

## Shapiro-Wilk Testi

Shapiro-Wilk (W) testi,  $X_1, X_2, \dots, X_n$  rastgele örnekleminin normal dağılıma sahip olup olmadığını test etmek amacıyla kullanılır. Sıfır ve alternatif hipotezler aşağıda gösterildiği gibidir:

$H_0$ : Veri setinin dağılımı normaldir

$H_1$ : Veri setinin dağılımı normal değildir.

W test istatistiğinin değerini hesaplamak için,

1. Öncelikle örneklemden elde edilen veriler küçükten büyüğe doğru sıralanır.

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

Burada,  $X_{(i)}$  i-inci sıra istatistiği olarak adlandırılır.

2.  $s^2 = \sum_{i=1}^n (X_{(i)} - \bar{X})^2$  değeri hesaplanır.

3. Eğer örneklem büyüklüğü  $n$  çift ise  $n = 2k$ , eğer  $n$  tek ise  $n = 2k + 1$  alınıp

$$b = \sum_{i=1}^k a_{(n-i+1)} (X_{(n-i+1)} - X_{(i)})$$

değeri hesaplanır. Burada,  $a_{(n-i+1)} = -a_{(i)}$  değerleri ilgili tabloda verilmiştir.

4.  $W_h = b^2/s^2$  değeri hesaplanır.

5. 0.01, 0.02, 0.05, 0.10, 0.50, 0.90, 0.95, 0.98 ve 0.99 yüzdelik değerleri  $W_{tablo}$  ile gösterilir ve ilgili tablodan bulunur.

Karar:

Eğer  $W_h < W_{tablo}$  ise  $H_0$  hipotezi reddedilir.

**Örnek:** Aşağıdaki verilerin normal dağılıma sahip olup olmadığını Shapiro Wilk testini kullanarak sınavınız.

$$X_i: 1.4172, 1.4897, 1.6302, -1.2075, -0.0631, 0.6715, \\ 0.7172, 0.7254, 1.4090, -0.2050, -0.1241, 0.7147$$

**Çözüm:**

1. Gözlemler küçükten büyüğe doğru sıralanır.

$$X_{(1)} = -1.2075 \quad X_{(7)} = 0.7172$$

$$X_{(2)} = -0.2050 \quad X_{(8)} = 0.7254$$

$$X_{(3)} = -0.1241 \quad X_{(9)} = 1.4090$$

$$X_{(4)} = -0.0631 \quad X_{(10)} = 1.4172$$

$$X_{(5)} = 0.6715 \quad X_{(11)} = 1.4897$$

$$X_{(6)} = 0.7147 \quad X_{(12)} = 1.6302$$

2)  $s^2$  değeri hesaplanır;

$$\bar{x} = \frac{1}{12} \sum_{i=1}^{12} X_i = 0.5979$$

$$s^2 = \sum_{i=1}^{12} (X_i - \bar{X})^2 = 8.1020$$

3)  $n = 12$  ise  $2k = 12$  olup  $k = 6$  dir. ( $n$  çift)

$$b = \sum_{i=1}^6 a_{(12-i+1)} (X_{(12-i+1)} - X_{(i)}) = -2.7174$$

olarak hesaplanır.

$a$  değerleri ilgili tablodan

$$a_{12} = -0.5475, a_{11} = -0.3325,$$

$$a_{10} = -0.2347, a_9 = -0.1586,$$

$$a_8 = -0.0922, a_7 = -0.0303$$

olarak bulunur.

4)  $W_h = \frac{b^2}{s^2} = 0.9114$  şeklinde hesaplanır.

5)  $n = 12$  durumunda

$\alpha = 0.10$  için  $W_{tablo} = 0.883$  ve

$\alpha = 0.05$  için  $W_{tablo} = 0.859$

dır.

$W_h < W_{tablo}$  ise  $H_0$  reddedilir.

Burada,

$0.9114 > 0.883$  olduğundan  $H_0$  hipotezi reddedilemez. ( $\alpha = 0.10$  için)

Burada,

$0.9114 > 0.859$  olduğundan  $H_0$  hipotezi reddedilemez. ( $\alpha = 0.05$  için)

Sonuç: Veriler normal dağılıma uymaktadır.

**Örnek:** Aşağıdaki verilerin normal dağılıma sahip olup olmadığını Shapiro Wilk testini kullanarak sınavınız.

$X_i$ : 0.4889, 6.0347, 0.7269, -0.3034, 0.2939, -0.7873  
0.8884, -1.1471, -1.0689, -0.8095, -2.9443, 6.4384, 0.3252

**Çözüm:**

1. Gözlemler küçükten büyüğe doğru sıralanır.

$$X_{(1)} = -2.9443 \quad X_{(7)} = 0.2939$$

$$X_{(2)} = -1.1471 \quad X_{(8)} = 0.3252$$

$$X_{(3)} = -1.0689 \quad X_{(9)} = 0.4889$$

$$X_{(4)} = -0.8095 \quad X_{(10)} = 0.7269$$

$$X_{(5)} = -0.7873 \quad X_{(11)} = 0.8884$$

$$X_{(6)} = -0.3034 \quad X_{(12)} = 6.0347$$

$$X_{(13)} = 6.4384$$

2.  $s^2$  değeri hesaplanır.

$$\bar{x} = \frac{1}{13} \sum_{i=1}^{13} X_i = 0.6258$$

$$s^2 = \sum_{i=1}^{13} (X_i - \bar{X})^2 = 87.0215$$

3)  $n = 13$  ise  $2k + 1 = 13$  olup  $k = 6$  dir. ( $n$  tek)

$$\begin{aligned} b &= a_{(13)}(X_{(13)} - X_{(1)}) + a_{(12)}(X_{(12)} - X_{(2)}) + a_{(11)}(X_{(11)} - X_{(3)}) \\ &\quad + a_{(10)}(X_{(10)} - X_{(4)}) + a_{(9)}(X_{(9)} - X_{(5)}) + a_{(8)}(X_{(8)} - X_{(6)}) \\ &= -8.3246 \end{aligned}$$

olarak hesaplanır.

$a$  değerleri ilgili tablodan

$$a_{(13)} = -0.5359, a_{(12)} = -0.3325,$$

$$a_{(11)} = -0.2412, a_{(10)} = -0.1707,$$

$$a_{(9)} = -0.1099, a_{(8)} = -0.0539$$

olarak bulunur.

4)  $W_h = \frac{b^2}{s^2} = 0.7963$  şeklinde hesaplanır.

5)  $n = 13$  durumunda,

$\alpha = 0.10$  için  $W_{tablo} = 0.889$  ve

$\alpha = 0.05$  için  $W_{tablo} = 0.866$

dır.

$W_h < W_{tablo}$  ise  $H_0$  reddedilir.

Burada,

$0.7963 < 0.889$  olduğundan  $H_0$  reddedilir. ( $\alpha = 0.10$  için)

Burada,

$0.7963 < 0.866$  olduğundan  $H_0$  reddedilir. ( $\alpha = 0.05$  için)

Sonuç: Veriler normal dağılıma uymamaktadır.