

## BÖLÜM 5. $\chi^2$ (Kİ-KARE) ÇÖZÜMLEMESİ

$\chi^2$  çözümlemesi; beklenen değer ile gözlenen değer arasındaki farkın araştırılması için kullanılır.(Aralarındaki fark anlamlı mı?)

Örneğin; Bir para atma deneyinde

$$y \rightarrow 1/2 \qquad T \rightarrow 1/2$$

olasılıkla ortaya çıkar.

$n = 100$  kez parayı atalım.

Beklenen	50Y	50T	(gözlenen-beklenen)
Gözlenen	45Y	55T	
	-5	5	fark vardır.

$\pm 5$  birimlik farkın istatistiksel olarak anlamlı olup olmadığı  $\chi^2$  çözümlemesi ile test edilir.

$\chi^2$  çözümlemesinin uygulandığı konular;

- a) Sıklık dağılımları ile olasılık dağılımları arasındaki farkın araştırılmasında  $\chi^2$  istatistiği kullanılır. Buna “Uyum İyiliği Testleri”denir.
- b) Sayımla belirtilen kitlelerde bağımsızlık kontrolünde kullanılır.

### ÇAPRAZ TABLOLAR

#### a) Tek Gruplu Çapraz Tablo

Genellikle uyum iyiliği testlerinde kullanılır. Burada incelenen grup belirli bir özelliğe göre düzeylere ayrılmıştır.

Sınıflar	Sıklıklar
1	$f_1$
2	$f_2$
3	$f_3$
4	$f_4$ gibi.

Kullanılan test istatistiği;

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - f'_i)^2}{f'_i}$$

dir. Burada;

$\chi^2$  : Ki-kare değeri

$f_i$  : Gözlenen (deneysel) sıklık

$f'_i$  : Beklenen (kuramsal) sıklık

k : Sınıf sayısı

$$\sum_{i=1}^k f_i = \sum_{i=1}^k f'_i = n$$

dir.

**b)  $R \times C$  Çapraz Tablosu**

R : Satır sayısı      C : Sütun sayısı

$\begin{matrix} \text{X} \\ \text{Y} \end{matrix}$	1	2	...	C	Toplam
1	$f_{11}$	$f_{12}$	...	$f_{1C}$	$n_{1.}$
2	$f_{21}$	$f_{22}$	...	$f_{2C}$	$n_{2.}$
$\vdots$	$\vdots$				$\vdots$
R	$f_{R1}$	$f_{R2}$	...	$f_{RC}$	$n_{R.}$
Toplam	$n_{.1}$	$n_{.2}$	...	$n_{.C}$	$n$

$$\sum_{i=1}^R n_{i.} = \sum_{j=1}^C n_{.j} = n$$

$$\sum_{i=1}^R \sum_{j=1}^C f_{ij} = \sum_{i=1}^R \sum_{j=1}^C f'_{ij} = n$$

$$\chi^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}}$$

$f_{ij}$  :gözlenen sıklık

$f'_{ij}$  :beklenen sıklık

Örneğin; lise mezunlarının okuma alışkanlığı üzerinde yapılan bir araştırmada çapraz tablo aşağıdaki gibi olabilir.

	Kitap	Gazete	Dergi	Bilimsel yayın	Toplam
Okuyan	$f_{11}$	$f_{12}$	$f_{13}$	$f_{14}$	$n_{1.}$
Okumayan	$f_{21}$	$f_{22}$	$f_{23}$	$f_{24}$	$n_{2.}$
Toplam	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	$n$

$$n_{.1} + n_{.2} + n_{.3} + n_{.4} = n_{1.} + n_{2.} = n$$

$R = 2, C = 4 \rightarrow 2 \times 4$  tipinde bir çapraz tablo

### c) $2 \times 2$ Çapraz Tablo

$R = 2, C = 2$  olduğu durumdur.

$$\chi^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}}$$

ya da

X \ Y	Hasta	Hasta değil	Toplam
Sigara içiyor	$f_{11} = \mathbf{a}$	$f_{12} = \mathbf{b}$	$n_{1.} = \mathbf{A}$
Sigara içmiyor	$f_{21} = \mathbf{c}$	$f_{22} = \mathbf{d}$	$n_{2.} = \mathbf{B}$
Toplam	$n_{.1} = \mathbf{C}$	$n_{.2} = \mathbf{D}$	$n$

olmak üzere  $2 \times 2$  çapraz tablo için;

$$\chi^2 = \frac{(ad - bc)^2 n}{ABCD}$$

ile hesaplanır.

**Not:** Bu formül ile beklenen sıklıkların hesaplanması gerekmiyor. Sadece gözlenen sıklıklar kullanılır.

## Özellikler

$\chi^2$  çözümlemesinde aşağıdaki özelliklere dikkat etmek gerekir:

- 1) Yapılacak tahminlerin güvenilir olması için beklenen sıklıkların en az 5 olması istenir. Eğer 5'ten az beklenen (kuramsal) sıklık değeri varsa, bu sıklığın yer aldığı satır ya da sütun tablodaki uygun bir satır ya da sütun ile birleştirilir.
- 2) **Yates Düzeltmesi:** Uygulamada genellikle denek sayısı  $n < 50$  ve serbestlik derecesi 1 olduğunda Yates düzeltmesi yapılması gerektiği söylenir. Bu durumda  $\chi^2$  değişkeninin düzeltilmiş ve düzeltilmemiş değerlerinin karşılaştırılmasında yarar vardır; eğer her iki  $\chi^2$  değeri birbirinden çok farklı ise denek sayısı yetersizdir, örneklem genişliği artırılmalıdır.

(a) için;

$$\chi^2 = \sum_{i=1}^k \frac{(|f_i - f'_i| - 0.5)^2}{f'_i}$$

(b) ve (c) için;

$$\chi^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(|f_{ij} - f'_{ij}| - 0.5)^2}{f'_{ij}}$$

(c) için;

$$\chi^2 = \frac{(|ad - bc| - n/2)^2 n}{ABCD}$$

- 3) i.inci satırdaki gözlenen sıklık ile i.inci satırdaki beklenen sıklık toplamaları birbirine eşit olmalıdır. Aynı şekilde j.inci sütundaki gözlenen sıklık ile j.inci satırdaki beklenen sıklık toplamaları birbirine eşit olmalıdır.
- 4)  $2 \times 2$  düzeninde çapraz tabloda 5'den az sayıda gözlenen sıklık varsa, önemlilik testi Fisher'in özel eşitliği uygulanarak yapılır.

## Sayımla Belirtilen Kitlelerde Bağımsızlık Kontrolü

İki ya da daha çok grubun bağımsız olup olmadığı kontrol edilir. Veriler çapraz tablo üzerinde gösterilebilir.

1. Hipotez kurulur.

$H_0$ : Değişkenler bağımsızdır.(Gruplar arası fark yoktur.)

$H_1$ : Değişkenler bağımlıdır.(Gruplar arası fark vardır.)

## 2. Test istatistiđi belirlenir.

$R \times C$  tablosu düzenlenir ve  $f'_{ij}$  kuramsal sıklığı (i inci satır, j inci sütuna düşmesi gereken sıklık) şöyle bulunur:

$$f'_{ij} = (p_{i.})(p_{.j})n$$

Burada,

$$p_{i.} = \frac{n_{i.}}{n} \quad \text{ve} \quad p_{.j} = \frac{n_{.j}}{n}$$

dir. Buradan,

$$f'_{ij} = \frac{(n_{i.})(n_{.j})}{n}, \quad i = 1, \dots, R(\text{satır}) \quad , \quad j = 1, \dots, C(\text{sütun})$$

$$\chi^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}}$$

$R$ :satır sayısı

$C$ :sütun sayısı

$f_{ij}$ :gözlenen sıklık

Yates Düzeltmeli Formülü

$$\chi_H^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(|f_{ij} - f'_{ij}| - 0.5)^2}{f'_{ij}}$$

$2 \times 2$  düzeninde çapraz tablolar için;

$$\chi_H^2 = \frac{(ad-bc)^2 n}{ABCD}, \quad \text{Yates düzeltmeli} \quad \chi_H^2 = \frac{(|ad-bc| - n/2)^2 n}{ABCD}$$

## 3. Yorum ve karar

$\chi_H^2 > \chi_T^2$  ise  $H_0$  red edilir.

$\chi_{T(\alpha, s'=(R-1)(C-1))}^2$  ile bulunur.

Bir değişkenin ikiden çok düzeyi olduğunda ya da ikiden çok grup karşılaştırıldığında, gruplar arası fark önemli ise, gruplardan hangisinin diğerlerinden önemli derecede farklı olduğunu araştırmak gerekir. Bu nedenle en büyük  $\chi^2$  değerine sahip olan grup işleminden çıkarılır. Geriye kalan gruplar için, hipotez kabul edilinceye kadar işlemlere devam edilir.

**Örnek 5.1.** Bir sağlık idarecisi difteri-boğmaca karma aşısı satın alacaktır. Piyasada ayrı firmalara ait 4 aşı vardır ve idareci en etkin olanını seçmek istemektedir. Bunun için bir araştırma yaparak bütün aşıları uygulamış ve sonuçları şöyle bulmuştur. Aşılar arasında fark var mıdır  $\alpha = 0.05$  anlam düzeyinde test ediniz.

	Korunan		Korunmayan		
Aşı	$f_{ij}$	$f'_{ij}$	$f_{ij}$	$f'_{ij}$	Toplam
1	82	$\frac{123 \times 245}{372} = 82$	41	$\frac{123 \times 127}{372} = 42$	123
2	70	$\frac{94 \times 245}{372} = 62$	24	$\frac{94 \times 127}{372} = 32$	94
3	45	$\frac{65 \times 245}{372} = 43$	20	$\frac{65 \times 127}{372} = 22$	65
4	48	$\frac{90 \times 245}{372} = 59$	42	$\frac{90 \times 127}{372} = 31$	90
Toplam	245	245	127	127	372

1.  $H_0$ : Koruyuculuk yönünden aşılar arasında fark yoktur.  
 $H_1$ : Koruyuculuk yönünden aşılar arasında fark vardır.

2. Her bir satır için  $\chi^2$  değeri bulunur.

$$\chi_1^2 = \frac{(82-81)^2}{81} + \frac{(41-42)^2}{42} = 0.036$$

$$\chi_2^2 = \frac{(70-62)^2}{62} + \frac{(24-32)^2}{32} = 3.032$$

$$\chi_3^2 = \frac{(45-43)^2}{43} + \frac{(20-22)^2}{22} = 0.275$$

$$\chi_4^2 = \frac{(48-59)^2}{59} + \frac{(42-31)^2}{31} = 5.954$$

$$\chi_H^2 = \chi_1^2 + \chi_2^2 + \chi_3^2 + \chi_4^2 = 0.036 + 3.032 + 0.275 + 5.954 = 9.297$$

3.  $\chi_{T(0.05,(4-1) \times (2-1)=3)}^2 = 7.815$

$\chi_H^2 > \chi_T^2$  olduğundan  $H_0$  red edilir.

**Yorum:** Aşılar arasında fark olduğu %95 güvenilirlikle söylenebilir.

Şimdi bu farklılığın hangi aşıdan kaynaklandığına bakılmalıdır. Bunun için en büyük  $\chi^2$  değerine sahip olan dördüncü aşı analiz dışı bırakılarak, diğer üç aşı arasında fark olup olmadığına bakılır.

Aşı	Korunan	Korunmayan	Toplam
1	82(86)	41(37)	123
2	70(66)	24(28)	94
3	45(45)	20(20)	65
<b>Toplam</b>	197	85	282

$$\chi_1^2 = \frac{(82-86)^2}{86} + \frac{(41-37)^2}{37} = 0.6185$$

$$\chi_2^2 = \frac{(70-66)^2}{66} + \frac{(24-28)^2}{28} = 0.8139$$

$$\chi_3^2 = \frac{(45-45)^2}{45} + \frac{(20-20)^2}{20} = 0$$

$$\chi_H^2 = 0.6185 + 0.8139 = 1.4324$$

$$\chi_{T(0.05,(3-1) \times (2-1)=2)}^2 = 5.99$$

$\chi_H^2 < \chi_T^2$  olduğundan  $H_0$  red edilemez. 1, 2, 3 nolu aşılar arasında fark yoktur.

**Örnek 5.2** 100 deney faresi iki gruba ayrılıyor. Birinci gruptaki 53 deney faresine bakteri ve daha sonra standart dozda anti serum veriliyor. İkinci gruptaki 47 fareye ise yalnız bakteri veriliyor. Belirli bir süre geçtikten sonra 81 fare canlı kalıyor. Bakteri ve anti serum verilen 8 fare ölüyor.%5 anlam düzeyinde farelerin ölümünün anti serum etkisinden bağımsız olup olmadığını araştırınız.

	Ölen	Canlı kalan	Toplam
<b>Bakteri ve anti serum</b>	8 10.07	45 42.93	53
<b>Bakteri</b>	11 8.93	36 38.07	47
<b>Toplam</b>	19	81	100

1.  $H_0$ : Fare ölümleri anti serum etkisinden bağımsızdır.  
 $H_1$ : Fare ölümleri anti serum etkisinden bağımsız değildir.

$$2. \chi_H^2 = \frac{(8-10.07)^2}{10.07} + \frac{(45-42.93)^2}{42.93} + \frac{(11-8.93)^2}{8.93} + \frac{(36-38.07)^2}{38.07} = 1.1198$$

$$3. \chi_{T(0.05, sd=1)}^2 = 3.84$$

$\chi_H^2 < \chi_T^2$  olduğundan  $H_0$  red edilemez.

**Yorum:** Fare ölümlerinin anti serum etkisinden bağımsız olduğu %95 güvenilirlikle söylenebilir.

**Örnek 5.3** Rasgele olarak seçilen 200 hasta cinsiyet ve kandaki kolestrol miktarına göre aşağıdaki tabloda özetlenmiştir.

	Kandaki Kolestrol Miktarı			
Cinsiyet	Normal Değerlerden Küçük	Normal Değerler İçinde	Normal Değerlerden Büyük	Toplam
Kadın	10(30)	55(50)	35(20)	100
Erkek	50(30)	45(50)	5(20)	100
Toplam	60	100	40	200

$\alpha = 0.01$  anlam düzeyinde cinsiyet ile kandaki kolestrol miktarı arasında ilişki olup olmadığını test ediniz.

1.  $H_0$ : Cinsiyet ile kandaki kolestrol miktarı arasında ilişki yoktur.  
 $H_1$ : Cinsiyet ile kandaki kolestrol miktarı arasında ilişki vardır.

$$2. \chi_H^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}}$$

$$= \frac{(10-30)^2}{30} + \frac{(55-50)^2}{50} + \dots + \frac{(5-20)^2}{20} = 50.167$$

$$3. \chi_{T(\alpha, sd=(R-1)(C-1)=0.01, (2-1) \times (3-1)=2)}^2 = 9.21$$

$\chi_H^2 > \chi_T^2$  olduğundan  $H_0$  red edilir.

**Yorum:** Cinsiyet ile kandaki kolestrol miktarı arasında bir ilişkinin olduğu %99 güvenilirlikle söylenebilir.



**Örnek 5.4** Çocukların davranışı normal ve patolojik olarak nitelendirildiğinde aşağıdaki verilere göre gözlemciler arası farklılığı araştırınız. ( $\alpha = 0.05$ )

<b>Psikiyatrist B \ Psikiyatrist A</b>	<b>Normal</b>	<b>Patolojik</b>	<b>Toplam</b>
<b>Normal</b>	$10 = a$	$8 = b$	$18 = A$
<b>Patolojik</b>	$20 = c$	$9 = d$	$29 = B$
<b>Toplam</b>	$30 = C$	$17 = D$	$47 = n$

1.  $H_0$ : İki gözlemci arasında fark yoktur.

$H_1$ : İki gözlemci arasında fark vardır.

$n < 50$ ,  $sd = 1$  olduğundan Yates düzeltmesi yapılır.

$$\chi_H^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(|f_{ij} - f'_{ij}| - 0.5)^2}{f'_{ij}}$$

ya da

$$\chi_H^2 = \frac{(|ad - bc| - n/2)^2 n}{ABCD}$$

$$= \frac{\left(|10 \times 9 - 8 \times 20| - \frac{47}{2}\right)^2 \times 47}{18 \times 29 \times 30 \times 17} = \frac{2162.25 \times 47}{18 \times 29 \times 30 \times 17} = 0.382$$

$$\chi_{T(0.05, (2-1) \times (2-1)=1)}^2 = 3.84$$

$\chi_H^2 < \chi_T^2$  olduğundan  $H_0$  red edilemez.