

# İSTATİSTİK

## 4. Ki-Kare ( $\chi^2$ ) Testi

# Ki-Kare Testi Ne İçin Kullanılır?

1. Sayısal olmayan (kategorik veya sıralı) iki değişken arasında bir ilişki olup olmadığını anlamak istediğimizde (BAĞIMSIZLIK TESTİ),
2. Bir veri kümesinin belirli bir istatistik dağılımına uyup uymadığını anlamak istediğimizde (İYİ UYUM TESTİ),
3. İki yada daha fazla örneklemin aynı ana kütleden çekilip çekilmediğini anlamak istediğimizde (HOMOJENLİK TESTİ)

Ki-Kare testi kullanılır

# Ki-Kare Testlerinde Hipotezler

## ▶ BAĞIMSIZLIK TESTİ:

- ▶  $H_0$ : İki değişken arasında ilişki yoktur
- ▶  $H_a$ : İki değişken arasında ilişki vardır

## ▶ İYİ UYUM TESTİ:

- ▶  $H_0$ : Veri kümesi iddia edilen istatistik dağılımına uyar
- ▶  $H_a$ : Veri kümesi iddia edilen istatistik dağılımına uymaz

## ▶ HOMOJENLİK TESTİ:

- ▶  $H_0$ : İki veri kümesi de aynı örneklemden çekilmiştir
- ▶  $H_a$ : İki veri kümesi farklı örneklemelerden çekilmiştir

# Ki-Kare Testi Nasıl Uygulanır

- Verilerin gözlemlenen ve beklenen değerleri hesaplanıp KONTENJANS Tablosunda gösterilir. Herhangi bir hücrenin beklenen değeri 5'ten küçük olmamalıdır!
- Kontenjans tablosundaki her bir hücre için Ki-Kare değerine katkısı hesaplanır  $[(G-B)^2/B]$ . Daha sonra bu değerler toplanarak Ki-Kare değeri bulunur.
- Bulunan toplam değer, belirlenen serbestlik derecesi  $[sd=(\text{satır}-1) \times (\text{sütun}-1)]$  ve anlamlılık seviyesine ( $\alpha=0,05$  veya  $\alpha=0,01$ ) göre Ki-Kare tablo değeri ile karşılaştırılır.
- Hesaplanan değer tablo değerinden büyükse  $H_0$  reddedilir (değişkenler arası ilişki vardır).

# Kontenjans Tablosu

- ▶ İncelenen iki değişkenin farklı kategorilerine düşen gözlenen frekansların yazıldığı satır ve sütunlardan oluşan bir tablodur. İki değişkene ait kategoriler satır ve sütunlarda gösterilir. Hücrelerde ise o kategorilere ait gözlemlenen ve beklenen değerler yazılır.
- ▶ Örneklem gözlemlerini sınıflandırmak için kullanılır
- ▶ Çapraz tablolar olarak da adlandırılır

# $\chi^2$ Bağımsızlık Testi

Ki-Kare testi:

$$\chi^2 = \sum \frac{(f_g - f_b)^2}{f_b}$$

■ burada:

$f_g$  = r x c tablosunda belirli bir hücrede gözlemlenen frekans

$f_b$  =  $H_0$  doğruysa beklenen frekans

$\chi^2$  r x c tablosu için  $\chi^2$  nin serbestlik derecesi  $(r-1) \times (c-1)$

# Beklenen Hücre Frekansları

- Beklenen hücre frekansları:

$$f_b = \frac{\text{Satır toplamı} \times \text{Sütun toplamı}}{n}$$

burada:

satır toplamı = Satırdaki bütün frekansların toplamı

sütun toplamı = Sütundaki bütün frekansların toplam

n = toplam örneklem büyüklüğü

# Karar Kuralı

► Karar Kuralı

$$\chi^2 > \chi^2_{\dagger} \text{ ise } H_0 \text{ red, değilse kabul}$$

# Örnek

- 200 öğrenci ve tercih ettikleri aktiviteler

| Sınıf  | Aktivite |      |         | Toplam |
|--------|----------|------|---------|--------|
|        | Gezi     | Spor | Eğlence |        |
| 1      | 24       | 32   | 14      | 70     |
| 2      | 22       | 26   | 12      | 60     |
| 3      | 10       | 14   | 6       | 30     |
| 4      | 14       | 16   | 10      | 40     |
| Toplam | 70       | 88   | 42      | 200    |

# Örnek

► Test edilecek hipotez

$H_0$ : Tercih edilen aktivite ve devam edilen sınıf bağımsızdır

$H_1$ : Tercih edilen aktivite ve devam edilen sınıf bağımlıdır

# Örnek: Beklenen Hücre frekansları

Gözlemlenen:

| Sınıf  | Aktivite |      |      | Toplam |
|--------|----------|------|------|--------|
|        | Gezi     | Spor | Eğl. |        |
| 1      | 24       | 32   | 14   | 70     |
| 2      | 22       | 26   | 12   | 60     |
| 3      | 10       | 14   | 6    | 30     |
| 4      | 14       | 16   | 10   | 40     |
| Toplam | 70       | 88   | 42   | 200    |

Bir hücre için örnek:

$$f_b = \frac{(sat\_top) \times (sut\_top)}{n}$$
$$= \frac{30 \times 70}{200} = 10.5$$

$H_0$  doğru iken beklenen hücre frekansları

| Sınıf  | Aktivite |      |      | Toplam |
|--------|----------|------|------|--------|
|        | Gezi     | Spor | Eğl. |        |
| 1      | 24.5     | 30.8 | 14.7 | 70     |
| 2      | 21.0     | 26.4 | 12.6 | 60     |
| 3      | 10.5     | 13.2 | 6.3  | 30     |
| 4      | 14.0     | 17.6 | 8.4  | 40     |
| Toplam | 70       | 88   | 42   | 200    |

# Örnek: Test İstatistiği

► Test istatistiğinin değeri:

$$\begin{aligned}\chi^2 &= \sum \frac{(f_g - f_b)^2}{f_b} \\ &= \frac{(24 - 24.5)^2}{24.5} + \frac{(32 - 30.8)^2}{30.8} + \dots + \frac{(10 - 8.4)^2}{8.4} = 0.709\end{aligned}$$

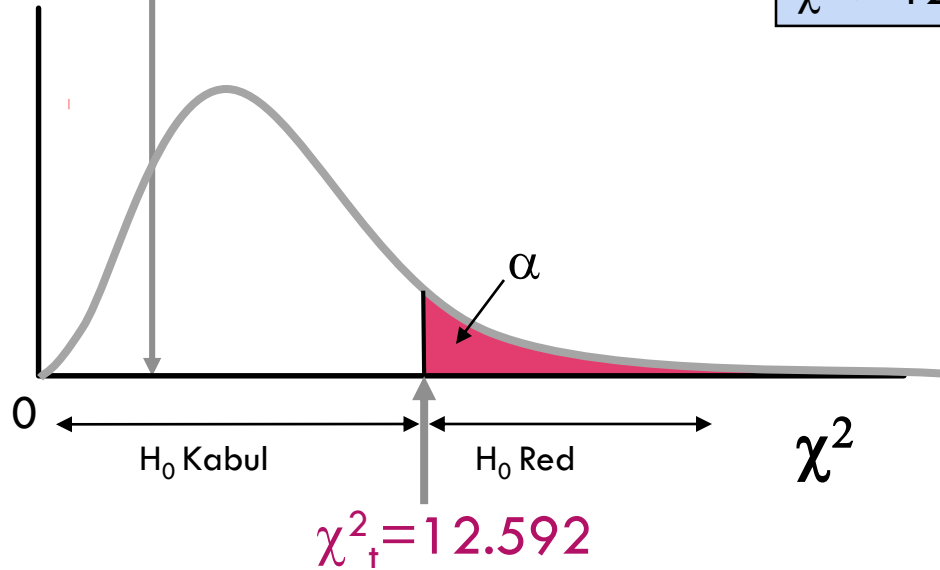
$\alpha = 0.05$  ve  $(4 - 1)(3 - 1) = 6$  serbestlik derecesi için  $\chi^2_{\alpha} = 12.592$

# Örnek: Karar ve Yorum

Test istatistiği  $\chi^2 = 0.709$ ,  $\chi^2_t = 12.592$  (s.d. = 6)

**Karar kuralı:**

$\chi^2 > 12.592$  ise,  $H_0$  red, değilse  $H_0$  kabul



Burada,

$\chi^2 = 0.709 < \chi^2_t = 12.592$ ,

Dolayısıyla  $H_0$  kabul

**Sonuç:** tercih edilen aktivite ile devam edilen sınıf arasında bir ilişki olduğuna dair  $\alpha = 0.05$  seviyesinde yeterince kanıt yoktur