

FISHER'İN KESİN OLASILIK TESTİ

BİYOİSTATİSTİK VE TIP BİLİŞİMİ A.D.

Beklenen deęerlerinin 5 ten Kk Olması Durumu İin Uygulanır

Örnek hacmi küçük olduğunda hücrelerin beklenen deęerlerinin 5 ten büyük bulunması çok zordur, dolayısıyla χ^2 dağılışına iyi yaklaşım sağlanamaz.

Bu nedenle 2x2 lik tablolarda hücrelerden herhangi birisinin beklenen deęeri 5 den küçük ise Fisher'in kesin olasılık testi uygulanması, Süreklilik düzeltmesi yapılmış olan Person χ^2 yönteminin uygulanmasına tercih edilir.

Fisher' in testi, gözlenen 2x2 lik bir tablonun kesin önem seviyesinin hesaplanmasına dayanır.

- Bu teste göre sıra ve sütun marjinal (kenar) toplamaları ve (n) değişmemek koşulu ile tablo içi hücre değerlerinin değişik kombinasyonları için söz konusu olasılıkları hesaplayarak toplanır. Bu son elde edilen değer gözlenen ve daha ekstrem tabloların elde edilme olasılığıdır. Toplamı (n) olan hücreleri a,b,c,d olan bir tablonun elde edilme olasılığı:
- Eğer sıfır hipotezi doğru ise en fazla elde edilen tablo kadar **a** değeri elde etme olasılığı esas alınır.

		I. Özellik		
		Var	Yok	Σ
II. Özellik	Var	a	b	a+b
	Yok	c	d	c+d
	Σ	a+c	b+d	n

		I. Özellik		
		Var	Yok	Σ
II. Özellik	Var	a	b	a+b
	Yok	c	d	c+d
	Σ	a+c	b+d	n

Bağımsızlık hipotezi altında birinci hücredeki sayı (A) hipergeometrik dağılışı gösterir. A'nın değişim aralığı

$\max(0, a-d); \min(a+b; a+c)$

Olasılığında:

$$\text{Prob}(A = a) = \frac{(a+b)! * (c+d)! * (a+c)! * (b+d)!}{n! * a! * b! * c! * d!}$$

Tek yönlü önem düzeyi p_1 hesabı,

$$p_1 = \begin{cases} \Pr(A \geq a) & \text{eğer } a > E(A) \\ \Pr(A \leq a) & \text{eğer } a \leq E(A) \end{cases}$$

$$\text{Prob}(A = a) = \frac{(a + b)! * (c + d)! * (a + c)! * (b + d)!}{n! * a! * b! * c! * d!}$$

ile hesaplanır.

Birinci sıra ve birinci sütundaki (a) hücrelerinin değeri her seferinde 1 küçültülerek yeni tablo ve bunun elde edilme olasılığı hesaplanır. Bu işlem a'nın değeri sıfır oluncaya kadar devam edilir.

ÖRNEK UYGULAMA:

Perikardial efüzyon olan ve olmayan hastaların hastalıklarının sigara içme durumu ile ilişkisi araştırılıyor.

60 kişi ile yapılan bir denemede aşağıdaki tablo elde ediliyor.

		Sigara içme		
		Yok	Var	Σ
Efüzyon olması	Yok	2	23	25
	Var	5	30	35
	Σ	7	53	60

Fisher böyle bir tablonun elde edilme olasılığını hipergeometrik dağılışıla bulunacağını göstermiştir.

$$p = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)! * (c+d)! * (a+c)! * (b+d)!}{a! * b! * c! * d! * n!}$$

Burada; **!**: faktöriyel alınacağını ifade ediyor.

$$\binom{n}{k} = \frac{n!}{(n-k)! * k!} \text{ dir.}$$

: Bunlar, Binom katsayılarıdır.

Pratik olarak bu tablonun elde edilme olasılığı:

$$p = \frac{(a + b)! * (c + d)! * (a + c)! * (b + d)!}{a! * b! * c! * d! * n!}$$

- Orijinal tablodan en küçük hücre değeri esas alınarak bu sayı sıfıra indirgenene kadar yeni alt tablolar elde edilir.
- Bu örnek tablodan 2 farklı alt tablo daha elde edilebilir.
- Bunlar tablodaki en küçük hücre değeri **a=2** hücrelerinin değeri (**0**) değerine kadar indirilerek **a=1**, **a=0** için iki yeni tablo daha elde edilir. Yani orijinal tablo dahil olasılık hesaplanacak **3 tablo** olur. Tablolar elde edilirken yan toplamlar değişmeyecektir.

İlgili tablolar ve olasılık hesabı

Orijinal veri :

a=2 için alt tablo:

		Sigara içme		
		Yok	Var	Σ
Efüzyon olması	Yok	2	23	25
	Var	5	30	35
	Σ	7	53	60

$$\text{Prob}(A = a) = \frac{(25)! * (35)! * (7)! * (53)!}{60! * 2! * 23! * 5! * 30!} = 0.252$$

Yeni alt tablolar ve olasılık hesabı

a=1 için alt tablo:

		Sigara içme		
		Yok	Var	Σ
Efüzyon olması	Yok	1	24	25
	Var	6	29	35
	Σ	7	53	60

$$\text{Prob}(a,b,c,d) = \frac{(25)! * (35)! * (7)! * (53)!}{60! * 1! * 24! * 6! * 29!} = 0.105$$

a=0 için alt tablo:

		Sigara içme		
		Yok	Var	Σ
Efüzyon olması	yok	0	25	25
	var	7	28	35
	Σ	7	53	60

$$\text{Prob}(a,b,c,d) = \frac{(25)! * (35)! * (7)! * (53)!}{60! * 0! * 25! * 7! * 28!} = 0.017$$

- Bu olasılıklar toplanırsa, Fisher'in **tek yönlü** kesin olasılığı elde edilmiş olur.

Çift yönlü olasılık istenirse bunun iki katı alınır, yani $p=0.688$ olur.

- Her ne kadar efüzyon oranı $5/7=0.71$, $30/53=0.566$ olsa da istatistiksel olarak
- $P_{(fisher)} = (0.017+0.105+0.252) = \mathbf{0,374}$

0,374 > **$\alpha=0.05$** olduğundan

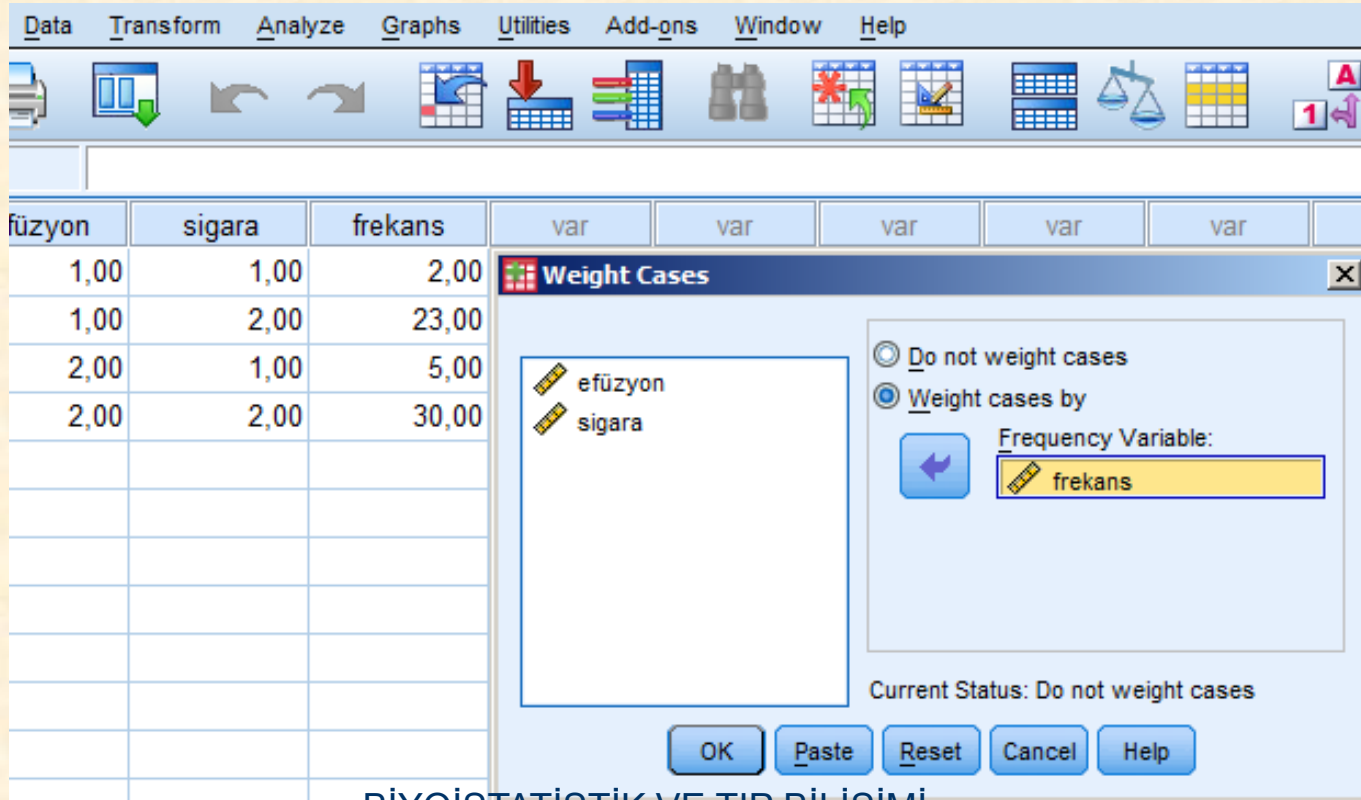
Karar: Sigara ile efüzyonun varlığı arasında bir ilişki olmadığı söylenebilir ($p=0.688$).

SPSS Çözümü

Önce frekansların tartılı olduğu belirtilir.

Veri girişinden sonra;

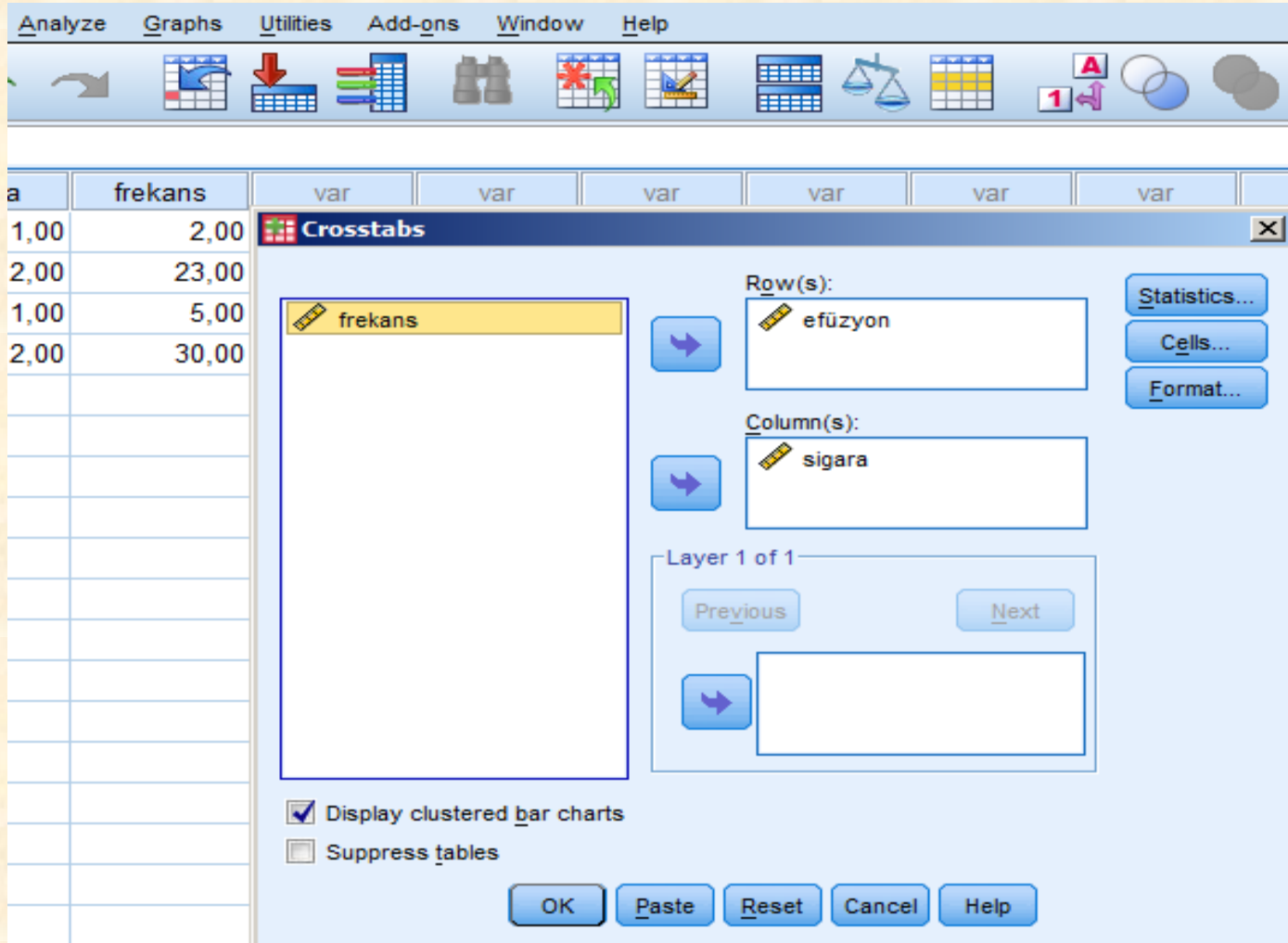
Data>Weight case> seçeneği kullanılır.



The screenshot shows the SPSS 'Weight Cases' dialog box. The 'Do not weight cases' option is unselected, and the 'Weight cases by' option is selected. The 'Frequency Variable' field contains 'frekans'. The current status is 'Do not weight cases'. The background shows a data table with columns 'füzyon', 'sigara', and 'frekans'.

füzyon	sigara	frekans
1,00	1,00	2,00
1,00	2,00	23,00
2,00	1,00	5,00
2,00	2,00	30,00

Analysis>Descriptives>Crosstabs seçilir ve açılan pencerelele ilgili deęişkenler konur.



The image shows the SPSS Crosstabs dialog box. The background is a spreadsheet with a table:

a	frekans	var	var	var	var	var	var
1,00	2,00						
2,00	23,00						
1,00	5,00						
2,00	30,00						

The Crosstabs dialog box is open, showing the following settings:

- Row(s): efüzyon
- Column(s): sigara
- Layer 1 of 1
- Display clustered bar charts
- Suppress tables

Buttons: OK, Paste, Reset, Cancel, Help, Statistics..., Cells..., Format...

Cells

Crosstabs: Cell Display

Counts

- Observed
- Expected

Percentages

- Row
- Column
- Total

Residuals

- Unstandardized
- Standardized
- Adjusted standardized

Noninteger Weights

- Round cell counts
- Round case weights
- Truncate cell counts
- Truncate case weights
- No adjustments

Continue Cancel Help

Statistics

Crosstabs: Statistics

Chi-square

Correlations

Nominal

- Contingency coefficient
- Phi and Cramer's V
- Lambda
- Uncertainty coefficient

Ordinal

- Gamma
- Somers' d
- Kendall's tau-b
- Kendall's tau-c

Nominal by Interval

- Eta

Kappa

Risk

McNemar

Cochran's and Mantel-Haenszel statistics

Test common odds ratio equals: 1

Continue Cancel Help

efüzyon * sigara Crosstabulation

		sigara		Total	
		ıçmıyor	ıçıyor		
Efüzyon	yok	Count	2	23	25
		Expected Count	2,9	22,1	25,0
		% within sigara	28,6%	43,4%	41,7%
	Var	Count	5	30	35
		Expected Count	4,1	30,9	35,0
		% within sigara	71,4%	56,6%	58,3%
		Count	7	53	60
		Expected Count	7,0	53,0	60,0
		% within sigara	100,0%	100,0%	100,0%

Çift yönlü test için

Çift yönlü test için

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,559 ^a	1	,455		
Continuity Correction ^b	,116	1	,734		
Likelihood Ratio	,581	1	,446		
Fisher's Exact Test				0,688	0,375
Linear-by-Linear Association	,550	1	,458		
N of Valid Cases	60				

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,92.

b. Computed only for a 2x2 table

P değeri, çift yönlü test

P değeri, tek yönlü test

2 hücrenin beklenen değeri
5 den küçük

- Aslında gamma veya log-gamma fonksiyonu kullanan daha hassas testler de vardır. Ancak standart istatistik paketlerde hipergeometrik, binomiyal dağılış gibi dağılışlara dayanan kesin olasılık testleri yer aldığından bunlar daha sık kullanılmaktadır.
- Kenar toplamların şansa bağılı olarak deęişmesi durumunda Barnard'ın kesin olasılık testi kullanılabilir. Fakat bu testle ilgili bazı eleştirel yönler vardır.