorik değişken arasındaki ilişkiyi test etmek için kullanılır.

---

### 4. Farklı Ölçme Düzeyleri Arasındaki Korelasyon Örnekleri

Örnekleri, sıklıkla karşılaşılan durumlara göre hazırlayalım:

#### Örnek 1: Sürekli ve Sürekli Değişkenler Arasındaki Korelasyon (Pearson Korelasyon)

- **Değişkenler:** Bir öğrencinin sınav puanı (0-100 arası) ve haftalık ders çalışma saati.

Çözüm: Pearson korelasyon testi kullanılarak puan ve çalışma saatleri arasındaki doğrusal ilişki ölçülür. Pozitif bir korelasyon beklenir; çünkü daha fazla çalışan öğrenciler genellikle daha yüksek puan alırlar.

#### Örnek 2: Sıralı ve Sürekli Değişkenler Arasındaki Korelasyon (Spearman Korelasyon)

- **Değişkenler:** Müşteri memnuniyet dereceleri (1: Çok memnun değil, 5: Çok memnun) ve müşteri tarafından harcanan toplam para miktarı (TL).

Çözüm: Bu durumda Spearman korelasyon testi uygundur. Çünkü memnuniyet sıralı bir ölçekteyken, harcanan miktar sürekli bir değişkendir. Korelasyon katsayısı pozitifse, memnuniyet arttıkça harcamanın da arttığını söyleyebiliriz.

### **Örnek 3: Kategorik ve Sürekli Değişkenler Arasındaki Korelasyon (Point-Biserial Korelasyon)**

- **Değişkenler:** Cinsiyet (Kadın = 0, Erkek = 1) ve bir sınavdan alınan puan.

Çözüm: Point-biserial korelasyon testi, cinsiyet ve puan arasındaki ilişkiyi incelemek için kullanılır. Test sonucunda anlamlı bir korelasyon bulunursa, cinsiyetin puan üzerinde bir etkisi olduğu düşünülebilir.

### **Örnek 4: Kategorik ve Kategorik Değişkenler Arasındaki Korelasyon (Phi Korelasyon)**

- **Değişkenler:** Bir ürünün satın alınıp alınmadığı (0: Hayır, 1: Evet) ve bir reklamın izlenip izlenmediği (0: Hayır, 1: Evet).

Çözüm: Phi korelasyon testi, iki kategorik değişken arasındaki ilişkiyi incelemek için kullanılır. Pozitif bir korelasyon varsa, reklamı izleyenlerin ürünü satın alma olasılığının daha yüksek olduğu söylenebilir.

---

Bu ders notları korelasyon analizinin temel prensiplerini, türlerini ve testlerini kapsamaktadır. Ölçme düzeylerine göre seçilen örnekler, farklı veri türleriyle nasıl çalışılabileceğini ve uygun korelasyon yöntemlerini göstermektedir.

## Örnek 1: Pearson Korelasyon Katsayısı Hesaplaması

Veri Seti:

Öğrenci	Çalışma Saati (X)	Sınav Puanı (Y)
1	5	70
2	3	50
3	8	90
4	6	80
5	4	60

Adım 1: Pearson Korelasyon Formülü

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

Hesaplayalım ve sonucu yorumlayalım.

Pearson korelasyon katsayısı  $r = 0.986$  olarak bulunmuştur. Bu, çalışma saati ile sınav puanı arasında çok güçlü ve pozitif bir ilişki olduğunu gösterir. Yani öğrenciler daha fazla çalıştıkça, sınav puanları da genellikle artmaktadır.

### Adım 1: Gerekli Değerlerin Hesaplanması

Pearson korelasyon formülü:

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

1.  $n$  (gözlem sayısı): 5
2.  $\sum X = 5 + 3 + 8 + 6 + 4 = 26$
3.  $\sum Y = 70 + 50 + 90 + 80 + 60 = 350$
4.  $\sum XY = (5 \times 70) + (3 \times 50) + (8 \times 90) + (6 \times 80) + (4 \times 60) = 350 + 150 + 720 + 480 + 240 = 1940$
5.  $\sum X^2 = 5^2 + 3^2 + 8^2 + 6^2 + 4^2 = 25 + 9 + 64 + 36 + 16 = 150$
6.  $\sum Y^2 = 70^2 + 50^2 + 90^2 + 80^2 + 60^2 = 4900 + 2500 + 8100 + 6400 + 3600 = 25500$

### Adım 2: Formülde Yerine Koyma

$$r = \frac{5(1940) - (26)(350)}{\sqrt{[5(150) - 26^2][5(25500) - 350^2]}}$$

$$r = \frac{9700 - 9100}{\sqrt{[750 - 676][127500 - 122500]}}$$

$$r = \frac{600}{\sqrt{74 \times 5000}}$$

$$r = \frac{600}{\sqrt{370000}}$$

$$r = \frac{600}{608.28}$$

$$r \approx 0.986$$

## Örnek 2: Spearman Sıra Korelasyonu Hesaplaması

Spearman korelasyonunu, sıralı ölçekte olan bir örnek veri setiyle yapalım.

Veri Seti:

Öğrenci	Memnuniyet (1-5)	Harcama Miktarı (TL)
1	5	1200
2	3	600
3	4	900
4	2	400
5	1	200

Bu veri seti üzerinden Spearman korelasyon katsayısını hesaplayalım. [↗]

Spearman korelasyon katsayısı  $\rho = 1.0$  olarak bulunmuştur. Bu, memnuniyet düzeyi ile harcama miktarı arasında mükemmel ve pozitif bir sıralı ilişki olduğunu gösterir. Yani memnuniyet düzeyi arttıkça harcama miktarı da artmaktadır.

Veri seti:

Öğrenci	Memnuniyet (1-5)	Harcama Miktarı (TL)	Memnuniyet Sırası	Harcama Sırası
1	5	1200	5	5
2	3	600	3	3
3	4	900	4	4
4	2	400	2	2
5	1	200	1	1

### Adım 1: Değişkenlerin Sıralamasını Belirleme

Öğrenci	Memnuniyet (X)	Harcama (Y)	Memnuniyet Sırası (Rx)	Harcama Sırası (Ry)	d = Rx - Ry	d <sup>2</sup>
1	5	1200	5	5	0	0
2	3	600	3	3	0	0
3	4	900	4	4	0	0
4	2	400	2	2	0	0
5	1	200	1	1	0	0

### Adım 2: Spearman Korelasyon Katsayısı Formülü

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Yerine koyarsak:

$$\rho = 1 - \frac{6(0)}{5(5^2 - 1)}$$

$$\rho = 1 - 0 = 1$$

**Sonuç:** Spearman korelasyon katsayısı  $\rho = 1$ , yani mükemmel bir sıralı ilişki vardır.

### Örnek 3: Point-Biserial Korelasyon Katsayısı

Veri Seti:

Katılımcı	Cinsiyet (0: Kadın, 1: Erkek)	Gelir (TL)
1	0	3000
2	1	4500
3	0	3200
4	1	5000
5	0	3500

#### Adım 1: Gerekli Değerlerin Hesaplanması

- Ortalama gelir erkekler için ( $M_1$ ):

$$M_1 = \frac{4500 + 5000}{2} = 4750$$

- Ortalama gelir kadınlar için ( $M_0$ ):

$$M_0 = \frac{3000 + 3200 + 3500}{3} = 3233.33$$

- Toplam gelir ortalaması:

$$\bar{Y} = \frac{3000 + 4500 + 3200 + 5000 + 3500}{5} = 3840$$

- Standart sapma ( $s$ ):

$$s = \sqrt{\frac{(3000 - 3840)^2 + (4500 - 3840)^2 + (3200 - 3840)^2 + (5000 - 3840)^2 + (3500 - 3840)^2}{5 - 1}}$$

Hesaplayalım:

$$s = \sqrt{\frac{(-840)^2 + (660)^2 + (-640)^2 + (1160)^2 + (-340)^2}{4}}$$

$$s = \sqrt{\frac{705600 + 435600 + 409600 + 1345600 + 115600}{4}}$$

$$s = \sqrt{\frac{3004000}{4}} = \sqrt{751000} \approx 866.91$$



## Adım 2: Point-Biserial Korelasyon Formülü

$$r_{pb} = \frac{(M_1 - M_0)}{s} \sqrt{\frac{n_1 n_0}{n^2}}$$

- $M_1 = 4750$
- $M_0 = 3233.33$
- $s = 866.91$
- $n_1 = 2$  (erkek),  $n_0 = 3$  (kadın)
- $n = 5$

$$r_{pb} = \frac{(4750 - 3233.33)}{866.91} \sqrt{\frac{(2)(3)}{5^2}}$$

$$r_{pb} = \frac{1516.67}{866.91} \sqrt{\frac{6}{25}}$$

$$r_{pb} = 1.75 \times \sqrt{0.24}$$

$$r_{pb} \approx 1.75 \times 0.49 = 0.856$$

**Sonuç:** Point-biserial korelasyon katsayısı  $r_{pb} \approx 0.856$ .