S1) Günlük içilen sigara miktarının belli yaş aralığına göre cinsiyetle arasında ilişki olup olmadığı araştırmak amacıyla bir çalışma yapılmıştır. Bu çalışmaya ilişkin olarak aşağıdaki bilgiler verilmiştir.

Günde 10'dan fazla sigara içen 70 kişi vardır.

20 yaş üstü olan 22 erkek günde 10'dan fazla sigara içiyor.

20 yaş ve altında olan erkek sayısı 35 iken, 20 yaş ve altında olan kadın sayısı da 25'dir.

20 yaş ve altı erkeklerin 18'i günde 10'dan fazla sigara içerken, 20 yaş ve altı kadınların 12'si günde 10'dan fazla sigara içmektedir.

20 yaş üstü erkeklerin 20'si günde 10'dan az sigara içerken, 20 yaş üstü kadınların 15'i günde 10'dan az sigara içmektedir.

a) Yukarıda verilen bilgilere ilişkin olarak çalışmada yer alan değişkenleri ve her bir değişkenin kategorilerini yazınız.

Degişkenler → Sigora Sayısı
→ Yaş
→ Cinsiyet

1 Sigora Sayısı → Günde 10'dan az, Günde 10'dan fazla

2 Yaş → 20 yaş ve altı, 20 yaş üstü

3 Cinsiyet → Kadın, Erkek

b) Yukarıda verilen bilgilere göre olumsallık (kontenjans) tablosunu oluşturunuz.

		Sigora Sayısı		
Cinsiyet	Yaş	Günde 10'dan az	Günde 10'dan fazla	Toplam
Erkek	20 yaş ve altı	17	18	35
	20 yaş üstü	20	22	42
Kadın	20 yaş ve altı	13	12	25
	20 yaş üstü	15	18	33
Toplam		65	70	135

S2) Aşağıda A ve B partisine oy veren kişilerin cinsiyetlere göre dağılımı verilmiştir.

Cinsiyet	A Partisi	B Partisi	Toplam
Kadın	28	48	76
Erkek	10	114	124
Toplam	38	162	200

Bu verilere ilişkin olarak;

a) Ortak olasılıkları bulunuz.

$$p_{11} = \frac{28}{200} = 0,14$$

$$p_{12} = \frac{48}{200} = 0,24$$

$$p_{21} = \frac{10}{200} = 0,05$$

$$p_{22} = \frac{114}{200} = 0,57$$

b) Marjinal olasılıkları bulunuz.

$$p_{1\cdot} = p_{11} + p_{12} = 0,14 + 0,24 = 0,38 \quad (\text{veya}) \quad p_{1\cdot} = \frac{76}{200} = 0,38$$

$$p_{2\cdot} = p_{21} + p_{22} = 0,05 + 0,57 = 0,62 \quad (\text{veya}) \quad p_{2\cdot} = \frac{124}{200} = 0,62$$

$$p_{\cdot 1} = p_{11} + p_{21} = 0,14 + 0,05 = 0,19 \quad (\text{veya}) \quad p_{\cdot 1} = \frac{38}{200} = 0,19$$

$$p_{\cdot 2} = p_{12} + p_{22} = 0,24 + 0,57 = 0,81 \quad (\text{veya}) \quad p_{\cdot 2} = \frac{162}{200} = 0,81$$

c) Cinsiyet bilindiğinde koşullu olasılıkları bulunuz.

$$P(\text{A partisi} / \text{Kadın}) = \frac{p_{11}}{p_{1\cdot}} = \frac{0,14}{0,38} = 0,368 \quad \left(\text{veya } \frac{28}{76} = 0,368\right)$$

$$P(\text{A partisi} / \text{Erkek}) = \frac{p_{21}}{p_{2\cdot}} = \frac{0,05}{0,62} = 0,08 \quad \left(\text{veya } \frac{10}{124} = 0,08\right)$$

$$P(\text{B partisi} / \text{Kadın}) = \frac{p_{12}}{p_{1\cdot}} = \frac{0,24}{0,38} = 0,631 \quad \left(\text{veya } \frac{48}{76} = 0,631\right)$$

$$P(\text{B partisi} / \text{Erkek}) = \frac{p_{22}}{p_{2\cdot}} = \frac{0,57}{0,62} = 0,919 \quad \left(\text{veya } \frac{114}{124} = 0,919\right)$$

S5) Aşağıda yeni doğan bebeklerin doğum ağırlığı ile doğumdan sonra ölüm olup olmamasına ilişkin olumsallık (kontenjans) tablosu bulunmaktadır.

	Ölüm	
Doğum Ağırlığı	**Var**	**Yok**
Çok Düşük ($< 1500\ gr$)	42	80
Düşük ($\gg 1500\ gr$)	43	302

Top. $\boxed{467}$

Bu verilere göre, doğum ağırlığı ile ölüm arasındaki bağımlılığı Üstünlük (Odds) oranı θ ile inceleyiniz. Ayrıca kitleye ilişkin $\acute{\theta}$ değerini %95 güvenle tahmin ediniz.

$$\theta = \frac{p_{11}\, p_{22}}{p_{12}\, p_{21}}$$

$$p_{11} = \frac{42}{467} = 0,089$$

$$p_{12} = \frac{80}{467} = 0,171$$

$$p_{21} = \frac{43}{467} = 0,092$$

$$p_{22} = \frac{302}{467} = 0,647$$

$$\theta = \frac{(0,089).(0,647)}{(0,171).(0,092)}$$

$$\theta = \frac{0,0576}{0,0157} = \underline{\underline{3,67}}$$

(ya da) $\theta = \frac{42(302)}{43(80)} \cong 3,68$

Doğum ağırlığı çok düşük olanların ölme olasılığı, doğum ağırlığı düşük olan ölme olasılığından yaklaşık 3,7 kat fazladır.

$$\ln\theta \mp z_{\alpha/2}\, s_{\ln\theta}$$

$$\ln\theta = \ln(3,67) = 1,3$$

$$s_{\ln\theta} = \sqrt{\frac{1}{f_{11}} + \frac{1}{f_{12}} + \frac{1}{f_{21}} + \frac{1}{f_{22}}} = \sqrt{\frac{1}{42} + \frac{1}{80} + \frac{1}{43} + \frac{1}{302}} = \sqrt{0,0628} = \underline{\underline{0,25}}$$

$$z_{\alpha/2} = z_{0,025} = 1,96$$

$$\left[1,3 \mp 1,96\,(0,25)\right] \Rightarrow [0,81\,;\,1,79]$$

$$e^{0,81} = 2,248$$

$$e^{1,79} = 5,989$$

$$P(2,248 < \hat{\theta} < 5,989) = 0,95$$

Kitleye ilişkin doğum ağırlığı ~~çok~~ çok düşük olanların ölme olasılığı, kitlede doğum ağırlığı düşük olanların ölme olasılığından yaklaşık en az 2,25, en fazla 5,99 katıdır.

S4) Rasgele olarak seçilen 134 hasta cinsiyet ve kandaki kolesterolü ölçme yöntemine göre aşağıdaki tabloda özetlenmiştir.

	Kolesterolü Ölçme Yöntemi			
Cinsiyet	**A**	**B**	**C**	Toplam
Kadın	10	37	12	59
Erkek	25	45	5	75
Toplam	35	82	17	134

α= 0,01 anlamlılık düzeyinde cinsiyet ile kandaki kolesterolü ölçme yönteminin bağımsız olup olmadığını test ediniz. (Hipotezleri ve adımları açıkça gösteriniz.)

Hipotezler:

H_0: Cinsiyet ile kandaki kolesterol miktarını ölçme yöntemi bağımsızdır.

H_1: " " " " " " " bağımsız değildir.

Test İstatistiği:

$$f'_{11} = \frac{59(35)}{134} = 15,4 \qquad f'_{21} = \frac{35(75)}{134} = 19,6$$

$$f'_{12} = \frac{59(82)}{134} = 36,1 \qquad f'_{22} = \frac{75(82)}{134} = 45,9$$

$$f'_{13} = \frac{59(17)}{134} = 7,5 \qquad f'_{23} = \frac{75(17)}{134} = 9,5$$

$$\chi^2 = \frac{(10-15,4)^2}{15,4} + \frac{(37-36,1)^2}{36,1} + \frac{(12-7,5)^2}{7,5} + \frac{(25-19,6)^2}{19,6} + \frac{(45-45,9)^2}{45,9} + \frac{(5-9,5)^2}{9,5} =$$

$$\chi^2 = 1,9 + 0,02 + 2,7 + 1,5 + 0,02 + 2,13 = \underline{\underline{8,27}}$$

Tablo Değeri:

$$\chi^2_{0,01;(2-1).(3-1)} = \chi^2_{0,01;2} = 9,21$$

Yorum:

$\chi^2_{hesap} = 8,27 < \chi^2_{0,01;2} = 9,21$ old. Ho hip reddedilemez.

Yani, %1 anlamlılıkla cinsiyet ile kandaki kolesterolü ölçme yöntemi bağımsızdır denir.

$$\frac{134}{2(2)} \geq 5$$

II. yol

$$G^2 = 2\sum f_{ij} \ln\left(\frac{f_{ij}}{f'_{ij}}\right) = 2.\left[10.\ln\left(\frac{10}{15,4}\right) + 37.\ln\left(\frac{37}{36,1}\right) + 12.\ln\left(\frac{12}{7,5}\right) + 25.\ln\left(\frac{25}{19,6}\right) + 45\ln\left(\frac{45}{45,9}\right) + 5.\ln\left(\frac{5}{9,5}\right)\right]$$