



**T.C.  
ONDOKUZ MAYIS ÜNİVERSİTESİ**

**FEN EDEBİYAT FAKÜLTESİ  
İSTATİSTİK BÖLÜMÜ**

**İST.482 PARAMETRİK OLMAYAN  
İSTATİSTİKSEL YÖNTEMLER**

**PROF. DR. YÜKSEL ÖNER**

**3. Hafta**

## 2.3 LILLIEFORS UYUM İYİLİĞİ TESTİ

Normal dağılıma uyumun anlamlı olup olmadığını göstermede kullanılan bir uyum iyiliği testidir. Eğer hipotezde öngörülen normal dağılımın parametreleri bilinmiyorsa ve bunlar örnek verisinden tahmin ediliyorsa normal dağılıma uyumun anlamlılığı testinde Lilliefors testinin kullanılması tercih edilmektedir. Test işleminin algoritması:

i) Hipotezler kurulur

a)  $H_0$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur

$H_1$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir

ii)  $n$  birimlik örnek  $X_1, X_2, \dots, X_n$  olmak üzere,  $H_0$  hipotezinde öngörülen normal dağılımın parametreleri tahmin edilir.  $\mu$  parametresinin tahmin edicisi  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  ve  $\sigma^2$  parametresinin tahmin edicisi  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  dir.

iii)  $H_0$  hipotezinde öngörülen normal dağılımın dağılım fonksiyonu bulunur.

Söz konusu dağılım fonksiyonu belirlemek amacıyla önce örneğin sıralı istatistikleri oluşturulur ve  $\forall x$  için  $F_0(x) = P(X \leq x) = P\left(Z \leq z = \frac{x - \bar{X}}{S}\right)$  değerleri bulunur.

iv) Örneğin dağılım fonksiyonu bulunur. Lilliefors testinde örneğin dağılım fonksiyonu standart değerlere ( $z$ ) dayanır. Sıralı istatistiklere göre  $\forall x$  için

$$s(x) = \frac{\left( \begin{array}{l} z' \text{ eşit veya daha küçük değerli} \\ \text{örnek birimlerinin sayısı} \end{array} \right)}{n} \quad (2.7)$$

şeklinde tanımlanır.

v) Test istatistiği belirlenir.

$F_0(x)$  sürekli ve  $s(x)$  kesikli fonksiyon olduğundan Lilliefors testi için test istatistiği;

$$D = \text{Sup}\{|s(x_j) - F_0(x_j)|, |s(x_{j-1}) - F_0(x_j)|\} \quad (2.8)$$

eşitliği ile verilir.

$H_0$  hipotezi doğru iken  $s(x)$  ve  $F_0(x)$  değerlerinin uyumlu ve böylece  $D$  istatistiğinin küçük bir değer alması beklenir. Test istatistiğinin örnekten hesaplanan değeri  $D_n$  olsun.

vi) Karar kuralı belirlenerek karar verilir ve yorumlanır.

Test istatistiğinin örnekleme dağılımından yararlanarak  $H_1$  hipotezine göre  $\alpha$  önem seviyesinde karar kuralı belirlenir.

$H_1$ 'e göre  $s(x)$  ve  $F_0(x)$  değerlerinin uyumlu olması beklenmez. Bu sebeple  $|s(x) - F_0(x)|$  mutlak farklarının büyük çıkması ve test istatistiğinin de büyük bir değer alması beklenir. Bu büyüklük kriteri,  $D$  istatistiğinin örnekleme dağılımından belirlenecek olan  $D_k$  kritik değeridir.  $D_k$ :  $n$  ve  $1 - \alpha$  için (T6) tablo değeri olmak üzere:

$D_h \geq D_k$  ise  $H_0$  ret edilir,  $D_h < D_k$  ise  $H_0$  ret edilemez.

**Örnek 2.6** Bir sınıftan seçilen 25 öğrencinin istatistik dersinden almış oldukları notlar aşağıdadır. Bu notların dağılımının  $N(\mu, \sigma^2)$  dağılımı ile uyumlu olup olmadığına %5 önem seviyesinde Lilliefors testi ile karar veriniz?

**Not ( $X_i$ ):** 60 40 45 90 70 40 40 70 90 80 70 70 20 10 20 80 60 40 70 90  
50 60 80 40 50

**Cözüm:** Değişken( $X$ ): Not (Puan)..... Nicel, sürekli ve ölçme düzeyi eşit aralıklı

i) Hipotezler kurulur

$H_0$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur

$H_1$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir

ii) Örnekten yararlanarak  $N(\mu, \sigma^2)$  dağılımının parametreleri tahmin edilir. Örnekte tekrarlanan gözlemler olduğundan veriler düzenlenerek tekrarlı frekans veri düzeni oluşturulur ve düzenlenmiş veriden parametreler tahmin edilir.

$x_j$	$f_j$	$f_j x_j$	$(x_j - \bar{x})$	$f_j (x_j - \bar{x})^2$	$z_j = \frac{x_j - 57,4}{22,69}$	$s(x_j)$	$F_0(x_j) = P(Z \leq z_j)$
10	1	10	-47,4	2246,76	-2,09	1/25	0,0183
20	2	40	-37,4	2797,52	-1,65	3/25	0,0495
40	5	200	-17,4	1513,80	-0,77	8/25	0,2206
45	1	45	-12,4	153,76	-0,55	9/25	0,2912
50	2	100	-7,4	109,52	-0,33	11/25	0,3707
60	3	180	2,6	20,28	0,11	14/25	0,5438
70	5	350	12,6	793,80	0,56	19/25	0,7123
80	3	240	22,6	1532,28	1,00	22/25	0,8413
90	3	270	32,6	3188,28	1,44	1,00	0,9251
Toplam	1435			12356			

Örneğin veri düzeni tekrarlı frekans veri düzeninde olduğundan,  $\mu$  kitle ortalamasının tahmin edicisi olan örnek ortalaması

$$\bar{X} = \frac{1}{n} \sum_{j=1}^c f_j X_j = \frac{1435}{25} = 57,4 \text{ puan iken,}$$

$\sigma^2$  kitle varyansının tahmin edicisi olan örnek varyansı

$$S^2 = \frac{1}{n-1} \sum_{j=1}^c f_j (X_j - \bar{X})^2 = \frac{12356}{24} = 514,83 \text{ ve böylece st. sapması } s = \sqrt{514,83} = 22,69 \text{ olarak bulunur.}$$

iii)  $H_0$  hipotezinde öngörülen normal dağılımın dağılım fonksiyonu ve örneğin dağılım fonksiyonu bulunur.

$x_j$	$ s(x_j) - F_0(x_j) $	$ s(x_{j-1}) - F_0(x_j) $
10	0,0217	0,0183
20	0,0705	0,0095
40	0,0994	0,1006
45	0,0688	0,0288
50	0,0693	0,0107
60	0,0162	0,1038
70	0,0477	0,1523
80	0,0387	0,0813
90	0,0749	0,0451

iv) Test istatistiğinin alabileceği değer hesaplanır.

$$D = \text{Sup}\{|s(x_j) - F_0(x_j)|, |s(x_{j-1}) - F_0(x_j)|\} \dots D_h = 0,1523$$

v) Karar kuralı belirlenerek karar verilir ve yorumlanır.

$\alpha = 0,05$ ,  $1 - \alpha = 0,95$  ve  $n = 25$  için (T6'dan)  $D_k = 0,173$  olup,  $D_h < D_k$  olduğundan  $H_0$  hipotezi ret edilemez. Buna göre mevcut örneğin dağılımı  $N(\mu, \sigma^2)$  dağılımı ile uyumludur.

## 2.4 SHAPIRO WILK UYUM İYİLİĞİ TESTİ

Normal dağılıma uyumlulukta yaygın olarak kullanılan uyum iyiliği testlerinden birisidir. İlgilenilen değişken nicel türden, sürekli ve en az eşit aralıklı ölçme düzeyine sahip olan bir değişken olmalıdır. Test işleminin algoritması:

i) Hipotezler kurulur

a)  $H_0$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur

$H_1$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir

ii)  $n$  birimlik örneğe ait ortalamadan farkların kareleri toplamı hesaplanır. Örnek birimleri  $X_1, X_2, \dots, X_n$  olmak üzere, ortalamadan farkların kareleri toplamı;

$D = \sum_{i=1}^n (X_i - \bar{X})^2$  veya  $D = \sum_{j=1}^c f_j (X_j - \bar{X})^2$  eşitliklerinden uygun olanı ile hesaplanır. Burada  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  veya  $\bar{X} = \frac{1}{n} \sum_{j=1}^c f_j X_j$  ile bulunur.

iii) Örneğin sıralı istatistikleri  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  oluşturulur ve test istatistiği belirlenir.

Shapiro- Wilk testi için test istatistiği:

$$W = \frac{1}{D} \left[ \sum_{i=1}^k a_i (X_{(n-i+1)} - X_{(i)}) \right]^2 \quad (2.9)$$

ile verilir. Burada  $k = \frac{n}{2}$  ve  $i = 1, 2, \dots, k$  için  $a_i$  katsayılar (T7) bulunacak olan Shapiro- Wilk katsayılarıdır. Test istatistiğinin örnekten hesaplanan değeri  $W_h$  olsun.

iv) Karar kuralı belirlenerek karar verilir ve yorumlanır.

Test istatistiğinin örnekleme dağılımından yararlanarak  $H_1$  hipotezine göre  $\alpha$  önem seviyesinde karar kuralı belirlenir.  $Pr(W < W_h) = p$  olmak üzere;  $p < \alpha$  ise  $H_0$  hipotezi ret edilir,  $p \geq \alpha$  ise  $H_0$  hipotezi ret edilemez. Burada  $p$  olasılığı iki farklı yolla bulunabilir.

**I.yol:** Test istatistiğinin örnekleme dağılımından yararlanarak yaklaşık olarak hesaplamak (T8.(a)) tablosu.

Örneğin  $n = 40$  ve  $W_h = 0,93$  olsun.  $Pr(W < 0,93) = p$  diyelim. T8(a)'dan  $n = 40$  iken,  $Pr(W < 0,928) = 0,02$  ve  $Pr(W < 0,94) = 0,05$  olup, buna göre  $0,02 < Pr(W < 0,93) < 0,05 = \alpha$ , yani  $p < \alpha$  dır.

**II.yol:** Dönüşüm yaparak test istatistiğinin örnekleme dağılımını standart normale dönüştürerek kesin olarak hesaplamak

Örnek hacmi  $n$  biliniyorken  $b_n, c_n$  ve  $d_n$  sabitler olmak üzere  $W$  istatistiği üzerinde uygulanan

$$G = b_n + c_n \ln \left( \frac{W - d_n}{1 - W} \right) \quad (2.10)$$

dönüşüm istatistiği için  $G \sim N(0, 1)$  dir. Burada  $b_n, c_n$  ve  $d_n$  sabitleri (T8(b))'den bulunabilir. Eşitlik (2,10) ile elde edilen değer  $G_h$  olmak üzere  $Pr(G < G_h) = p$  olasılığı standart normal dağılımdan kesin olarak bulunabilir.

**Örnek 2.7** Bir sınıftan seçilen 25 öğrencinin istatistik dersinden almış oldukları notlar aşağıdadır. Bu notların dağılımının  $N(\mu, \sigma^2)$  dağılımı ile uyumlu olup olmadığına %5 önem seviyesinde Shapiro-Wilk testi ile karar veriniz?

**Not ( $X_i$ ):** 60 40 45 90 70 40 40 70 90 80 70 70 20 10 20 80 60 40 70 90  
50 60 80 40 50

**Cözüm:** Değişken( $X$ ): Not (Puan)..... Nicel, sürekli ve ölçme düzeyi eşit aralıklı

i) Hipotezler kurulur

$H_0$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur

$H_1$ : Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir

ii) Örneğe ait ortalamadan farkların kareleri toplamı, örnek veri düzeni tekrarlı frekans veri düzeninde olduğundan Örnek 2.6'da  $D = \sum_{j=1}^c f_j (X_j - \bar{X})^2 = 12356$  olarak bulundu.

iii) Örneğin sıralı istatistikleri ve test istatistiğinin alabileceği değer bulunur.  $k = \frac{n}{2} = \frac{25}{2} \cong 13$

$x_j$	$f_j$	$i$	$x_{(i)}$	$a_i$	$X_{(n-i+1)} - X_{(i)}$	$a_i(X_{(n-i+1)} - X_{(i)})$
10	1	1	10	0,4450	90-10=80	35,600
20	2	2	20	0,3069	90-20=70	21,483
40	5	3	40	0,2543	90-20=70	17,801
45	1	4	45	0,2148	80-40=40	8,592
50	2	5	50	0,1822	80-40=40	7,288
60	3	6	60	0,1539	80-40=40	6,156
70	5	7	70	0,1283	70-40=30	3,849
80	3	8	80	0,1046	70-40=30	3,138
90	3	9	90	0,0823	70-45=25	2,0575
		10		0,0610	70-50=20	1,220
		11		0,0403	70-50=20	0,806
		12		0,0200	60-60=0	0,000
		13		0,0000	60-60=0	0,000
Top.	25					<b>107.9905</b>

Shapiro- Wilk testi için test istatistiğinin alabileceği değer:

$$W_h = \frac{1}{D} \left[ \sum_{i=1}^k a_i (X_{(n-i+1)} - X_{(i)}) \right]^2 = \frac{(107.9905)^2}{12356} = 0,944 \text{ bulunur.}$$

iv) Karar kuralı belirlenerek karar verilir ve yorumlanır.

$H_1$  hipotezine göre  $\alpha$  önem seviyesinde karar kuralı belirlenir.  $Pr(W < W_h) = p$  olmak üzere  $p < \alpha$  ise  $H_0$  hipotezi ret edilir,  $p \geq \alpha$  ise  $H_0$  hipotezi ret edilemez.  $p$  olasılığının belirlenmesi için iki yol mevcuttur.

**I.Yol:** Yaklaşık hesaplama;  $n = 25$  ve  $\alpha = 0,05$  iken (T8(a)) tablosundan  $p = Pr(W \leq W_h) = Pr(W \leq 0,944) = ?$

$Pr(W < 0,931) = 0,10$  ve  $Pr(W < 0,964) = 0,50$  olduğundan  $0,10 < p < 0,50$  yani  $p > \alpha$  elde edilir ve böylece  $H_0$  hipotezi ret edilemez. Buna göre Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur.

**II.Yol:** Kesin hesaplama;  $n = 25$  iken T8(b)'den,  $b_n = -5,704$ ,  $c_n = 1,876$  ve  $d_n = 0,2063$  olmak üzere  $G$  istatistiğinin alabileceği değer;

$$G_h = b_n + c_n \ln \left( \frac{W_h - d_n}{1 - W_h} \right) = -5,704 + 1,876 \ln \left( \frac{0,944 - 0,2063}{1 - 0,944} \right) = -0,89 \text{ bulunur. } \alpha = 0,05 \text{ ve}$$

$G \sim N(0, 1)$  olduğundan  $p = Pr(G \leq G_h) = Pr(G \leq -0,89) = 1 - 0,8133 = 0,1867$  olup,  $p > \alpha$  elde edilir ve böylece  $H_0$  hipotezi ret edilemez. Buna göre Örnek verisinin dağılımı,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur.

## 2.5 SPSS'DE UYUM İYİLİĞİ ANALİZİ

### 2.5.1 Ki-Kare Uyum İyiliği Testi Algoritması

Spss'de Ki-Kare uyum iyiliği testini uygulayabilmek için takip edilecek algoritmanın adımları şu şekilde verilebilir.

**1.Adım** Değişkenler ve özelliklerinin tanımlamaları yapılır. **Variable View** sayfasında sınıflama düzeyinde ölçülen değişkenin kategorileri (sınıfları) **Values** işlem penceresinde kodlanır. Gözlenen frekanslar için frekans değişkeni tanımlanır.

**2.Adım** Frekans değişkeni ağırlıklandırılır. Bu işlem için **Data > Weight Cases...** seçenekleri sonucunda açılan ekranda **Weight cases by** seçeneği işaretlenerek frekans değişkeni **Frequency Variable** işlem kutusuna aktarılır. **Ok** tıklanarak ağırlıklandırma işlemi tamamlanmış olur, bu durumda **Data View** sayfasının sağ alt köşesinde **Weight On** uyarısı görülür.

**3.Adım I. YOL:** **Analyze > Nonparametric Tests > Legacy Dialogs > Chi Square...** yolu izlenerek açılan ekranda değişkenler listesinde yer alan ve kategorileri gösteren değişken seçilerek **Test Variable List** işlem kutusuna aktarılır. Aynı ekranda bulunan **Expected Values** kutusundan uygun olan seçenek tercih edilir. Eğer tüm kategoriler eşit oranlarda tercih ediliyorsa **All categories equal** seçeneği, eğer sınıflara ait beklenen frekanslar belli ise **Values** seçeneği işaretlenerek aktif hale gelen pencereden sırası ile her bir sınıfa ait beklenen frekanslar girilir.

Eğer sınıflara ait beklenen frekanslar bilinmiyorsa, SPSS yardımıyla hesaplanabilir. Beklenen frekansları hesaplamak için önce  $p_j, j = 1, 2, \dots, c$  sınıf olasılıkları hesaplanır. Bunun için **Transform > Compute Variable** tıklanarak açılan pencerede **Target Variable** hücreğine  $p_j$  yazılır. Sonra aynı pencerede **Function group** bölümünden **PDF&Noncentral PDF** seçeneği ve **Functions and special variables** bölümünden  $H_0$  hipotezinde öngörülen teorik dağılıma ait **Pdf(öngörülen dağılım)** seçeneği seçilir. Bu seçim bu bölümün sol tarafındaki boş hücrede görüntülenir. Görüntülenen bilgi aktarma butonu ile **Numeric Expression** işlem kutusuna aktarılır. Buraya aktarılan fonksiyonda bulunan soru işaretleri yerine istenen bilgiler girilir ve **Ok** tıklanır. Böylece  $p_j$  olasılıkları hesaplanarak **Data View** sayfasında yeni bir değişken olarak görüntülenir. Beklenen frekansları hesaplamak için **Transform > Compute Variable** tıklanarak açılan pencerede **Target Variable** hücreğine  $B_j$  yazılır. Sonra **Numeric Expression** işlem kutusuna  $p_j \times n$  işlemi yazılarak  $B_j = p_j \times n$  denklemi kurulur ve **Ok** tıklanır. Böylece  $B_j$  beklenen frekansları hesaplanarak **Data View** sayfasında yeni bir değişken olarak görüntülenir. Hesaplanan beklenen frekanslar **Values** seçeneği işaretlenerek aktif hale gelen pencereden sırayla girilir.

**4.Adım Ok** tıklanarak işlem bitirilir, analiz sonuçları **Output** (çıktı) sayfasında tablolar halinde verilir. Tablolardaki bulgular değerlendirilerek “ $p < \alpha$  ise  $H_0$  ret edilir,  $p \geq \alpha$   $H_0$  ret edilemez” karar kuralına göre karar verilir ve yorumlama yapılır.

**3.Adım II. YOL: Analyze > Nonparametric Tests > One Sample** yolu izlenerek açılan ekranda **Fields** kutusu tıklanır ve açılan pencerede **Fields** bölümünden ilgilenilen değişken seçilerek **Test Fields** penceresine aktarılır. **Setting** kutusu tıklanır ve açılan pencerede **Customize Tests** işaretlenir ve açılan seçenekler içerisinde **Chi-Square Test** seçilir. Aktif konumdaki **Options...** kutusu tıklanır ve açılan pencerede, eğer tüm kategoriler eşit oranlarda tercih ediliyorsa **All categories have equal probability** seçeneği ve **Ok** tıklanır, önceki ekrana dönülür. Eğer sınıflara ait beklenen frekanslar belli ise **Customize expected probability** seçeneği işaretlenir ve **Expected Probabilities** işlem kutusunda **Category** kutusuna alt alta kategoriler ve karşısındaki **Relative Frequency** kutusuna her bir kategoriye karşılık gelen beklenen frekanslar girilir ve **Ok** tıklanır, önceki ekrana dönülür. Eğer sınıflara ait beklenen frekanslar bilinmiyorsa yukarıda açıklandığı şekli ile SPSS yardımıyla hesaplanabilir. Hesaplanan beklenen frekanslar **Customize expected probability** seçeneği işaretlenir ve **Expected Probabilities** işlem kutusunda **Category** kutusuna alt alta kategoriler ve karşısındaki **Relative Frequency** kutusuna her bir kategoriye karşılık gelen beklenen frekanslar girilir ve **Ok** tıklanır, önceki ekrana dönülür.

**4.Adım Run** tuşu tıklanarak işlem bitirilir, analiz sonuçları çıktı sayfasında bir tabloda sunulur. Tablonun üzerine çift tıkladığı zaman detaylı analiz sonuçları grafik ve tablosuna ulaşılır. Bu grafik ve tablolara bakılarak bulgular değerlendirilerek “ $p < \alpha$  ise  $H_0$  ret edilir,  $p \geq \alpha$   $H_0$  ret edilemez” karar kuralına göre karar verilir ve yorumlama yapılır.

## 2.5.2 Kolmogorov-Smirnov Uyum İyiliği Testi Algoritması

### I. Algoritma

**1.Adım** Değişkenler ve özelliklerinin tanımlamaları **Variable View** sayfasında yapılır. **Data View** sayfasında her bir değişkene ait veriler girilir ve küçükten büyüğe doğru sıralanır. Aynı değerli gözlemler varsa bunlar tekrarlanan gözlem sayısı kadar alt alta girilir.

**2.Adım Analyze > Nonparametric Tests > One Sample** yolu izlenerek açılan ekranda **Fields** kutusu tıklanır ve açılan pencerede **Fields** bölümünden ilgilenilen değişken seçilerek **Test Fields** penceresine aktarılır. **Setting** kutusu tıklanır ve açılan pencerede **Customize Tests** işaretlenir ve açılan seçenekler içerisinde **Kolmogorov-Smirnov Test** seçilir. Aktif konumdaki **Options...** kutusu tıklanır ve açılan pencerede uyum iyiliği testi uygulanacak dağılım seçilir. Dağılımın parametreleri / ortalaması örnek veri kullanılacaksa **Use Sample Data** seçeneği, bilinen parametreler / ortalama girilecekse **Custom** seçeneği işaretlenir ve ilgili kutulara bilinene parametreler / ortalama girilir. **Ok** tıklanır, önceki ekrana dönülür.

**3.Adım Run** tuşu tıklanarak işlem bitirilir, analiz sonuçları çıktı sayfasında bir tabloda sunulur. Tablonun üzerine çift tıkladığı zaman detaylı analiz sonuçları grafik ve tablosuna ulaşılır. Bu grafik ve tablolara bakılarak bulgular değerlendirilerek “ $p < \alpha$  ise  $H_0$  ret edilir,  $p \geq \alpha$   $H_0$  ret edilemez” karar kuralına göre karar verilir ve yorumlama yapılır.



## II. Algoritma

**1.Adım** Değişkenler ve özelliklerinin tanımlamaları **Variable View** sayfasında yapılır. **Data View** sayfasında her bir değişkene ait veriler girilir ve küçükten büyüğe doğru sıralanır. Aynı değerli gözlemler varsa bunlar tekrarlanan gözlem sayısı kadar alt alta girilir.

**2.Adım Analyze > Nonparametric Tests > Legacy Dialogs > 1-sample K-S...** yolu izlenerek açılan ekranda değişkenler listesinde yer alan ilgilenilen değişken seçilerek **Test Variable List** işlem kutusuna aktarılır. **Test distribution** bölümünden uyum iyiliği testi uygulanacak dağılım seçilir.

**3.Adım Ok** tıklanarak işlem bitirilir, analiz sonuçları **Output** (çıktı) sayfasında tablolar halinde verilir. Tablolardaki bulgular değerlendirilerek “ $p < \alpha$  ise  $H_0$  ret edilir,  $p \geq \alpha$   $H_0$  ret edilemez” karar kuralına göre karar verilir ve yorumlama yapılır.

### **2.5.2 Normallik için Lilliefors ve Shapiro Wilk Testi Algoritması**

**Adım:1** Değişken ve özellikleri tanımlanıp veriler girilir. Eğer birden fazla grup için aynı anda normallik testi uygulanacaksa, ayrıca grup değişkeni adı ile yeni bir değişken daha tanımlanmalıdır. Bu yeni değişken bir kategorik değişken olacağından, bu değişkene ait kategorilerinde **Value** bölümünde girilmesi gerekir.

**Adım:2 Analyze > Descriptive Statistics > Explore** yolu izlenerek açılan ekranda ilgilenilen değişken **Depent List** işlem kutusuna aktarılır. Eğer birden fazla grup için aynı anda normallik testi uygulanacaksa grup değişkeni olarak tanımlanan kategorik değişken **Factor List** işlem kutusuna aktarılır.

**Adım:3 Plots** seçeneği seçilerek açılacak olan ekranda **Normality plots with tests** seçeneği işaretlenir.

**Adım:4 Continue > OK** diyerek işlem bitirilir. Sonuçlar bir tablo halinde **Output** sayfasında sunulur. Tablolardaki bulgular değerlendirilerek “ $p < \alpha$  ise  $H_0$  ret edilir,  $p \geq \alpha$   $H_0$  ret edilemez” karar kuralına göre karar verilir ve yorumlama yapılır.

**Örnek 2.8** Örnek:2.1-2.7'nin istenen çözümlerini SPSS programı ile yineleyiniz? Bulduğunuz sonuçları önceki çözümlerden elde edilen sonuçlarla karşılaştırınız?

**Çözüm: Örnek 2.1** Ki-Kare uyum iyiliği analizi: **(I.Yol)**

$H_0$ : Örnek, öngörülen kesikli düzgün dağılıma sahip kitleden çekilmiştir (Gözlenen frekanslar ile beklenen frekanslar uyumludur)

$H_1$ : Örnek, öngörülen kesikli düzgün dağılıma sahip kitleden çekilmemiştir (Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir)

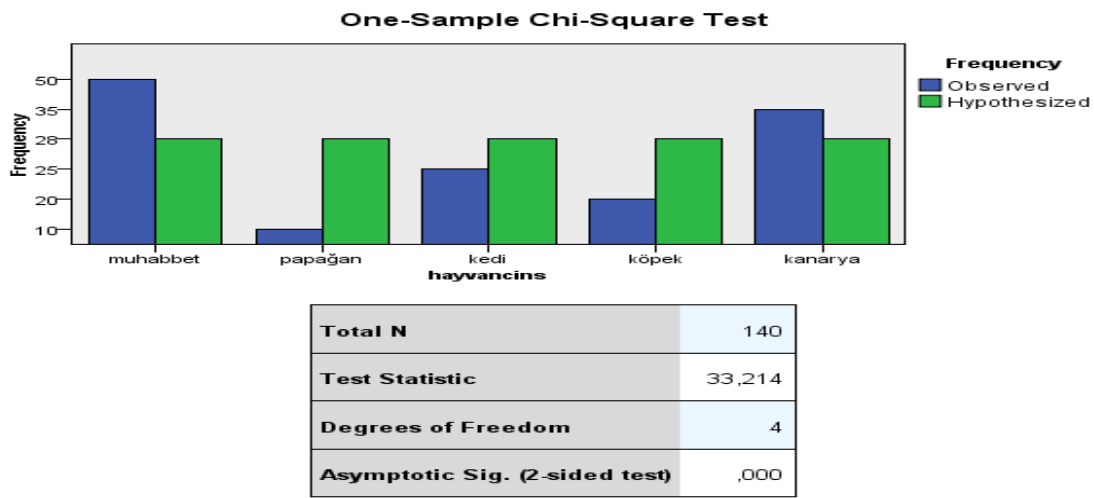
Hayvan Cinsi	Observed N	Expected N	Residual
muhabbet kuşu	50	28,0	22,0
papağan	10	28,0	-18,0
kedi	25	28,0	-3,0
köpek	20	28,0	-8,0
kanarya	35	28,0	7,0
Total	140		

Test Statistics	hayvancins
Chi-Square	33,214 <sup>a</sup>
df	4
Asymp. Sig.(p)	,000

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

$p < 0,05$  olduğundan ise  $H_0$  ret edilir.

Ki-Kare uyum iyiliği analizi: (II.Yol)



1. There are 0 cells (0%) with expected values less than 5. The minimum expected value is 28.

$p = 0.000 < 0,05$  olduğundan ise  $H_0$  ret edilir.

**Örnek 2.2** Ki-Kare uyum iyiliği analizi: (I.YOL)

$H_0$ : Hastaların iyileşme süresine ait dağılım,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur (Gözlenen frekanslar ile beklenen frekanslar uyumludur)

$H_1$ : Hastaların iyileşme süresine ait dağılım,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir (Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir)

Eğer sınıflara ait beklenen frekanslar bilinmiyorsa, SPSS yardımıyla hesaplanabilir. Normal dağılıma uyumluluk testinde gruplandırılmış frekans verilerinde beklenen frekansları hesaplamak için önce  $p_j, j = 1, 2, \dots, c$  sınıf olasılıkları hesaplanır. Bunun için öncelikle sınıf sınırlarının standart değerleri bulunmalıdır. Orijinal sınıf sınırları için alt ve üst sınırları ayrı birer değişken gibi girilir ve bu sınırlar üzerinde standartlaştırma dönüşümü  $Z = \frac{X-\mu}{\sigma}$  uygulanarak sınıfların standart sınırları bulunur. Alt sınırlar (**Zalt**) ve üst sınırlar (**Züst**) yeni

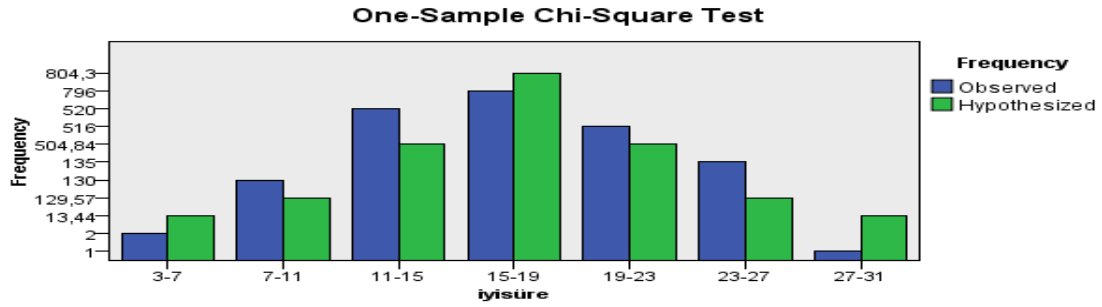
değişkenleri tanımlanır. Sonra hem alt sınırlar için hem de üst sınırlar için ayrı ayrı olasılık hesaplamaları yapılır. Alt sınırlar için hesaplanacak olasılıkları ( $p_{j1}$ ) ile üst sınırlar için hesaplanacak olasılıkları ( $p_{j2}$ ) ile gösterelim. Bu olasılıkları hesaplariken izlenecek yol şu şekildedir. **Transform > Compute Variable** tıklanarak açılan pencerede **Target Variable** hücreğine  $p_{j1}$  yazılır. Sonra aynı pencerede **Function group** bölümünden **CDF&Noncentral CDF** seçeneği ve **Functions and special variables** bölümünden **Cdfnorm** seçeneği seçilir. Bu seçim bu bölümün sol tarafındaki boş hücrede **CDFNORM(zvalue)...** şeklinde görüntülenir. Görüntülenen bilgi aktarma butonu ile **Numeric Expression** işlem kutusuna aktarılır. Buraya aktarılan fonksiyonda bulunan soru işareti yerine **Zalt** girilir ve **Ok** tıklanır. Böylece  $p_{j1}$  olasılıkları hesaplanarak **Data View** sayfasında yeni bir değişken olarak görüntülenir. Aynı yol izlenerek  $p_{j2}$  olasılıkları da hesaplanır. Daha sonra birinci sınıf için  $p_{j1}$  olasılığı 0 (sıfır) olarak ve son sınıf için  $p_{j2}$  olasılığı 1 (bir) olarak değiştirilir ve  $p_j = p_{j2} - p_{j1}$  fonksiyonu uygulanarak istenen  $p_j, j = 1, 2, \dots, c$  olasılıkları hesaplanmış olur. Beklenen frekansları hesaplamak için **Transform > Compute Variable** tıklanarak açılan pencerede **Target Variable** hücreğine  $B_j$  yazılır. Sonra **Numeric Expression** işlem kutusuna  $p_j \times n$  işlemi yazılarak  $B_j = p_j \times n$  denklemi kurulur ve **Ok** tıklanır. Böylece  $B_j$  beklenen frekansları hesaplanarak **Data View** sayfasında yeni bir değişken olarak görüntülenir. Hesaplanan beklenen frekanslar **Values** seçeneği işaretlenerek aktif hale gelen pencereden sırayla girilir.

iyisüre				Test Statistics	
	Observed N	Expected N	Residual		iyisüre
3-7	2	13,4	-11,4	Chi-Square	22,196 <sup>a</sup>
7-11	130	129,6	,4	df	6
11-15	520	504,9	15,1	Asymp. Sig.	,001
15-19	796	804,1	-8,1	a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 13,4.	
19-23	516	504,9	11,1		
23-27	135	129,6	5,4		
27-31	1	13,4	-12,4		
Total	2100				

## (II.Yol)

$H_0$ : Hastaların iyileşme süresine ait dağılım,  $N(\mu, \sigma^2)$  dağılımı ile uyumludur (Gözlenen frekanslar ile beklenen frekanslar uyumludur)

$H_1$ : Hastaların iyileşme süresine ait dağılım,  $N(\mu, \sigma^2)$  dağılımı ile uyumlu değildir (Gözlenen frekanslar ile beklenen frekanslar uyumlu değildir)



<b>Total N</b>	2.100
<b>Test Statistic</b>	22,269
<b>Degrees of Freedom</b>	6
<b>Asymptotic Sig. (2-sided test)</b>	,001

1. There are 0 cells (0%) with expected values less than 5. The minimum expected value is 13,440.

$p = 0.001 < 0,05 = \alpha$  olduğundan  $H_0$  hipotezi ret edilir. Yani hastaların iyileşme süresine ait dağılım  $N(\bar{X} = 17, S^2 = 16,18)$  dağılımı ile uyumlu değildir.

### **Örnek:2.3** Ki-Kare uyum iyiliği (I.Yol)

$H_0$ : Herhangi bir grupta iyileşen hasta sayısının dağılımı  $Binom(k = 5; p = 0,95)$  dağılımı ile uyumludur

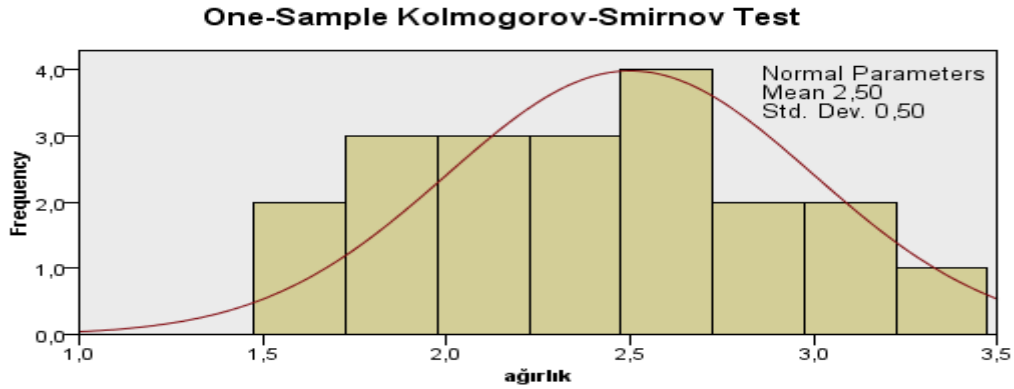
$H_1$ : Herhangi bir grupta iyileşen hasta sayısının dağılımı  $Binom(k = 5; p = 0,95)$  dağılımı ile uyumlu değildir

Hasta Sayısı	Observed N	Expected N	Residual
,00	1	,0	1,0
1,00	5	,0	5,0
2,00	10	,1	9,9
3,00	50	2,1	47,9
4,00	30	20,4	9,6
5,00	4	77,4	-73,4
Total	100		

Test Statistics	hastasayısı
Chi-Square	28253,518 <sup>a</sup>
df	5
Asymp. Sig.	P=0,000

a. 4 cells (66,7%) have expected frequencies less than 5. Beklenen frekansı 5'den küçük olan sınıfların birleştirilmesi gerekir. Bu işlemden sonra analiz tekrarlanmalıdır.  $p < 0,05$  olduğundan ise  $H_0$  ret edilir.

**Örnek:2.4** (ALGORİTMA-I) parametreler biliniyor



<b>Total N</b>		20
<b>Most Extreme Differences</b>	<b>Absolute</b>	,188
	<b>Positive</b>	,188
	<b>Negative</b>	-,036
<b>Test Statistic</b>		,841
<b>Asymptotic Sig. (2-sided test)</b>		,478

**Örnek:2.5** (ALG-II)

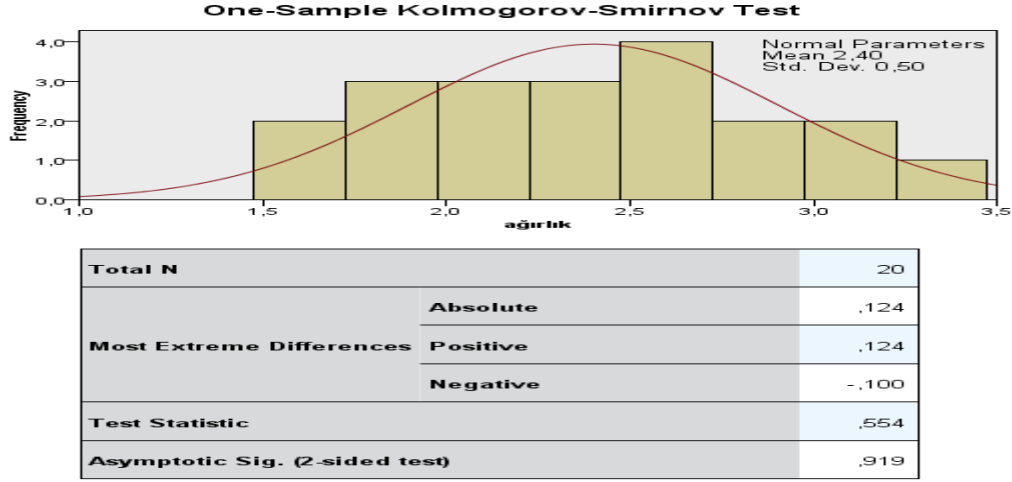
**One-Sample Kolmogorov-Smirnov Test (Parametreler bilinmiyor)**

		ağırlık
N		20
Normal Parameters <sup>a,b</sup>	Mean	2,4000
	Std. Deviation	,50498
	<b>Absolute</b>	<b>,124</b>
Most Extreme Differences	Positive	,124
	Negative	-,100
Kolmogorov-Smirnov Z		,554
Asymp. Sig. (2-tailed)		<b>,919</b>

a. Test distribution is Normal. b. Calculated from data.  
)

$p = 0,478 > 0,05$  veya  $\left( p = \frac{0,919}{2} = 0,460 > 0,05 \right)$  olduğundan  $H_0$  ret edilemez

Örnek:2.5 (ALG-I parametreler bilinmiyor)



$p = \frac{0,919}{2} = 0,460 > 0,05, H_0$  olduğundan ret edilemez

Örnek: (2.6-2.7)

**Tests of Normality**

	Lilliefors Testi			Shapiro-Wilk Testi		
	Statistic	df	Sig.(p)	Statistic	df	Sig.(p)
puan	.151	25	.147	.945	25	.197

$p > 0,05$  olduğundan  $H_0$  ret edilemez.