BAĞIMSIZLIK VE HOMOJENLİK İÇİN Kİ-KARE TESTLERİ

BÖLÜM 8

1

Bağımsızlık İçin Ki-Kare Testi

İkinci değişkenin düzeyleri	Birinci değişkenin düzeyleri						Toplam
	1	2	⋯	j	⋯	c	
1	n_{11}	n_{12}		n_{1j}		n_{1c}	$n_{1.}$
2	n_{21}	n_{22}		n_{2j}		n_{2c}	$n_{2.}$
⋮	⋮	⋮		⋮		⋮	⋮
i	n_{i1}	n_{i2}		n_{ij}		n_{ic}	$n_{i.}$
⋮	⋮	⋮		⋮		⋮	⋮
r	n_{r1}	n_{r2}		n_{rj}		n_{rc}	$n_{r.}$
Toplam	$n_{.1}$	$n_{.2}$	⋯	$n_{.j}$	⋯	$n_{.c}$	n

H_0: Değişkenler bağımsızdır.
H_1: Değişkenler bağımsız değildir.

H_0: Değişkenler arasında ilişki yoktur.
H_1: Değişkenler arasında ilişki vardır.

2

Bağımsızlık İçin Ki-Kare Testleri

$$\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{\left(G_{ij}-B_{ij}\right)^2}{B_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

c : Sütun değişkeninin düzey sayısı
r : Satır değişkeninin düzey sayısı

Karar kuralı	
$\chi_h^2 > \chi^2_{(r-1)(c-1),\alpha}$	ise H_0 reddedilir
$\chi_h^2 \leq \chi^2_{(r-1)(c-1),\alpha}$	ise H_0 reddedilemez

3

Örnek

	Uyruk					
Harcama Düzeyi ($/Gün)	**Alman**	**İngiliz**	**İtalyan**	**Amerikalı**	**Diğer**	**Toplam**
50'den az	30	30	60	20	10	150
51-100	40	20	15	40	20	135
100'den fazla	30	10	5	60	10	115
Toplam	100	60	80	120	40	400

H_0: Harcama düzeyi ve uyruk değişkenleri bağımsızdır.
H_1: Harcama düzeyi ve uyruk değişkenleri bağımsız değildir.

H_0: Harcama düzeyi ve uyruk arasında ilişki yoktur.
H_1: Harcama düzeyi ve uyruk arasında ilişki vardır.

4

Örnek

Harcama Düzeyi ($/Gün)	Uyruk: Alman	İngiliz	İtalyan	Amerikalı	Diğer	Toplam
50'den az	30 (37.50)	30 (22.50)	60 (30.00)	20 (45.00)	10 (15.00)	150
51-100	40 (33.75)	20 (20.25)	15 (27.00)	40 (40.50)	20 (13.50)	135
100'den fazla	30 (28.75)	10 (17.25)	5 (23.00)	60 (34.50)	10 (11.50)	115
Toplam	100	60	80	120	40	400

$$\chi_h^2 = 95.4171$$

$\chi_h^2 > \chi^2_{(r-1)(c-1),\alpha}$, yani $95.4171 > 15.507$,

5

Cramer V

$$v = \sqrt{\frac{\chi_h^2}{n \cdot \min\{(c-1),(r-1)\}}}$$

v	:	İlişki katsayısının tahmini
n	:	Örnek hacmi
$min\{(c-1),(r-1)\}$	:	$(c-1)$ ve $(r-1)$'den küçük olan

6

Örnek

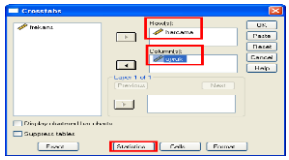
$$v = \sqrt{\frac{95.4171}{400 \cdot min\{(5-1),(3-1)\}}}$$

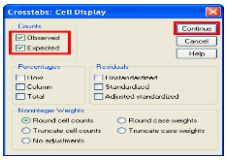
$$v = \sqrt{\frac{95.4171}{400(2)}}$$

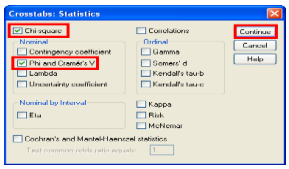
$$v = 0.345$$

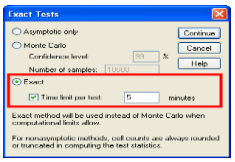
SPSS Uygulama

	harcama	uyruk	frekans
1	1	1	30
2	1	2	30
3	1	3	60
4	1	4	20
5	1	5	10
6	2	1	40
7	2	2	20
8	2	3	15
9	2	4	40
10	2	5	20
11	3	1	30
12	3	2	10
13	3	3	5
14	3	4	60
15	3	5	10









Harcama * Uyruk Crosstabulation

			Uyruk					
			Alman	İngiliz	İtalyan	Amerikalı	Diğer	Total
Harcama	50'den az	Count	30	30	60	20	10	150
		Expected Count	37,50	22,50	30,00	45,00	15,00	150,00
	51-100	Count	40	20	15	40	20	135
		Expected Count	33,75	20,25	27,00	40,50	13,50	135,00
	100'den fazla	Count	30	10	5	60	10	115
		Expected Count	28,75	17,25	23,00	34,50	11,50	115,00
Total		Count	100	60	80	120	40	400
		Expected Count	100,00	60,00	80,00	120,00	40,00	400,00

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	95,417[a]	8	,000	,[b]	
Likelihood Ratio	96,692	8	,000	,[b]	
Fisher's Exact Test	,[b]			,[b]	
Linear-by-Linear Association	6,246	1	,012	,[b]	,[b]
N of Valid Cases	400				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,50.

b. Cannot be computed because there is insufficient memory.

Örnek

	Tedavi yöntemi				
Hastanın durumu	**A**	**B**	**C**	**D**	**Toplam**
İyileşti	25	30	40	23	118
İyileşmedi	25	20	10	27	82
Toplam	50	50	50	50	200

H_0: Tedavi yöntemi ve tedavi sonrası durum değişkenleri bağımsızdır.
H_1: Tedavi yöntemi ve tedavi sonrası durum değişkenleri bağımsız değildir.

H_0: Tedavi yöntemi ve tedavi sonrası durum arasında ilişki yoktur.
H_1: Tedavi yöntemi ve tedavi sonrası durum arasında ilişki vardır.

9

Örnek

	Tedavi yöntemi				
Hastanın durumu	**A**	**B**	**C**	**D**	**Toplam**
İyileşti	25 (29.5)	30 (29.5)	40 (29.5)	23 (29.5)	118
İyileşmedi	25 (20.5)	20 (20.5)	10 (20.5)	27 (20.5)	82
Toplam	50	50	50	50	200

$$\chi_h^2 = 14.3034$$

$$\chi_h^2 > \chi_{(r-1)(c-1),\alpha}^2 \text{, yani } 14.3034 > 7.815,$$

10

Örnek

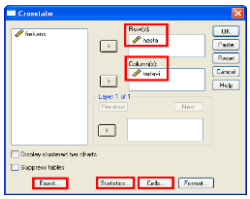
$$v = \sqrt{\frac{14.3034}{200 \cdot min\{(4-1),(2-1)\}}}$$

$$v = \sqrt{\frac{14.3034}{200}}$$

$$v = 0.2674$$

SPSS Uygulama

	hasta	tedavi	frekans
1	1	1	25
2	1	2	30
3	1	3	40
4	1	4	23
5	2	1	25
6	2	2	20
7	2	3	10
8	2	4	27



hasta * tedavi Crosstabulation

			tedavi				
			A	B	C	D	Total
hasta	İyileşti	Count	25	30	40	23	118
		Expected Count	29,5	29,5	29,5	29,5	118,0
	İyileşmedi	Count	25	20	10	27	82
		Expected Count	20,5	20,5	20,5	20,5	82,0
Total		Count	50	50	50	50	200
		Expected Count	50,0	50,0	50,0	50,0	200,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	14,303[a]	3	,003	,003		
Likelihood Ratio	15,093	3	,002	,002		
Fisher's Exact Test	14,777			,002		
Linear-by-Linear Association	,066[b]	1	,798	,848	,424	,049
N of Valid Cases	200					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 20,50.

b. The standardized statistic is -,257.

Homojenlik İçin Ki-Kare Testi

H_0: Örnekler aynı yığından gelmiştir.
H_1: Örnekler aynı yığından gelmemiştir.

$$\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{(G_{ij}-B_{ij})^2}{B_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

$\chi^2_h > \chi^2_{(r-1)(c-1),\alpha}$ ise H_0 hipotezi reddedilir.

13

Örnek

	Kalite durumu		
Makine	**Kaliteli**	**Kalitesiz**	**Toplam**
A	90	10	100
B	70	30	100
Toplam	160	40	200

H_0: Örnekler aynı yığından gelmiştir.
H_1: Örnekler aynı yığından gelmemiştir.

$$\chi^2_h = \frac{(90-80)^2}{80} + \frac{(70-80)^2}{80} + \frac{(10-20)^2}{20} + \frac{(30-20)^2}{20} = 12.5$$

14

Örnek

$\chi_h^2 > \chi^2_{(r-1)(c-1),\alpha}$, yani $12.5 > 3.841$,

olduğundan H_0 hipotezi reddedilir.

Bu örneklerin aynı yığından gelmedikleri %95 güvenle söylenebilir.

Bu problem için

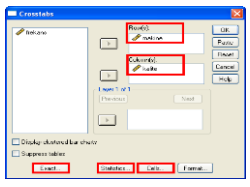
$$p\text{-değeri} = P(\chi_1^2 > 12.5)$$

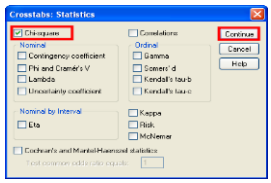
olur.

15

SPSS Uygulama

	makine	kalite	frekans
1	1	1	90
2	1	2	10
3	2	1	70
4	2	2	30





makine * kalite Crosstabulation

			kalite		
			Kaliteli	Kalitesiz	Total
makine	A	Count	90	10	100
		Expected Count	80,0	20,0	100,0
	B	Count	70	30	100
		Expected Count	80,0	20,0	100,0
Total		Count	160	40	200
		Expected Count	160,0	40,0	200,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	12,500[b]	1	,000	,001	,000	
Continuity Correction[a]	11,281	1	,001			
Likelihood Ratio	12,972	1	,000	,001	,000	
Fisher's Exact Test				,001	,000	
Linear-by-Linear Association	12,438[c]	1	,000	,001	,000	,000
N of Valid Cases	200					

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 20,00.

c. The standardized statistic is 3,527.

16