# XIII. BÖLÜM
# PARAMETRİK OLMAYAN İLİŞKİ KATSAYILARI

## XIII.1 Giriş

Parametrik olmayan korelasyon kaysayıları parametrik korelasyon (Pearson çarpım momentleri korelasyonu) katsayısının parametrik oölmayan tahminidir. Eşit aralıklı, oransal ve sıralama düzeylerindeki verilere uygulanabilirler. Değişkeblerin gözlem değerleri normal dağılım göstermediğinde ve küçük örneklemlerde kullanılmaları daha uygun olur. Burada sadece en yaygın kullanıma sahip olan iki tanesi verilecektir. Bunlar Kendall'ın Tausu ( $\tau$) ve Spearman'ın Rho ($\rho_S$) sıra korelasyonudur.

## XIII.2 Kkendall'ın Tau ( $\tau$) İlişki Katsayısı

Kendall, orijinal çalışmasında $n$ adet gözlem ikilisinin herhangi bir örneklem için $(x_i, y_i)$ ve $(x_j, y_j)$ noktaları arasında olası karşılaştırma sayısının $S = \frac{1}{2}[n(n-1)]$ olduğu bilgisinden hareketele karesel çapraz tablolar için de kullanılabilecek bir ilişki katsayı geliştirmiştir. Kendall'ın Tau katsayısı, uyumlu çift sayısı $P$ ile uyumsuz çift sayısı$Q$ arasındaki farkın ( yani; $P-Q$ ), olası karşılaştırma sayısına (yani; $S$'ye) oranı ile elde edilir. Bu nicelik ise $\hat{\tau} = \frac{2(P-Q)}{n(n-1)} = \frac{P-Q}{S}, -1 \leq \hat{\tau} \leq 1$ ile verilir.

$X$ değişkenini orijinal bazı değerleri veya $Y$ değişkeninin orijinal bazı değerleri ya da hem $X$ değişkenini ve hem de $Y$ değişkeninin orijinal bazı değerleri birbirine eşit olduğunda Kendall'ın Tau ilişki katsayısı $\hat{\tau} = \frac{(P-Q)}{\sqrt{\left[\frac{1}{2}n(n-1)-U\right]\left[\frac{1}{2}n(n-1)-V\right]}}$ eşitliği ile hesaplanır. Bu eşitlikte $U = \frac{1}{2}\sum_{x \text{ eşit gözlem}} u(u-1)$ ve $U = \frac{1}{2}\sum_{y \text{ eşit gözlem}} v(v-1)$ değerlerini göstermektedir.

♦ **Örnek:** İstatistik Bölümü birinci sınıf öğrencilerinin matematik notları ile istatistik notları arasında aynı yönde bir ilişki olduğu düşünülmektedir. Bu amaçla yola çıkan bir araştırmacı 10 öğrencinin 100 üzerinden matematik notu ($X$) ile istatistik notu ( $Y$) arasındaki sıra sayıları takipte verildiği gibi elde edilmiştir. (a) Her iki değişken değerlerinde de eşit gözlem değeri yokken, (b) Birici ve sekizinci öğrencilerin matematik notları eşit istatistik nokları aynı iken, (c) Birinci , üçüncü ve beşinci öğrencilerin istatistik motları eşit matematik notları aynı iken (d) (b) ve (c) seçeneklerinde iki durum birlikte geçerli iken Kendall'ın Tau karsayısını hesaplayınız.

| Seçenek | Öğrenci No | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (a) | $x_i$ | 7 | 4 | 3 | 10 | 6 | 2 | 9 | 8 | 1 | 5 |
| | $y_i$ | 5 | 7 | 3 | 10 | 4 | 1 | 6 | 8 | 2 | 9 |
| (b) | $x_i$ | 7.5 | 4 | 3 | 10 | 6 | 2 | 9 | 7.5 | 1 | 5 |
| | $y_i$ | 5 | 7 | 3 | 10 | 4 | 1 | 6 | 8 | 2 | 9 |
| (c) | $x_i$ | 7 | 4 | 3 | 10 | 6 | 2 | 9 | 8 | 1 | 5 |
| | $y_i$ | 4 | 7 | 4 | 10 | 4 | 1 | 6 | 8 | 2 | 9 |
| (d) | $x_i$ | 7.5 | 4 | 3 | 10 | 6 | 2 | 9 | 7.5 | 1 | 5 |
| | $y_i$ | 4 | 7 | 4 | 10 | 4 | 1 | 6 | 8 | 2 | 9 |

♦ **Çözüm:** (a) Verilen sıra sayılarını $x_i$' e göre düzenlersek yeni sıralama ve hesaplama tablosu takipteki gibi olur.

| Öğrenci No | 9 | 6 | 3 | 2 | 10 | 5 | 1 | 8 | 7 | 4 | Toplam |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | - |
| $y_i$ | 2 | 1 | 3 | 7 | 9 | 4 | 5 | 8 | 6 | 10 | - |
| $P$ | 8 | 8 | 7 | 3 | 1 | 4 | 3 | 1 | 1 | - | 36 |
| $Q$ | 1 | 0 | 0 | 3 | 4 | 0 | 0 | 1 | 0 | - | 9 |

$\hat{\tau} = \frac{36-9}{45} = 0.60$ bulunur. değişkenlerin iki özeeliği arasında aynı yönde % 60 bir ilişki olduğu söylenir. (b) (a) seçeneğinde benzer işlemler yapılırsa yeni tablo ve hesaplamalar,

| Öğrenci No | 9 | 6 | 3 | 2 | 10 | 5 | **1** | **8** | 7 | 4 | Toplam |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | **7.5** | **7.5** | 9 | 10 | - |
| $y_i$ | 2 | 1 | 3 | 7 | 9 | 4 | **5** | **8** | 6 | 10 | - |
| $P$ | 8 | 8 | 7 | 3 | 1 | 4 | 2 | 1 | 1 | - | 35 |
| $Q$ | 1 | 0 | 0 | 3 | 4 | 0 | 0 | 1 | 0 | - | 9 |

$x$ değerlerinde bir adet iki eşit gözlem değeri olup $U = \frac{1}{2}2(1) = 1$ ve $y$ değerlerinde eşit gözlem değeri olmadığından $V = 0$ olup $\hat{\tau} = \frac{(35-9)}{\sqrt{[45-1][45-0]}} = \frac{26}{\sqrt{44(45)}} \cong 0.584$ bulunur. Düzeltmesiz bakıldığında ise $\hat{\tau} = \frac{2(P-Q)}{n(n-1)} = \frac{26}{45} \cong 0.578$ bulunur. (c) (a) seçeneğinde benzer işlemler yapılırsa yeni tablo ve hesaplamalar,

| Öğrenci No | 9 | 6 | **3** | 2 | 10 | **5** | **1** | 8 | 7 | 4 | Toplam |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 2 | **3** | 4 | 5 | **6** | **7** | 8 | 9 | 10 | - |
| $y_i$ | 2 | 1 | **4** | 7 | 9 | **4** | **4** | 8 | 6 | 10 | - |
| $P$ | 8 | 8 | 5 | 3 | 1 | 3 | 3 | 1 | 1 | - | 33 |
| $Q$ | 1 | 0 | 0 | 3 | 4 | 0 | 0 | 1 | 0 | - | 9 |

$x$ değerlerinde eşit gözlem değeri olmadığından $U = 0$ ve $y$ değerlerinde bir adet üç eşit gözlem değeri olup $V = \frac{1}{2}3(2) = 3$ olup $\hat{\tau} = \frac{(33-9)}{\sqrt{[45-0][45-3]}} = \frac{24}{\sqrt{42(45)}} \cong 0.552$ bulunur. Düzeltmesiz bakıldığında ise $\hat{\tau} = \frac{2(P-Q)}{n(n-1)} = \frac{24}{45} \cong 0.533$ bulunur. (d) Önceki seçeneklerdeki benzer işlemler yapılırsa yeni tablo ve hesaplamalar,

| Öğrenci No | 9 | 6 | **3** | 2 | 10 | **5** | **1** | **8** | 7 | 4 | Toplam |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | **7.5** | **7.5** | 9 | 10 | - |
| $y_i$ | 2 | 1 | **4** | 7 | 9 | **4** | **4** | **8** | 6 | 10 | - |
| $P$ | 8 | 8 | 5 | 3 | 1 | 3 | 2 | 1 | 1 | - | 32 |
| $Q$ | 1 | 0 | 0 | 3 | 4 | 0 | 0 | 1 | 0 | - | 9 |

$x$ değerlerinde bir adet eşit gözlem değeri olup $U = \frac{1}{2}2(1) = 1$ ve $y$ değerlerinde bir adet üç eşit gözlem değeri olup $V = \frac{1}{2}3(2) = 3$ olup $\hat{\tau} = \frac{(32-9)}{\sqrt{[45-1][45-3]}} = \frac{23}{\sqrt{42(44)}} \cong 0.535$ bulunur. Düzeltmesiz bakıldığında ise $\hat{\tau} = \frac{2(P-Q)}{n(n-1)} = \frac{23}{45} \cong 0.511$ bulunur.

### XIII.3 Spearman Sıra Korelasyon Katsayısı

Spearman'ın sıralama korelasyon katsayısı (veya Spearman'ın rho'su), bu istatistiksel ölçüyü ilk ortaya atan Amerikan istatistikçi Charles Spearman'a atfen adlandırılmıştır. Bir parametrik olmayan ilişki ölçüsüdür. Aralarında ilişki araştırılan değişkenlerin her ikisi de sıralı değişken olduğunda ya da değişkenler arasında doğrusal ilişki varken değişkenlerden en az biri normal dağılımdan uzak olduğunda Spearman Rho ikişki katsayısı kullanılır. Pearson korelasyon katsayısının parametrik olmayan karşılığı olarak bilinir. Aralarındaki en önemli fark Pearson'ın ham değerleri, Spearman Rho'nun sıra sayıları dikkate almasıdır. Spearman Rho'ya aşağıdaki örneklerde sıklıkla baş vurulur:

- Değişkenlerin doğrudan sıralı olarak elde edildiği ya da belli bir kritere göre sıralandığı durumlarda
- Verilerin kesikli sayısal veri türünde olduğu zamanlarda
- Veri sürekli iken gözlem sayısının az olduğu durumlarda
- Değişkenlerin (en az birinin) normal dağılmadığı durumlarda
- İlişkinin doğrusallıktan hafif bir şekilde saptığı durumlarda.

Örneğin: 10 sporcunun, 30 saniye içindeki şınav sayıları ile dikey sıçrama sayıları arasındaki ilişki, BKI (Beden Kütle İndeksi) ile kandaki şeker düzeyi arasındaki ilişki vb örneklerde Spearman Rho ile hesaplanır.

Spearman'ın sıra korelasyon katsayısı $\rho_S$ ile gösterilir. Pearson çarpım-moment korelasyon katsayısının parametrik olmayan özel bir halidir. $\rho_S$ değerinin hesaplanması için iki değişken ($X$ ve $Y$) içinde örneklem verilerinin sıralama düzeninde olmaları gereklidir. Gerçek yaşamda örneklem verileri için bu şart genel olarak uygun değildir. Yani, veriler sıralama düzeninde olabileceği gibi oransal ölçekli veya eşit aralıklı ya da sırasalı ölçekli olarak da bulunur. Hangi durumda olursa olsun orijinal verler bir şekilde sıralama düzeni haline dönüştürülmelidir. Böylece elde edilen sıralana düzeyinde elde edilen veriler yardımı ile $r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$ eşitliği yardımı ile tahmin edilir. Burada $d_i = x_i - y_i, i = 1, 2, \cdots, n$ ile hesaplanır. Yani; değişkenlerin sıra sayıları arasındaki farktır. Eğer değişkemlerin orijinal değerleri arasında bazı gözlemler eşit ise bu tür verilere aynı olan gözlemlerin sıra numaralarının ortalaması ortak sıra numarası olarak atanır.

Bu katsayı, -1 ile 1 arasında tanımlıdır. Katsayının değeri 1'e eşit ise, değişkenler arasında pozitif yönde tam bir ilişkiden ve eğer -1 eşit olursa test yönde tam bir ilişkiden söz edilir. -1 veya 1'e ne kadar yakın bir katsayı elde edilirse, o kadar güçlü bir doğrusal ilişkiden söz edlir. Eğer katsayı değeri 0 ise iki değişkenin sıraları arasında doğrusal bir ilişki yoktur, denir.

♦ **Örnek:** Kendall'ın Tau katsayısında verilen örnekteki veriler için Spearman sıra korelasyon katsayısını hesplayınız.

♦ **Çözüm:** (a) $r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} = 1 - \frac{6(44)}{990} \cong 0.733$, (b) $r_s = 1 - \frac{6(46.5)}{990} \cong 0.718$, (c) $r_s = 1 - \frac{6(50)}{990} \cong 0.697$ ve (d) $r_s = 1 - \frac{6(53.5)}{990} \cong 0.676$ bulunur.

| **Şık→** | **(a)** | | **(b)** | | **(c)** | | **(d)** | |
|---|---|---|---|---|---|---|---|---|
| **Öğrenci No↓** | $d_i$ | $d_i^2$ | $d_i$ | $d_i^2$ | $d_i$ | $d_i^2$ | $d_i$ | $d_i^2$ |
| 1 | 2 | 4 | 2,5 | 6,25 | 3 | 9 | 3,5 | 12,25 |
| 2 | -3 | 9 | -3 | 9 | -3 | 9 | -3 | 9 |
| 3 | 0 | 0 | 0 | 0 | -1 | 1 | -1 | 1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 2 | 4 | 2 | 4 | 2 | 4 | 2 | 4 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 3 | 9 | 3 | 9 | 3 | 9 | 3 | 9 |
| 8 | 0 | 0 | -0,5 | 0,25 | 0 | 0 | -0,5 | 0,25 |
| 9 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 |
| 10 | -4 | 16 | 2,5 | 6,25 | -4 | 16 | -4 | 16 |
| Toplam | 0 | 44 | 0 | 46.5 | 0 | 50 | 0 | 53.5 |

### XIII.4 İlişki Katsayıları İçin Hipotez Testi

#### XIII.4.1 Kendall'ın Tau Katsayısı İçin Hipotez Testi

Kendall'ın Tau karsayısı için hipotez test işlem akışı şöyle verilebilir.

a) **Hipotezler:** $H_0: \tau = 0$ hipotezine karşılık üç farklı iddiada bulunulabilir. Bunlar: i) $H_A: \tau < 0$ (ilişki ters yöndedir), ii) $H_A: \tau > 0$ (ilişki aynı yöndedir) ve $H_A: \tau \neq 0$ (ilişki yoktur) şeklinde kurulur.

b) **Test İstatistiği:** $H_0$'ın doğruluğu varsayımı altında ve $n \geq 10$ iken test istatistiği, $z_H = \frac{3\hat{\tau}\sqrt{n(n-1)}}{\sqrt{2(2n+5)}} \sim N(0; 1)$'dir.

c) **Karar:** Eğer araştırma hipotezi tek yönlü ( i ve ii'de verildiği gibi) ise $z_T = z_{\alpha}$ ve eğer araştırma hipotezi çift yönlü (iii'de verildiği gibi) ise $z_T = z_{\frac{\alpha}{2}}$ kritik tablo değerini göstersin. Eğer $|z_H| \leq |z_T|$ => $H_0$ hipotezi kabul edilir ve eğer $|z_H| > |z_T|$ => $H_0$ hipotezi ret edilir.

d) **Yorum:** Karara ve iddiaya göre yorum yapılır.

♦ **Örnek:** Klasik matematik-İstatistik not örneğinde (a) şıkkında bulunan korelasyon için değişkenler arasında aynı yönde ilişki olduğunu % 5 önem seviyesinde söylemek mümkün müdür?

♦ **Çözüm:**

**a)** Hipotezler: $H_0: \tau = 0$ hipotezine karşılık araştırma hipotezi $H_A: \tau > 0$ (ilişki aynı yöndedir) şeklinde kurulur.

**b)** Test İstatistiği: $H_0$'ın doğruluğu varsayımı altında ve $n \geq 10$ iken test istatistiği, $z_H = \frac{3(0.6)\sqrt{10(9)}}{\sqrt{2(25)}} \cong 2.415$ bulunur.

**c)** Karar: $z_T = z_{0.05} = 1.645$ kritik tablo değeri ( tablodan) bulunur. Böylece $|z_H| > |z_T|$ olduğundan $H_0$ hipotezi ret edilir.

**d)** Yorum: % 95 güvenilirlikle iki değişken arasında aynı yönde bir ilşki olduğu söylenebilir.

### XIII.4.2 Spearman'ın Rho Katsayısı İçin Hipotez Testi

Kendall'ın Tau karsayısı için hipotez test işlem akışı şöyle verilebilir.

**a) Hipotezler:** $H_0: \rho_S = 0$ hipotezine karşılık üç farklı iddiada bulunulabilir. Bunlar: i) $H_A: \rho_S < 0$ (ilişki ters yöndedir), ii) $H_A: \rho_S > 0$ (ilişki aynı yöndedir) ve $H_A: \rho_S \neq 0$ (ilişki yoktur) şeklinde kurulur.

**b) Test İstatistiği:** $H_0$'ın doğruluğu varsayımı altında ve $10 \leq n < 30$ iken test istatistiği, $t_H = \frac{r_S\sqrt{(n-2)}}{\sqrt{1-r_S^2}} \sim t_{(n-2)}$'dir ve $n \geq 30$ iken test istatistiği, , $z_H = r_S\sqrt{(n-1)} \sim N(0; 1)$'dir.

**c) Karar:** Eğer araştırma hipotezi tek yönlü ( i ve ii'de verildiği gibi) ise $t_T = t_{(n-2);\,\alpha}$ veya $z_T = z_{\alpha}$ ve eğer araştırma hipotezi çift yönlü (iii'de verildiği gibi) ise $t_T = t_{(n-2);\,\frac{\alpha}{2}}$ veya $z_T = z_{\frac{\alpha}{2}}$ kritik tablo değerini göstersin. Eğer$|t_H| \leq |t_T|$ veya $|z_H| \leq |z_T|$ => $H_0$ hipotezi kabul edilir ve eğer$|t_H| > |t_T|$ veya $|z_H| > |z_T|$ => $H_0$ hipotezi ret edilir.

**d) Yorum:** Karara ve iddiaya göre yorum yapılır.

♦ **Örnek:** Klasik matematik-İstatistik not örneğinde (a) şıkkında bulunan korelasyon için değişkenler arasında aynı yönde ilişki olduğunu % 5 önem seviyesinde söylemek mümkün müdür?

♦ **Çözüm:**

a) Hipotezler: $H_0: \rho_S = 0$ hipotezine karşılık araştırma hipotezi $H_A: \rho_S > 0$ (ilişki aynı yöndedir) şeklinde kurulur.

b) Test İstatistiği: $H_0$'ın doğruluğu varsayımı altında ve $n \geq 10$ iken test istatistiği,$t_H = \frac{r_S\sqrt{(n-2)}}{\sqrt{1-r_S^2}} \cong \frac{0.733\sqrt{8}}{\sqrt{1-(0.733)^2}} \cong 3.048$ bulunur.

c) Karar: $t_T = t_{8;\,0.05} = 1.860$ kritik tablo değeri ( tablodan) bulunur. Böylece $|t_H| > |t_T|$ olduğundan $H_0$ hipotezi ret edilir.

d) Yorum: % 95 güvenilirlikle iki değişken arasında aynı yönde bir ilşki olduğu söylenebilir.

♦ **Çalışma Sorusu:**

Hava kirliliği ile sebep olduğu hastalıklar arasında aynı yönde bir ilişki olduğu iddia edilmektedir. Bu amaçla yapılan bir araştırmada 10 şehir tesadüfen seçilmiş ve sözü edilen özellikler bakımından (1= En kötü; 10= En iyi) 1'den 10'a kadar sıralanmışlardır. Sonuçlar takipteki tabloda verildiği gibi elde edilmiştir.

| Şehir | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| Kirlilik ($x_i$; sıra) | 4 | 7 | 9 | 1 | 2 | 10 | 3 | 5 | 6 | 8 |
| Hastalık ($y_i$, sıra) | 5 | 4 | 7 | 3 | 1 | 10 | 2 | 8 | 6 | 9 |

Değişkenler arasında aynı yönde ilişki olduğunu % 5 önem seviyesinde: (a) Kendall'ın Tau katsayısı be (b) Spearman'ın Rho katsayısı ile ayrı ayrı söylemek mümkün müdür?

(Cevap: (a) $\hat{\tau} \cong 0.644$, $z_H \cong \frac{3(0.644)\sqrt{10(90)}}{\sqrt{2(25)}} \cong 2.592$, $H_0$ hipotezi ret; (b) $r_S = 1 - \frac{6(30)}{10(99)} \cong 0.818$, $t_H = \frac{r_S\sqrt{(n-2)}}{\sqrt{1-r_S^2}} \cong \frac{0.818\sqrt{8}}{\sqrt{1-(0.818)^2}} \cong 4.022$, $H_0$ hipotezi ret)

## Kaynaklar


1. Spearman, C. (1904). The proof and measurement of association between two things. *Amer.J.Psychol.* C.15 say.72–101
2. Kendall, M. G. (1962). *Rank correlation methods*, Griffin.
3. Hollander, M. and Wolfe, D. A. (1973). *Nonparametric statistical methods*, New York: Wiley
4. Myers, J. L. and Well, A. D. (2003). *Research Design and Statistical Analysis (2.ed.)*, Lawrence Erlbaum.
5. Siegel, S. and Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences 2. ed.* New York: McGraw-Hill.