

BÖLÜM V Kİ KARE TESTLERİ

V.1 Bağımsızlık İçin Ki-Kare Testi

$r * c$ şeklindeki çapraz tablolarda sınıflama ya da sıralama düzeyinde ölçülen 2 veya daha fazla kategoriye sahip iki değişkenin (X ve Y gibi) kategorilerinin birbirleri ile bağımlı/bağımsız olup olmadıklarını ya da sınıflama veya sıralama düzeyinde ölçülen iki değişken arasındaki ilişkinin önemli olup olmadığını test etmek için kullanılır. Burada r : X değişkenine (birinci değişkene) ait kategori sayısı ($r \geq 2$) ve c : Y değişkenine (ikinci değişkene) ait kategori sayısı ($c \geq 2$) olsun. Örneklemdeki birim sayısı n olmak üzere, bu birimlerin değişkenlerin kategorilerine göre dağılımı, yani örneklem veri düzeni bir çapraz tablo ile verilebilir.

Tablo 5.1 Çapraz Tablo Veri Düzeni

	Y					
		1	2	...	c	Toplam($T_{i.}$)
X	1	G_{11}	G_{12}	...	G_{1c}	$T_{1.}$
	2	G_{21}	G_{22}	...	G_{2c}	$T_{2.}$

	r	G_{r1}	G_{r2}	...	G_{rc}	$T_{r.}$
	Toplam($T_{.j}$)	$T_{.1}$	$T_{.2}$...	$T_{.c}$	$T_{..} = n$

Bu veri düzeninde:

G_{ij} : X değişkeninin i . nci ve Y değişkeninin j . nci kategorisinde yer alan birim sayısı (gözlenen frekans) ($i = 1, 2, \dots, r; j = 1, 2, \dots, c$)

$T_{i.}$: X değişkeninin i . nci kategorisinde yer alan toplam birim sayısı

$$T_{i.} = \sum_{j=1}^c G_{ij}, (i = 1, 2, \dots, r) \quad (5.1)$$

$T_{.j}$: Y değişkeninin j . nci kategorisinde yer alan toplam birim sayısı

$$T_{.j} = \sum_{i=1}^r G_{ij}, (j = 1, 2, \dots, c) \quad (5.2)$$

$T_{..}$: Örneklemdeki toplam birim sayısı

$$T_{..} = \sum_{i=1}^r \sum_{j=1}^c G_{ij} = \sum_{i=1}^r T_{i.} = \sum_{j=1}^c T_{.j} = n \quad (5.3)$$

dir.

Test edilecek hipotezler:

H_0 : Değişkenler birbirinden bağımsızdır

H_1 : Değişkenler birbirinden bağımsız değildir. (5.4)

veya

H_0 : Değişkenler arasındaki ilişki önemsizdir

H_1 : Değişkenler arasındaki ilişki önemlidir. (5.5)

şeklinde oluşturulur. Test istatistiği ise

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(G_{ij} - B_{ij})^2}{B_{ij}} \sim \chi_{(r-1)(c-1)}^2 \quad (5.6)$$

şeklinde tanımlıdır. Burada

$$B_{ij} = \frac{T_{i.} \cdot T_{.j}}{n}, \quad i = 1, 2, \dots, r; j = 1, 2, \dots, c \quad (5.7)$$

formülü ile hesaplanan beklenen (teorik) frekanstır. Test istatistiğinin örneklemden hesaplanan değeri χ_h^2 olsun.

Karar α önem seviyesinde kritik değer $\chi_t^2 = \chi_{(r-1)(c-1)}^2; \alpha$ olmak üzere, eğer; $\chi_h^2 > \chi_t^2$ (yani $p < \alpha$) ise H_0 hipotezi ret edilir, $\chi_h^2 \leq \chi_t^2$ (yani $p \geq \alpha$) ise H_0 hipotezi ret edilemez.

SPSS uygulama algoritması

Bağımsızlık için Ki-Kare testinin SPSS uygulama algoritması veri düzenine göre farklılık gösterir.

Algoritma:1 (Veri düzeni $r * c$ çapraz tablo düzeninde bir frekans verisi ise)

Adım.1 Veri girişinde satır kategori kodları, sütun kategori kodları ve frekanslar farklı kolonlara girilir.

Adım.2 Frekanslar ağırlıklandırılır Bunun için **Data > Weight cases > Weight cases by** seçilip, frekans değişkeni **frequency variable** kutusuna aktarılır. **OK** tıklanır.

Adım.3 **Analyze > Descriptive Statistics > Crosstabs** seçenekleri seçilir ve yeni bir ekran açılır.

Adım.4 Açılan yeni ekranda değişkenler listesinden satır değişkeni (X) seçilerek **Row(s)** işlem kutusuna, sütun değişkeni (Y) seçilerek **Column(s)** işlem kutusuna aktarılır.

Adım.5 Ekranda yer alan **Exact...** penceresi açılır ve gelen ekrandan **Asymptotic only** seçeneği işaretlenir. **Continue** ile bu pencere kapatılır.

Adım.6 Ekranda yer alan **Statistics...** penceresi açılır ve **Chi-square** ve **Phi and Cramer's V** seçenekleri işaretlendikten sonra, **Continue** ile bu pencere kapatılır.

Adım.7 Ekranda yer alan **Cells...** penceresi açılır ve **Counts** kutusundan **Observed(gözlenen frekans)** ile **Expected(beklenen frekans)** tercihleri, **Percentages (yüzdelikler)** kutusundan **Row(satır)**, **Column(sütun)** ile **Total(toplam)** tercihleri yapıldıktan sonra, **Continue** ile bu pencere de kapatılır.

Adım.8 OK tıklanarak analiz bitirilir. Sonuçlar **Output** sayfasında tablolar halinde verilir. Bu tablolar değerlendirilerek yorumlar yapılır. Eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez

Algoritma:II (Veri düzeni tüm değişkenler için her bir birime ait ölçümler, yani ham veri düzeninde ise)

Adım.1 Her bir sütuna farklı bir değişken ve özellikleri tanımlandıktan sonra, tüm değişkenler için her bir birime ait değerler veri giriş sayfasında girilir.

Adım.2 Analyze > Descriptive Statistics > Crosstabs seçenekleri seçilir ve yeni bir ekran açılır.

Adım.3 Açılan ekranda değişkenler listesinden satır değişkeni (X) seçilerek **Row(s)** işlem kutusuna, sütun değişkeni/ değişkenleri seçilerek **Column(s)** işlem kutusuna aktarılır.

Adım.4 Ekranda yer alan **Exact...** penceresi açılır ve gelen ekrandan **Asymptotic only** seçeneği işaretlenir. **Continue** ile bu pencere kapatılır.

Adım.5 Ekranda yer alan **Statistics...** penceresi açılır ve **Chi-square** ve **Phi and Cramer's V** seçenekleri işaretlendikten sonra, **Continue** ile bu pencere kapatılır.

Adım.6 Ekranda yer alan **Cells...** penceresi açılır ve **Counts** kutusundan **Observed(gözlenen frekans)** ile **Expected(beklenen frekans)** tercihleri, **Percentages (yüzdelikler)** kutusundan **Row(satır)**, **Column(sütun)** ile **Total(toplam)** tercihleri yapıldıktan sonra, **Continue** ile bu pencere de kapatılır.

Adım.7 OK tıklanarak analiz bitirilir. Sonuçlar **Output** ayfasında tablolar halinde verilir. Bu tablolar değerlendirilerek yorumlar yapılır. Eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez

NOT: Ki-Kare analizlerinde her bir hücredeki beklenen frekanslar 5'den küçük olmamalıdır. Ya da büyük örneklemelerde beklenen frekansı 5'ten küçük olan hücre sayısı, tüm hücre sayısının %20'sini geçmemelidir. Eğer böyle hücrelerin sayısı %20'den fazla ise sınıflar (kategoriler) arasında bu sorunu giderecek şekilde uygun birleştirmeler yapılır. Eğer sınıf birleştirme olacaksa birbirine en yakın olan sınıflar birleştirilmelidir.

ÖRNEK 5.1 Stafilakok'a karşı yaygın olarak kullanılan 5 ilaç türünün etkinliklerini araştırmak için Stafilakok üremiş 213 vasat alınmış ve Antibiyogram yapılmıştır. Stafilakok'a karşı 5 antibiyotiğin etki düzeylerine (etki yok, düşük düzeyde etki, orta düzeyde etki ve yüksek düzeyde etki) göre 213 vasatın dağılımı için aşağıdaki çapraz tablo elde edilmiştir. İlaç türü ile etki düzeyi arasında ilişkinin önemli olup olmadığı araştırılmak istenmektedir. Buna göre;

- a) Değişkenleri ve özelliklerini belirtiniz?
- b) Test edilecek hipotezleri oluşturunuz?
- c) SPSS programı ile test işlemi için Ki-Kare testinin çözümünü yapınız?
- d) Elde edilen sonuca göre test istatistiğinin alabileceği değeri ve %1 önem seviyesinde, H_0 hipotezi hakkındaki kararınızı belirtiniz ve sonuçları yorumlayınız?

Tablo 5.2 Antibiyotik Türlerinin Etki Düzeylerine Göre Vasatların Dağılımı

Antibiyotik	Etki Düzeyleri				TOPLAM
	Yok(1)	Düşük(2)	Orta(3)	Yüksek(4)	
A(1)	9	11	16	4	40
B(2)	6	8	9	9	32
C(3)	21	19	6	10	56
D(4)	21	11	6	5	43
E(5)	6	10	15	11	42
TOPLAM	63	59	52	39	213

Cözüm: a) Satır Değişkeni: Antibiyotik türü....Nitel türden ve ölçme düzeyi sınıflama
Sütun Değişkeni: Etki Düzeyi....Nitel türden ve ölçme düzeyi sıralama

b) Test edilecek hipotezler;

H_0 : Stafilokok'a karşı Antibiyotik türü ile Etki düzeyleri arasındaki ilişki önemsizdir

H_1 : Stafilokok'a karşı Antibiyotik türü ile Etki düzeyleri arasındaki ilişki önemlidir

c) Veri düzeni $r = 5$ ve $c = 4$ olmak üzere $r * c$ çapraz tablo düzeninde olduğundan SPSS analizinde Algoritma-I uygulanır. Analiz sonucu Tablo 5.3'deki gibi elde edilir. Bu tablo veriye ilişkin çapraz tablodur. Tabloda gözlenen frekanslarla birlikte, beklenen (teorik) frekanslar, her bir antibiyotik türünün kendi içerisindeki yüzdelik değerleri, her bir etki düzeyinin kendi içerisindeki yüzdelik değerlerine ait bilgiler yer almaktadır.

Örneğin; 213 adet vasattan 40 tanesinde A antibiyotiği kullanıldığı ve bunların 9 tanesinde hiç etki olmadığı, 11 tanesinde düşük düzeyde etki olduğu, 16 tanesinde orta düzeyde etki olduğu ve 4 tanesinde de yüksek düzeyde etki olduğu gözlenmiştir.

Buna karşılık H_0 hipotezi doğru iken yani değişkenler arasındaki ilişki önemsiz ya da değişkenler bağımsız iken; A antibiyotiğinin kullanıldığı 40 vasattan 11,8 tanesinde hiç etki olmaması, 11,1 tanesinde düşük düzeyde etki olması, 9,8 tanesinde orta düzeyde etki olması ve 7,3 tanesinde de yüksek düzeyde etki olması beklenmektedir.

Tablo 5.3 Antibiyotik * Etkidüzeyi Çapraz Tablo Veri Düzeni ve Analizi

		etkidüzeyi				Total
		yok	düşük	orta	yüksek	
Antibiyotik	Count	9	11	16	4	40
	Expected Count	11,8	11,1	9,8	7,3	40,0
	A % within Antibiyotik	22,5%	27,5%	40,0%	10,0%	100,0%
	% within etkidüzeyi	14,3%	18,6%	30,8%	10,3%	18,8%
	% of Total	4,2%	5,2%	7,5%	1,9%	18,8%
	Count	6	8	9	9	32
	Expected Count	9,5	8,9	7,8	5,9	32,0
	B % within Antibiyotik	18,8%	25,0%	28,1%	28,1%	100,0%
	% within etkidüzeyi	9,5%	13,6%	17,3%	23,1%	15,0%
	% of Total	2,8%	3,8%	4,2%	4,2%	15,0%
	Count	21	19	6	10	56
	Expected Count	16,6	15,5	13,7	10,3	56,0
	C % within Antibiyotik	37,5%	33,9%	10,7%	17,9%	100,0%
	% within etkidüzeyi	33,3%	32,2%	11,5%	25,6%	26,3%
	% of Total	9,9%	8,9%	2,8%	4,7%	26,3%
	Count	21	11	6	5	43
	Expected Count	12,7	11,9	10,5	7,9	43,0
	D % within Antibiyotik	48,8%	25,6%	14,0%	11,6%	100,0%
	% within etkidüzeyi	33,3%	18,6%	11,5%	12,8%	20,2%
	% of Total	9,9%	5,2%	2,8%	2,3%	20,2%
	Count	6	10	15	11	42
	Expected Count	12,4	11,6	10,3	7,7	42,0
	E % within Antibiyotik	14,3%	23,8%	35,7%	26,2%	100,0%
	% within etkidüzeyi	9,5%	16,9%	28,8%	28,2%	19,7%
	% of Total	2,8%	4,7%	7,0%	5,2%	19,7%
Total	Count	63	59	52	39	213
	Expected Count	63,0	59,0	52,0	39,0	213,0
	% within Antibiyotik	29,6%	27,7%	24,4%	18,3%	100,0%
	% within etkidüzeyi	100,0%	100,0%	100,0%	100,0%	100,0%
	% of Total	29,6%	27,7%	24,4%	18,3%	100,0%

Ayrıca; A antibiyotiğinin kullanıldığı 40 vasatın $9/40= \%22,5$ inde hiç etki olmadığı, $11/40= \%27,5$ inde düşük düzeyde etki olduğu, $16/40= \%40$ ında orta düzeyde etki olduğu ve $4/40= \%10$ unda yüksek düzeyde etki olduğu anlaşılmaktadır.

Yine bu tabloya göre A-antibiyotik türü; hiç etki görülmeyen 63 vasatın $9/63=\%14,3$ ünde, düşük düzeyde etki görülen 59 vasatın $11/59=\%18,6$ sında, orta düzeyde etki görülen 52 vasatın $16/52=\%30,8$ inde ve yüksek düzeyde etki görülen 39 vasatın $4/39=\%10,3$ ünde kullanılmıştır.

Diğer antibiyotik türleri için de benzer değerlendirmeler yapılabilir.

Diğer taraftan; 5 antibiyotik türünün kullanıldığı 213 vasattan 63 ünde hiç etki görülmediği, 59 unda düşük düzeyde etki görüldüğü, 52 sinde orta düzeyde etki görüldüğü ve 39 unda ise yüksek düzeyde etki görüldüğü anlaşılmaktadır. Diğer bir ifadeyle; 5 antibiyotik türünün kullanıldığı 213 vasattan $63/213=\%29,6$ sinde hiç etki görülmediği, $59/213=\%27,7$ sinde düşük düzeyde etki görüldüğü, $52/213=\%24,4$ ünde orta düzeyde etki görüldüğü ve $39/213=\%18,3$ inde ise yüksek düzeyde etki görüldüğü söylenebilir.

Son olarak; 213 vasattan 40 tanesinde A-antibiyotiği, 32 tanesinde B-antibiyotiği, 56 tanesinde C-antibiyotiği, 43 tanesinde D-antibiyotiği ve 42 tanesinde E-antibiyotiği kullanılmıştır. Diğer bir ifadeyle; A-antibiyotiğinin tüm vasatların $40/213=\%18,8$ inde, B-antibiyotiğinin tüm vasatların $32/213=\%15$ inde, C-antibiyotiğinin tüm vasatların $56/213=\%26,3$ ünde, D-antibiyotiğinin tüm vasatların $43/213=\%20,2$ sinde ve E-antibiyotiğinin tüm vasatların $42/213=\%19,7$ sinde kullanıldığı söylenebilir.

Ki-Kare testi analiz sonuçları Tablo 5.4 ile İlişki katsayısı değeri ise Tablo 5.5’de verildi.

Tablo 5.4 Ki-Kare Testi Sonuçları

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	31,277 ^a	12	,002
Likelihood Ratio	31,922	12	,001
Linear-by-Linear Association	,003	1	,953
N of Valid Cases	213		

Tablo 5.5 İlişki Katsayıları Değerleri (Symmetric Measures)

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	,394	,001	. ^c
	Cramer's V	,228	,001	. ^c
	Contingency Coefficient	,367	,001	. ^c
N of Valid Cases		213		

d) Test istatistiğinin alabileceği değer Chi-square test tablosuna göre $\chi_h^2 = 31,277$ ve Asymp.Sig.= $2p=0,002$ olup buradan $p=0,001$ dir. $\alpha = 0,01$ önem seviyesi için $p < \alpha$ olup H_0 ret edilir. Bu sonuca göre antibiyotik türü ile etki düzeyleri arasındaki ilişki önemli ve bu iki

değişken bağımsız değildir. Bu iki değişken arasındaki ilişki önemli olduğundan, söz konusu ilişkinin derecesi Symmetrics Measures tablosuna göre $v = 0,228$ dir.

NOT: Veri düzeni tüm değişkenler için her bir birime ait ölçümler düzeninde ise bu durumda analiz için Algoritma-II uygulanır. Örnek 5.1’de verilen çapraz tablo verisi bu düzende hazırlanarak Örnek 5.2 Algoritma-II ile çözüldüğünde yukarıdaki sonuçların aynısı elde edildi.

ÖRNEK 5.2 Örnek5.1’de verilen çapraz tablo verisini, iki değişkene ait gözlemleri 213 birimin hepsi için düzenleyerek değişkenler arasındaki ilişkinin önemli olup olmadığını Algoritma-II’yi uygulayarak Ki-Kare analizi ile test ediniz?

Çözüm: Sonuçlar Örnek 5.1’in çözümü ile aynıdır.

ÖRNEK 5.3 Rastgele seçilen 400 yabancı uyruklu Türkiye’de kaldıkları süre içerisinde yaptıkları harcama(\$ ve uyrukları sorulmuş ve alınan cevaplardan aşağıdaki veri düzenlemesi yapılmıştır. Buna göre Türkiye’de kalan yabancı uyrukluların, yaptıkları harcama ile uyrukları arasındaki ilişkinin önemli olup olmadığını SPSS programı ile analiz ederek %5 önem düzeyinde karar veriniz ve ilişki ölçüsünü hesaplayınız?

Tablo 5.6 Yabancı Uyruklu*Harcama Veri Düzeni

Harcama Düzeyi (\$)	UYRUK				
	Alman	İngiliz	İtalyan	Amerikalı	Diğer
50’den az	30	30	60	20	10
51-100	40	20	15	40	20
100’den fazla	30	10	5	60	10

Çözüm: Satır Değişkeni: Harcama düzeyi....Nicel türden ve ölçme düzeyi sıralama ($r = 3$)

Sütun Değişkeni: Uyruk....Nitel türden ve ölçme düzeyi sınıflama ($c = 5$)

Test edilecek hipotezler;

H_0 : Uyruk ve harcama değişkenleri arasındaki ilişki önemsizdir

H_1 : Uyruk ve harcama değişkenleri arasındaki ilişki önemlidir

Veri düzeni $r = 3$ ve $c = 5$ olmak üzere $r * c$ çapraz tablo düzeninde olduğundan SPSS analizinde Algoritma-I uygulanır. Analiz sonucu Tablo 5.7’deki gibi elde edilir.

Tablo 5.7 Harcama * Uyruk Çapraz Tablo Veri Düzeni ve Analizi

		uyruk					Total
		alman	ingiliz	italyan	amerikalı	diğer	
harcama	Count	30	30	60	20	10	150
	Expected Count	37,5	22,5	30,0	45,0	15,0	150,0
	50 denaz % within harcama	20,0%	20,0%	40,0%	13,3%	6,7%	100,0%
	% within uyruk	30,0%	50,0%	75,0%	16,7%	25,0%	37,5%
	% of Total	7,5%	7,5%	15,0%	5,0%	2,5%	37,5%
	Count	40	20	15	40	20	135
	Expected Count	33,8	20,3	27,0	40,5	13,5	135,0
	51-100 % within harcama	29,6%	14,8%	11,1%	29,6%	14,8%	100,0%
	% within uyruk	40,0%	33,3%	18,8%	33,3%	50,0%	33,8%
	% of Total	10,0%	5,0%	3,8%	10,0%	5,0%	33,8%
	Count	30	10	5	60	10	115
	Expected Count	28,8	17,3	23,0	34,5	11,5	115,0
	100 den fazla % within harcama	26,1%	8,7%	4,3%	52,2%	8,7%	100,0%
	% within uyruk	30,0%	16,7%	6,2%	50,0%	25,0%	28,7%
	% of Total	7,5%	2,5%	1,2%	15,0%	2,5%	28,7%
Total	Count	100	60	80	120	40	400
	Expected Count	100,0	60,0	80,0	120,0	40,0	400,0
	% within harcama	25,0%	15,0%	20,0%	30,0%	10,0%	100,0%
	% within uyruk	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
	% of Total	25,0%	15,0%	20,0%	30,0%	10,0%	100,0%

Bu tablo veriye ilişkin çapraz tablodur. Tabloda gözlenen frekanslarla birlikte, beklenen (teorik) frekanslar, her bir harcama grubunun kendi içerisindeki yüzdelik değerlerine, her bir uyruk için kendi içerisindeki yüzdelik değerlerine ve her bir ortak grup için genele göre yüzde değerlerine ait bilgiler yer almaktadır.

Örneğin; 400 yabancı uyrukludan 150 tanesinin 50\$ dan az harcadığı ve bunların 30'unun Alman, 30'unun İngiliz, 60'ının İtalyan, 20'sinin Amerikalı ve 10'unun da diğer uyruklara mensup olduğu gözlenmiştir. Harcama için diğer gruplarda benzer şekilde değerlendirilebilir. Ayrıca; 400 yabancı uyrukludan 100 tanesinin Alman uyruklu olduğu ve bunların 30 tanesinin Türkiye'de kaldıkları sürede 50\$' dan az, 40 tanesinin 51-100\$ arasında ve 30 tanesinin ise 100\$'dan fazla harcama yaptıkları gözlenmiştir. Diğer uyruk gruplarında da benzer şekilde değerlendirmeler yapılabilir

Buna karşılık H_0 hipotezi doğru iken yani harcama ve uyruk değişkenleri arasındaki ilişki önemsiz ya da değişkenler bağımsız iken; 50\$'dan daha az harcama yapan 150 yabancı

uyrukludan 37,5'unun Alman, 22,5'unun İngiliz, 30'unun İtalyan, 45'inin Amerikalı ve 15'inin de diğer uyruklara mensup olması beklenmektedir. Harcama için diğer gruplarda da beklenen frekanslar benzer şekilde değerlendirilebilir. Ayrıca; 100 Alman uyrukludan 37,5'unun 50\$' dan az, 33,8'inin 51-100\$ arasında ve 28,8'inin ise 100\$'dan fazla harcama yapması beklenir. Beklenen frekanslarla ilgili olarak diğer uyruklar için de benzer değerlendirmeler yapılabilir.

Diğer taraftan; 50\$'ın altında harcama yapan 150 yabancı uyrukludan %20 si Alman, %20 si İngiliz, %40'ı İtalyan, %13,3'ü Amerikalı ve %6,7'si de diğer uyruklara mensuptur. Harcama kategorisinin diğer sınıflarında da yüzdeler üzerinden benzer değerlendirmeler yapılabilir. Ayrıca 100 Almandan %30'u 50\$'dan az, %40'ı 51-100\$ arasında ve %30'u da 100\$'dan fazla harcama yapmışlardır. Benzer şekilde diğer uyruklar içinde yüzdeler yönünden değerlendirmeler yapılabilir

Yine bu tabloya göre 100 Alman'dan %30'u, 60 İngilizden %50'si, 80 İtalyan'dan %75'i, 120 Amerikalı'dan %16,7'si ve 40diğer uyrukludn %25'i 50\$'ın altında harcama yapmışlardır. Diğer harcama grupları içerisinde yabancı uyrukluların yüzdeleri de benzer şekilde bulunabilir.

Diğer taraftan; 400 yabancı uyrukludan 100'ü Alman, 60'ı İngiliz, 80'i İtalyan, 120'si Amerikalı ve 40'ı da diğer uyruklularadan oluşmaktadır. Diğer bir ifadeyle; 400 yabancı uyrukludan $100/400=\%25$ ' Alman. $60/400=\%15$ 'i İngiliz, $80/400=\%20$ 'si İtalyan, $120/400=\%30$ 'u Amerikalı ve $40/400=\%10$ 'u da diğer uyruklulardan oluşmaktadır.

Son olarak; 400 yabancı uyrukludan 150'si 50\$'ın altında, 135'i 51-100\$ arasında ve 115'i de 100\$'ın üstünde harcama yapmışlardır. Diğer bir ifadeyle; 50\$'ın altında harcama yapanlar tüm yabancı uyrukluların $150/400=\%37,5$ 'üğünü, 51\$-100\$ arasında harcama yapanlar tüm yabancı uyrukluların $135/400=33,8$ 'ini ve 100\$'ın üstünde harcama yapanlar tüm yabancı uyrukluların $115/400=\%28,7$ 'sini oluşturmaktadır.

Son olarak, Ki-Kare testi analiz sonuçları Tablo 5.8'de ve İlişki katsayısı değeri ise Tablo 5.9'da verildi.

Tablo 5.8 Ki-Kare Testi Sonuçları

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	95,417 ^a	8	,000
Likelihood Ratio	96,692	8	,000
Linear-by-Linear Association	6,246	1	,012
N of Valid Cases	400		

Tablo 5.9 İlişki Katsayıları Değerleri (Symmetric Measures)

		Value	Approx. Sig.
Nominal by Nominal	Phi	,488	,000
	Cramer's V	,345	,000
N of Valid Cases		400	

Test istatistiğinin alabileceği değer Ki-Kare testi sonuç tablosuna göre $\chi_h^2 = 95,417$ ve Asymp. Sig.= $2p=0,000$ olup buradan $p=0,0000$ dir. $\alpha = 0,01$ önem seviyesi için $p < \alpha$ olup H_0 ret edilir. Bu sonuca göre uyruk ile harcama miktarı arasındaki ilişki önemlidir ve bu iki değişken bağımsız değildir. Bu iki değişken arasındaki ilişki önemli olduğundan, söz konusu ilişkinin derecesi Symmetrics Measures tablosuna göre $v = 0,345$ dir.

V.2 Homojenlik İçin Ki-Kare Testi

Bölüm IV.6'da verilen iki bağımsız grup karşılaştırmalarında bağımlı değişken nicel, sürekli ve ölçme düzeyi eşit aralıklı veya oranlamadır. Bu durumda normal dağılım varsayımının sağlanıp sağlanmamasına göre kullanılacak olan istatistiksel analizler o bölümde tanıtıldı. Bu analizlerde amaç iki bağımsız örneklemin (grubun) aynı kitleden gelip gelmediğini araştırmaktır. Bazı iki bağımsız grup karşılaştırmalarında bağımlı değişkenin ölçme düzeyi sınıflama ya da sıralama düzeyinde olabilir. Örneğin; rastgele seçilen bir gruptaki hastalara A ilacı ve yine rastgele seçilen bir diğer gruptaki hastalara da B ilacı verilerek uygulanan tedavi süreci sonunda, bu iki bağımsız gruptaki hastalar “iyileşti” veya “iyileşmedi” şeklinde değerlendirilerek 2*2 düzeninde bir veri düzenlemesi yapılabilir. Bu veri düzeni Tablo 5.10'daki gibi verilir.

Tablo 5.10 2*2 Veri Düzeni

İlaç Türü	Tedavi Sonucu	
	İyileşti	İyileşmedi
A	G_{11}	G_{12}
B	G_{21}	G_{22}

Bu tür problemlerde bu iki bağımsız örneğin (A ilacı ile tedavi edilenler ve B ilacı ile tedavi edilenler) aynı kitleden gelip gelmediğinin belirlenmesi için kullanılan istatistiksel analize Homojenlik (birimlerin benzerliği) için Ki-Kare testi denir.

Homojenlik için Ki-Kare testinde test edilecek hipotezler:

H_0 : Örnekler aynı kitleden gelmiştir (Birimler homojen dağılımlıdır)

H_1 : Örnekler aynı kitleden gelmemiştir (Birimler homojen dağılımlı değildir)

şeklinde oluşturulur. Bu hipotezleri test etmek için kullanılacak olan test istatistiği, Bağımsızlık için Ki-Kare testinde kullanılan test istatistiği ile aynıdır. Yani;

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

Test istatistiğinin örneklemden hesaplanan değeri χ_h^2 olsun.

Karar α önem seviyesinde kritik değer $\chi_t^2 = \chi_{(r-1)(c-1)}^2; \alpha$ olmak üzere, eğer; $\chi_h^2 > \chi_t^2$ (yani $p < \alpha$) ise H_0 hipotezi ret edilir, $\chi_h^2 \leq \chi_t^2$ (yani $p \geq \alpha$) ise H_0 hipotezi ret edilemez.

Homojenlik için Ki-Kare testinin SPSS uygulama algoritması da Bağımsızlık için Ki-Kare testinde verilen algoritma/algoritmalar ile aynıdır.

ÖRNEK 5.4 A ve B makineleri tarafından üretilen ürünlerden rastgele seçilen 100'er birimlik numunelerin kalite durumuna göre dağılımı aşağıda verilmiştir. Bu örneklerin aynı kitleden geldikleri, yani A ve B makineleri tarafından üretilen ürünlerin kalite durumuna göre dağılımlarının aynı olduğu %5 önem seviyesinde söylenebilir mi?

Tablo 5.11 Makine* Kalite Durumu

Makine	Kalite Durumu	
	Kaliteli(1)	Kalitesiz(2)
A(1)	90	10
B(2)	70	30

Cözüm: Satır Değişkeni: Üretim yapılan makine....Nitel türden ve ölçme düzeyi sınıflama ($r = 2$)

Sütun Değişkeni: Kalite durumu....Nitel türden ve ölçme düzeyi sınıflama ($c = 2$)

Test edilecek hipotezler;

H_0 : Örnekler aynı kitleden gelmiştir (A ve B makinelerinde üretilen ürünler kalite yönünden homojen (benzer) dağılımlıdır)

H_1 : Örnekler aynı kitleden gelmemiştir (A ve B makinelerinde üretilen ürünler kalite yönünden homojen (benzer) dağılımlı değildir)

Veri düzeni $r = 2$ ve $c = 2$ olmak üzere $r * c$ çapraz tablo düzeninde olduğundan SPSS analizinde Algoritma-I uygulanır. Analiz sonuçları Tablo 5.12 -5.13 ve 5.14'de verildi.

Tablo 5.12 Makine * Kalite Durumu Çapraz Tablo Veri Düzeni ve Analizi

		kalite		Total
		kaliteli	kalitesiz	
makine	A makinesi			
	Count	90	10	100
	Expected Count	80,0	20,0	100,0
	B makinesi			
	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

Tablo 5.13 Ki-Kare Testi Sonuçları

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

Test istatistiğinin alabileceği değer Chi-square tests tablosuna göre $\chi_h^2 = 12,500$ ve Exact Sig. (1-sided)= $p=0,0$ dir. $\alpha = 0,01$ önem seviyesi için $p < \alpha$ olup H_0 ret edilir. Bu sonuca göre örnekler aynı kitleden gelmemiştir. Yani A makinesi ait üretim ile B makinesi ait üretim kalite yönünden homojen (benzer) değildir. A makinesine ait üretim ile B makinesine ait üretimin kalite yönünden aralarındaki ilişki $v=0,25$ olup, $p = 0.00 < \alpha$ olduğundan önemli fakat zayıf bir ilişkidir.

Tablo 5.14 İlişki Katsayıları Değerleri (Symmetric Measures)

	Value	Approx. Sig.
Nominal by Nominal		
Phi	,250	,000
Cramer's V	,250	,000
N of Valid Cases	200	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

V.3 Uyum İyiliği (Goodness of Fit) İçin Ki-Kare Testi

Sınıflama ya da sıralama düzeyinde ölçülen bir değişkenin kitledeki dağılımı m tane sınıf/kategoriden oluşsun. Böyle bir kitleden rasgele çekilen n birimlik bir örnekleme ait veri düzeni Tablo 5.15’de verildiği gibi olacaktır. Böyle bir veri düzeninde kategorilere ait gözlenen frekansların $(G_j, j = 1, 2, \dots, m)$ dağılımının, belli bir teorik olasılık dağılımı yardımı ile hesaplanacak olan beklenen frekanslar $(B_j, j = 1, 2, \dots, m)$ ile uyumlu olup olmadığı araştırılmak istenebilir. Diğer bir ifade ile gözlenen frekanslara ait dağılımın, belli bir olasılık dağılımı ile uyumlu olup olmadığı Ki-Kare testi ile test edilebilir. Bu amaçla uygulanan Kİ-Kare testine uyum iyiliği için Ki-Kare testi adı verilir.

Tablo 5.15 Uyum İyiliği Testi İçin Veri Düzeni

Sınıf/Kategori	1	2	...	m	Toplam
Gözlenen Frekans (G_j)	G_1	G_2	...	G_m	$n = \sum_{j=1}^m G_j$
P_j	P_1	P_2	...	P_m	$\sum_{j=1}^m P_j = 1$
Beklenen Frekans (B_j)	B_1	B_2	...	B_m	$n = \sum_{j=1}^m B_j$

Burada $j = 1, 2, \dots, m$ için P_j : herhangi bir gözlemin j . sınıfa düşme olasılığı olup, H_0 : hipotezinde öngörülen teorik olasılık dağılımı yardımı ile hesaplanabilir. Test edilecek hipotezler;

H_0 : Gözlenen frekanslar ile beklenen frekanslar uyumludur (Gözlenen frekansların dağılımı öngörülen teorik olasılık dağılımı ile uyumludur)

H_1 : Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir (Gözlenen frekansların dağılımı öngörülen teorik olasılık dağılımı ile uyumlu değildir)

şeklinde ifade edilir. H_0 : hipotezi altında hesaplanacak olan P_j olasılıkları birbirine eşit, yani $P_1 = P_2 = \dots = P_m$ olabileceği gibi birbirinden farklı da olabilirler. Bu durumda yine H_0 : hipotezi altında hesaplanacak olan beklenen frekanslar $B_j = n * P_j, (j = 1, 2, \dots, m)$ eşitliği ile bulunur. Test istatistiği ise gözlenen frekanslar ile beklenen frekanslar arasındaki uyumun bir ölçüsü olarak ifade edilen Ki-Kare istatistiğidir ve

$$\chi^2 = \sum_{j=1}^m \frac{(G_j - B_j)^2}{B_j} \sim \chi_{m-k-1}^2$$

eşitliği ile verilir. Burada $k \geq 0$ olup, örneklemden tahmin edilen parametre sayısını gösteren bir doğal sayıdır.

Test istatistiğinin örneklemden hesaplanan değeri χ_h^2 olsun.

Karar α önem seviyesinde kritik değer $\chi_t^2 = \chi_{m-k-1}^2; \alpha$ olmak üzere, eğer; $\chi_h^2 > \chi_t^2$ (yani $p < \alpha$) ise H_0 hipotezi ret edilir, $\chi_h^2 \leq \chi_t^2$ (yani $p \geq \alpha$) ise H_0 hipotezi ret edilemez.

SPSS uygulama algoritması

Uyum iyiliği için Ki-Kare testinin SPSS uygulama algoritması veri düzenine göre farklılık gösterir.

Algoritma:I.1 (Veri düzeni m sınıflı/kategorili bir frekans verisi ise)

Adım.1 Veri girişinde sınıf/kategori kodları ve frekanslar farklı kolonlara girilir.

Adım.2 Frekanslar ağırlıklandırılır Bunun için **Data > Weight cases > Weight cases by** seçilip, frekans değişkeni **frequency variable** kutusuna aktarılır. **OK** tıklanır.

Adım.3 **Analyze > Nonparametric Tests > One Sample...** seçenekleri seçilir ve yeni bir ekran açılır.

Adım.4 Açılan yeni ekrandan **Fields** penceresi tıklanarak değişkenlerin listelendiği ekrana ulaşılır. Bu ekranda yer alan değişkenler listesinden **sınıf** değişkeni seçilerek aktarma butonu ile **Test Fields** işlem kutusuna aktarılır.

Adım.5 Aynı ekranda yer alan **Settings** penceresi açılır ve gelen ekrandan **Customize tests** seçeneği işaretlenir. Böylece aktif konuma gelen test seçeneklerinden **Compare observed probabilities to hypothesized (Chi- Square Test)** seçeneği işaretlenir ve **Options...** penceresi tıklanır. Açılan pencereden P_j sınıf olasılıklarının hesaplanmasına ilişkin seçenek belirlenir. Burada sunulan seçeneklerden birincisi bu olasılıkların tüm sınıflar için eşit olduğu durumunu kapsarken, ikincisi söz konusu olasılıkların birbirinden farklı olduğu durumunu kapsamaktadır. Eğer sınıf olasılıkları farklı ise her bir sınıf için olasılık değeri tercih yapıldıktan sonra aktif konuma gelen **Expected probabilities** işlem kutularına girilmek suretiyle tanımlanır.

Adım.6 **OK** tıklanarak analiz bitirilir. Sonuç **Output** sayfasında bir tablo ile verilir. Bu tablonun üzerine Mouse'un sağ tarafı ile çift tıklama yapıldığında analiz sonuçlarının yer aldığı daha detaylı bilgilere ulaşmak mümkündür. Bu bilgiler sınıflara göre gözlenen ve beklenen frekansların birlikte gösterildiği frekans histogramı grafiği ile test işlemi sonuçlarının yer aldığı bir tablodan oluşmaktadır. Bu tabloda verilen p olasılığı (önem seviyesi) dikkate alındığında, eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez.

Algoritma I (yani veri düzeni m sınıflı/kategorili bir frekans verisi ise) aşağıdaki adımları uygulamak suretiyle de uyum iyiliği için Ki-Kare testi yapılabilir:

Algoritma:I.2

Adım.1 Veri girişinde sınıf/kategori kodları ve frekanslar farklı kolonlara girilir.

Adım.2 Frekanslar ağırlıklandırılır Bunun için **Data > Weight cases > Weight cases by** seçilip, frekans değişkeni **frequency variable** kutusuna aktarılır. **OK** tıklanır.

Adım.3 **Analyze > Nonparametric Tests > Legacy dialogs > Chi-Square...** seçenekleri seçilir ve yeni bir ekran açılır.

Adım.4 Açılan yeni ekranda yer alan değişkenler listesinden **sınıf** değişkeni seçilerek aktarma butonu ile **Test Variable List** işlem kutusuna aktarılır.

Adım.5 Aynı ekranda yer alan **Expected Values** (beklenen frekanslar) bölümünde sunulan iki seçenekten uygun olanı tercih edilir. Burada sunulan seçeneklerden birincisi P_j olasılıkların ve böylece B_j beklenen frekanslarının tüm sınıflar için eşit olduğu durumu kapsarken, ikincisi söz konusu olasılıkların ya da beklenen frekansların birbirinden farklı olduğu durumu kapsamaktadır. Eğer beklenen frekanslar farklı ise her bir sınıf için beklenen frekanslar sırası ile **Values** seçeneği kullanılarak girilir.

Adım.6 **OK** tıklanarak analiz bitirilir. Sonuç **Output** sayfasında tablolar halinde verilir. Bu tablolardan birincisinde gözlenen frekanslar, beklenen frekanslar ve bunların farkları (yani $G_j - B_j$ farkları) her bir sınıf için yer alırken, ikincisinde Ki-Kare testine ait analiz sonuçları yer alır. Bu tabloda verilen p olasılığı (önem seviyesi) dikkate alındığında, eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez.

Algoritma:II.1 (Veri düzeni tüm değişkenler için her bir birime ait ölçümler, yani ham veri düzeninde ise)

Adım.1 Sınıflama veya sıralama düzeyinde ölçülen değişken ve özellikleri tanımlandıktan sonra, bu değişken için her bir birime ait değerler veri giriş sayfasında girilir.

Adım.2 **Analyze > Nonparametric Tests > One Sample...** seçenekleri seçilir ve yeni bir ekran açılır.

Adım.3 Açılan yeni ekrandan **Fields** penceresi tıklanarak değişkenlerin listelendiği ekrana ulaşılır. Bu ekranda yer alan değişkenler listesinden **sınıf** değişkeni seçilerek aktarma butonu ile **Test Fields** işlem kutusuna aktarılır.

Adım.4 Aynı ekranda yer alan **Settings** penceresi açılır ve gelen ekrandan **Customize tests** seçeneği işaretlenir. Böylece aktif konuma gelen test seçeneklerinden **Compare observed probabilities to hypothesized (Chi- Square Test)** seçeneği işaretlenir ve **Options...** penceresi tıklanır. Açılan pencereden P_j sınıf olasılıklarının hesaplanmasına ilişkin seçenek belirlenir. Burada sunulan seçeneklerden birincisi bu olasılıkların tüm sınıflar için eşit olduğu durumu kapsarken, ikincisi söz konusu olasılıkların birbirinden farklı olduğu durumu kapsamaktadır. Eğer sınıf olasılıkları farklı ise her bir sınıf için olasılık değeri tercih

yapıldıktan sonra aktif konuma gelen **Expected probabilities** işlem kutularına girilmek suretiyle tanımlanır.

Adım.5 OK tıklanarak analiz bitirilir. Sonuç **Output** sayfasında bir tablo ile verilir. Bu tablonun üzerine Mouse'un sağ tarafı ile çift tıklama yapıldığında analiz sonuçlarının yer aldığı daha detaylı bilgilere ulaşmak mümkündür. Bu bilgiler sınıflara göre gözlenen ve beklenen frekansların birlikte gösterildiği frekans histogramı grafiği ile test işlemi sonuçlarının yer aldığı bir tablodan oluşmaktadır. Bu tabloda verilen p olasılığı (önem seviyesi) dikkate alındığında, eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez.

Algoritma II (yani veri düzeni tüm değişkenler için her bir birime ait ölçümler, yani ham veri düzeninde ise) aşağıdaki adımları uygulamak suretiyle de uyum iyiliği için Ki-Kare testi yapılabilir:

Algoritma:II.2

Adım.1 Sınıflama veya sıralama düzeyinde ölçülen değişken ve özellikleri tanımlandıktan sonra, bu değişken için her bir birime ait değerler veri giriş sayfasında girilir.

Adım.2 Analyze > Nonparametric Tests > Legacy dialogs > Chi-Square... seçenekleri seçilir ve yeni bir ekran açılır.

Adım.3 Açılan yeni ekranda yer alan değişkenler listesinden **sınıf** değişkeni seçilerek aktarma butonu ile **Test Variable List** işlem kutusuna aktarılır.

Adım.4 Aynı ekranda yer alan **Expected Values** (beklenen frekanslar) bölümünde sunulan iki seçenekten uygun olanı tercih edilir. Burada sunulan seçeneklerden birincisi P_j olasılıkların ve böylece B_j beklenen frekanslarının tüm sınıflar için eşit olduğu durumu kapsarken, ikincisi söz konusu olasılıkların ya da beklenen frekansların birbirinden farklı olduğu durumu kapsamaktadır. Eğer beklenen frekanslar farklı ise her bir sınıf için beklenen frekanslar sırası ile **Values** seçeneği kullanılarak girilir.

Adım.5 OK tıklanarak analiz bitirilir. Sonuç **Output** sayfasında tablolar halinde verilir. Bu tablolardan birincisinde gözlenen frekanslar, beklenen frekanslar ve bunların farkları (yani $G_j - B_j$ farkları) her bir sınıf için yer alırken, ikincisinde Ki-Kare testine ait analiz sonuçları yer alır. Bu tabloda verilen p olasılığı (önem seviyesi) dikkate alındığında, eğer $p < \alpha$ ise H_0 hipotezi ret edilir, $p \geq \alpha$ ise H_0 hipotezi ret edilemez.

ÖRNEK 5.5 Trafik düzenlemesi amacıyla belli bir yoldan geçen araçların sayısının haftanın günlerine göre farklılık gösterip göstermediği öğrenilmek isteniyor. Bu amaçla haftanın her günü bu yoldan geçen araçlar sayılıyor ve günlere göre bu yoldan geçen araç sayıları aşağıdaki gibi gözlenmiştir. Belirlenen yoldan haftanın günlerine göre geçen araç sayılarının aynı olup olmadığına %5 önem seviyesinde karar veriniz?

Tablo 5.16 Haftanın Günlerine Göre Geçen Araç Sayısı

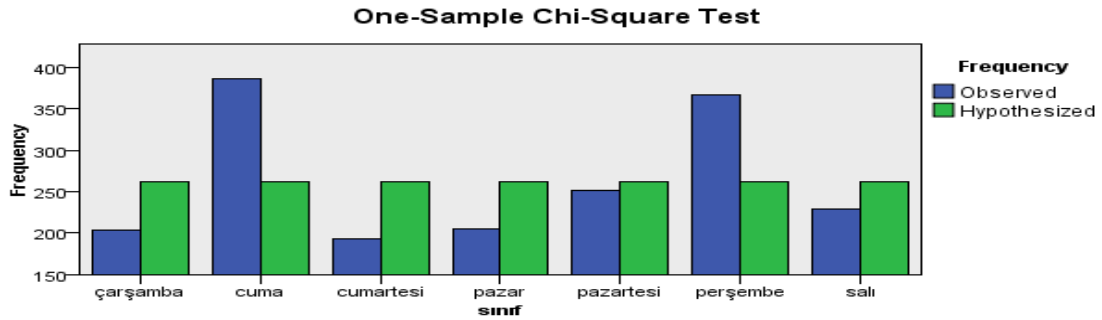
Günler	Gözlenen frekans (G_j)	p_j	B_j
Pazartesi (1)	251	1/7	262,14
Salı (2)	229	1/7	262,14
Çarşamba (3)	204	1/7	262,14
Perşembe (4)	367	1/7	262,14
Cuma (5)	386	1/7	262,14
Cumartesi (6)	193	1/7	262,14
Pazar (7)	205	1/7	262,14
Toplam	1835	1,00	1835

Cözüm: Test edilecek hipotezler;

H_0 : Belli bir yoldan geçen araç sayıları haftanın günlerine göre aynıdır (Gözlenen frekanslar ile beklenen frekanslar uyumludur)

H_1 : Belli bir yoldan geçen araç sayıları haftanın günlerine göre aynı değildir (Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir)

olup, analiz uyum iyiliği için Ki-Kare testi ile yapılır ve Algoritma I.1 veya Algoritma I.2 uygulanabilir. Algoritma I.1'e göre analiz sonuçları:



Total N	1.835
Test Statistic	148,716
Degrees of Freedom	6
Asymptotic Sig. (2-sided test)	,000

1. There are 0 cells (0%) with expected values less than 5. The minimum expected value is 262,143.

şeklinde elde edilir. H_0 doğru iken bir aracın belirlenen yoldan haftanın herhangi bir gününde geçme olasılıkları birbirine eşit olup $j = 1, 2, \dots, 7$ için $p_j = 1/7$ dir ve böylece herbir sınıf için beklenen frekanslar $B_j = n * p_j = 1835 * \frac{1}{7} = 262,14$ olarak bulunur. Burada $2p=0.000$ olduğundan $p=0.00$ olup, $\alpha = 0,05$ önem seviyesi için $p < \alpha$ olduğundan H_0 hipotezi ret edilir. Yani haftanın günlerine göre belirlenen yoldan geçen araç sayıları farklılık göstermektedir.

Algoritma I.2'ye göre analiz sonuçları:

Tablo 5.17 Haftanın Günlerine Göre Geçen Gözlenen ve Beklenen Araç Sayısı

	Observed N	Expected N	Residual
pazartesi	251	262,1	-11,1
salı	229	262,1	-33,1
çarşamba	204	262,1	-58,1
perşembe	367	262,1	104,9
cuma	386	262,1	123,9
cumartesi	193	262,1	-69,1
pazar	205	262,1	-57,1
Total	1835		

Tablo 5.18 Uyum İyiliği İçin Ki-kare Testi

	sınıf
Chi-Square	148,716 ^a
df	6
Asymp. Sig.	,000

a. 0 cells (0,0%) have expected frequencies less than 5.

Burada $p=0.00$ olup, $\alpha = 0,05$ önem seviyesi için $p < \alpha$ olduğundan H_0 hipotezi ret edilir. Yani haftanın günlerine göre belirlenen yoldan geçen araç sayıları farklılık göstermektedir.

ÖRNEK 5.6 Bir yerleşim yerinde yapılacak olan belediye başkanlığı seçimlerine A, B ve C adlarında üç aday katılıyor. Yapılan kamuoyu araştırmalarında geçerli oyların 1/5 ini A adayı alırken, geriye kalan oyların da diğer iki aday arasında eşit olarak paylaşıldığı iddia edilmektedir. Seçimde elde edilen 200 geçerli oyun adaylara göre dağılımı Tablo 5.19'de verilmiştir. Buna göre kamuoyu yoklamasında yapılan tahmin hakkında %5 önem seviyesinde ne söyleyebilirsiniz?

Tablo 5.19 Oyların adaylara dağılımı

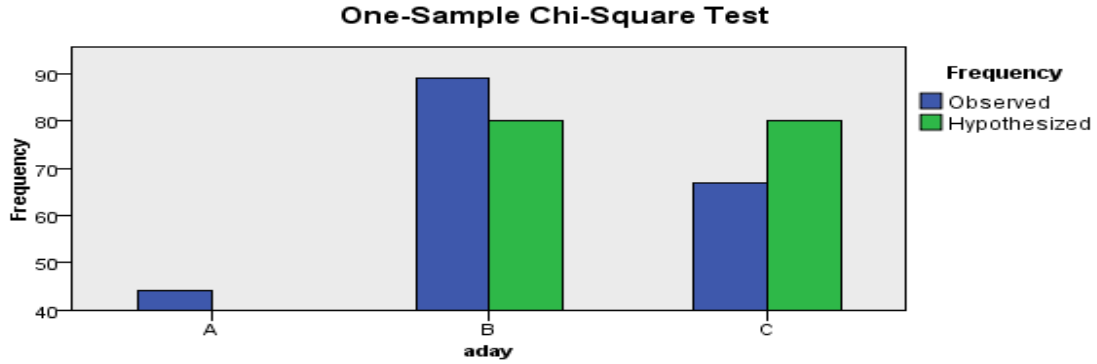
Aday	Seçim sonucu(G_j)	p_j	$B_j=n*p_j$
A	44	0,20	40
B	89	0,40	80
C	67	0,40	80
Toplam	200	1,00	200

Cözüm: Test edilecek hipotezler;

H_0 : Gözlenen frekanslar ile beklenen frekanslar uyumludur

H_1 : Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir

olup, analiz uyum iyiliği için Ki-Kare testi ile yapılır ve Algoritma I.1 veya Algoritma I.2 uygulanabilir. Algoritma I.1'e göre analiz sonuçları:



Total N	200
Test Statistic	3,525
Degrees of Freedom	2
Asymptotic Sig. (2-sided test)	,172

1. There are 0 cells (0%) with expected values less than 5. The minimum expected value is 40.

şeklinde elde edilir. Burada $2p=0.172$ olduğundan $p=0.086$ olup, $\alpha = 0,05$ önem seviyesi için $p > \alpha$ olduğundan H_0 hipotezi ret edilemez. Buna göre kamuoyu yoklamasında yapılan tahmin sonuçları geçerlidir.

Algoritma I.2'ye göre analiz sonuçları:

**Tablo 5.20 Haftanın Günlerine Göre Geçen
Gözlenen ve Beklenen Araç Sayısı**

	Observed N	Expected N	Residual
A	44	40,0	4,0
B	89	80,0	9,0
C	67	80,0	-13,0
Total	200		

**Tablo 5.21 Uyum İyiliği İçin
Ki-kare Testi**

	aday
Chi-Square	3,525 ^a
df	2
Asymp. Sig.	,172

a. 0 cells (0,0%) have expected
frequencies less than 5.

Burada $2p=0.172$ olduğundan $p=0.086$ olup, $\alpha = 0,05$ önem seviyesi için $p > \alpha$ olduğundan H_0 hipotezi ret edilemez. Buna göre kamuoyu yoklamasında yapılan tahmin sonuçları geçerlidir.