NORMAL DAĞILIM İÇİN UYUM İYİLİĞİ TESTLERİ VE BİR SİMÜLASYON ÇALIŞMASI

Nurcan YILDIRIM

YÜKSEK LİSANS TEZİ
İSTATİSTİK

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ŞUBAT 2013
ANKARA

Nurcan YILDIRIM tarafından hazırlanan “NORMAL DAĞILIM İÇİN UYUM İYİLİĞİ TESTLERİ VE BİR SİMÜLASYON ÇALIŞMASI” adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Fikri GÖKPINAR ………………………..
Tez Danışmanı, İstatistik Bölümü, Gazi Üniversitesi

Bu çalışma, jürimiz tarafından oy birliği ile İstatistik Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Prof. Dr. Berrin ÖZKAYA ………………………..
Gıda Mühendisliği, Ankara Üniversitesi

Doç. Dr. Fikri GÖKPINAR ………………………..
İstatistik Bölümü, Gazi Üniversitesi

Prof. Dr. Hülya BAYRAK ………………………..
İstatistik Bölümü, Gazi Üniversitesi

Tez Savunma Tarihi: 04/02/2013

Bu tez ile G.Ü. Fen Bilimleri Enstitüsü Yönetim Kurulu Yüksek Lisans derecesini onamıştır.

Prof. Dr. Şeref SAĞIROĞLU ………………………..
Fen Bilimleri Enstitüsü Müdürü

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Nurcan YILDIRIM

NORMAL DAĞILIM İÇİN UYUM İYİLİĞİ TESTLERİ VE BİR SİMÜLASYON ÇALIŞMASI

(Yüksek Lisans Tezi)

Nurcan YILDIRIM

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Şubat 2013

ÖZET

İstatistiksel bir modelin uyum iyiliği gözlenen bir veri setinin istatistiksel modele uyumluluğunu test eder. Bu çalışmada, uyum iyiliği testlerinden Ki Kare, Cramer-von Mises, Kolmogorov- Smirnov, Anderson-Darling, Watson, Shapiro-Wilk, Jarque Bera, Zhang, Esteban ve diğerleri tarafından verilen testler tanıtılmıştır. Ayrıca bu testlerin testin gücü bakımından hangi durumlarda birbirlerine göre daha iyi oldukları belirlenmiştir. Bu testler gamma, üstel, lognormal, tekdüze, beta, *t* ve uçdeğer altında kıyaslanmıştır.

Bilim Kodu : 205.1.066
Anahtar Kelimeler :Uyum İyiliği Testi, Normallik, 1.Tip Hata, Testin Gücü
Sayfa Adedi :49
Tez Yöneticisi : Doç. Dr. Fikri GÖKPINAR

GOODNESS OF FIT TESTS FOR NORMAL DISTRIBUTION AND A SIMULATION STUDY

(M.Sc.Thesis)

Nurcan YILDIRIM

GAZI UNIVERSITY
INSTITUTE SCIENCE AND TECHNOLOGY
February 2013

ABSTRACT

The goodness of fit of a statistical model tests how well it fits a set of observations. In this study, some goodness of fit tests called Chi-Square, Cramer-von Mises, Kolmogorov- Smirnov, Anderson- Darling, Watson, the Shapiro-Wilk, Jarque Bera, Zhang, Esteban et al. are investigated. In addition, a power comparison is made to determined which tests under what circumstances they are superior to the others. These test are compared under gamma, exponential, lognormal, uniform, beta, t and Extremum distributions.

Science Code	**: 205.1.066**
Key Words	**: Goodness of fit, Normality, Type I Error, Power of test.**
Page Number	**: 49**
Adviser	**: Assoc.Prof. Dr. Fikri GÖKPINAR**

TEŞEKKÜR

Çalışmalarım boyunca değerli yardım ve katkılarıyla beni yönlendiren danışmanım Sayın Hocam Doç. Dr. Fikri GÖKPINAR’ a, tecrübelerinden faydalandığım hocam Prof. Dr. Hülya BAYRAK’ a, çalışmalarım boyunca yardımlarını esirgemeyen arkadaşlarıma ve ayrıca manevi destekleri ile beni hiçbir zaman yalnız bırakmayan aileme teşekkürü bir borç bilirim.

İÇİNDEKİLER

Sayfa

Sayfa

ÇİZELGELERİN LİSTESİ

SİMGELER ve KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	**Açıklama**
n	Örnek çapı
c	Serbestlik derecesi
B_j	Beklenen frekans
G_j	Gözlenen frekans
X_i	i’ nci örnek gözlemi
$X_{(i)}$	i’ nci sıralı istatistik
α	1. tip hata oranı
$F(x)$	Yığının dağılım fonksiyonu
$F_0(x)$	Yokluk hipotezinde belirtilen dağılım fonksiyonu
$F_n(x)$	Gözlenen dağılım fonksiyonu
C_α	Watson’ ın U_n^2 testi için kritik değer
S	Örnek çarpıklık ölçüsü
K	Örnek basıklık ölçüsü
$\overline{X}$	Örnek ortalaması
s^2	Örnek varyansı
$V = (v_{ij})$	Sıra istatistiklerinin varyans kovaryans matrisi

Kısaltmalar	**Açıklama**
χ^2	Ki-Kare Test istatistiği
W_n^2	Cramer von Mises test istatistiği
D	Kolmogorov-Smirnov Test istatistiği
A^2	Anderson Darling Test istatistiği
A^{2*}	Değiştirilmiş Anderson Darling Test istatistiği
U_n^2	Watson Test istatistiği
W	Shapiro-Wilk *Test istatistiği*
JB	Jarque-Bera Test istatistiği
Z_K, Z_A, Z_C	Zhang Test istatistikleri
$S_{n,A}^1$, $S_{n,B}^1$, $S_{n,C}^1$	Esteban ve diğerleri Test istatistikleri

1.GİRİŞ

İstatistiksel hesaplamalarda örneklerin geldikleri yığınların dağılımlarının bilinmesi parametrik testlerin uygulanabilmesi için önemli bir varsayımdır. Diğer bir ifade ile örneğin dağılım biçimi bilinmiyorsa, parametrik testlerin kullanılması doğru olmaz. Bu durumda parametrik olmayan testlerin kullanılması önerilmektedir.

İstatistiksel hesaplamalarda, n hacimli bir örneğin belirtilen bir dağılımdan gelip gelmediğini belirlemek için ‘ Uyum İyiliği’ testleri kullanılır. Uyum iyiliği testlerinde n hacimli bir örneğin belirtilen bir dağılımdan gelip gelmediği incelenir.

Uyum iyiliği testlerinin amacı, verilerin varsayılmış bir modelden ne kadar saptığını bir ölçü birimi yardımı ile ölçmek ve bu farkı yokluk hipotezi altındaki dağılımdan elde edilen değerle kıyaslamaktır.

Uyum iyiliği konusu 1900 yılı civarında çalışılmaya başlanmış ve günümüzde de halen çalışılan popüler bir konudur. Bu konuyla ilgili olarak araştırmacılar tarafından birçok test istatistiği önerilmiştir. Bu çalışmalardan bazıları Pearson (1895, 1900, 1905), Kolmogorov-Smirnov (1933,1939), Cramer (1928), von-Mises (1931), Anderson-Darling (1952, 1954, 1962), Cox (1961,1962), Watson (1961,1962), Kuiper (1962), Lilliefors (1967), Shapiro ve Francia (1972), Vasicek (1976), Pettitt (1977), Cressie (1978,1979), Jarque-Bera (1980, 1981), Epps (1982, 1983, 2005), Baringhaus ve Henze (1988), Ledwina (1994), Fan (1996, 1997, 1998), Owen (2001), Bera ve Bilias (2002), Zhang (2002), Matsui ve Tamura (2005), Esteban ve diğerleri (2001, 2007), Dong ve Giles (2007), Towhidi ve Salmanpour (2007), Romao ve diğerleri (2009), Shan ve diğerleri (2011) tarafından yapılmıştır.

Literatürde geçen uyum iyiliği testleri dört ayrı kategoride sınıflandırılabilir: Ki-Kare ile ilişkili testler, deneysel dağılımla ilişkili testler, korelasyon tabanlı testler ve karakteristik (moment üreten) fonksiyonu temeline dayalı testler.

Uyum iyiliği testlerinden en yaygın olarak kullanılan ve bilinen Ki-Kare ve Kolmogorov-Smirnov testleridir. [Pearson, 1900; Kolmogorov, 1933].

Ki-Kare testi, örneği kategorilere göre sınıflandırmaya ve gözlenen frekans ile beklenen frekans arasındaki farkın karesinin farz edilen dağılıma göre ölçülmesine dayanır. Bu testin zayıf yönleri sınıflandırma işlemine bağlı olarak ortaya çıkan bilgi kaybıdır. Aynı zamanda geniş bir örnek ölçüsüne ihtiyaç duyulmasıdır. [Zghoul, 2010].

Kolmogorov- Smirnov, Cramer-von Mises ve Anderson-Darling tarafından önerilen testler, deneysel dağılım fonksiyonuyla ilişkili testlerdir ve sürekli dağılımları test etmek için daha uygundurlar. Bu testler yokluk hipotezi altında deneysel dağılım fonksiyonu ve kümülatif dağılım fonksiyonu arasındaki farkı ölçmeye yöneliktir. Kolmogorov- Smirnov testi deneysel dağılım fonksiyonu ile kümülatif dağılım fonksiyonu arasındaki en geniş farkı ölçerken, Cramer-von Mises testi ise bu farkların karelerinin ağırlıklı ortalamasını ölçer. Anderson- Darling testi kümülatif dağılım fonksiyonunun uzantılarına daha çok ağırlık verdiği için diğer testlerden daha hassastır. Klar (2001) ise teorik dağılım fonksiyonu ve deneysel dağılım fonksiyonu arasındaki farkı ortaya koymayı amaçlayan bir test önermiştir. [Zghoul, 2010].

Massey (1951) yaptığı çalışmada dağılımın parametreleri örnekten tahmin edildiğinde Kolmogorov- Smirnov testi kullanılırsa, sonucun güvenilir olmayacağını ve doğru olan H_0 hipotezini red etme olasılığının Kolmogorov-Smirnov test istatistiğinin kritik değer tablosunda verilenden daha büyük olacağı sonucunu göstermiştir. Çünkü bir gözlem setinin yokluk hipotezinde belirtilen dağılımdan gelip gelmediğini belirlemek için Kolmogorov- Smirnov testi için oluşturulan kritik değer tablosu kullanılır. Eğer bir ya da daha fazla parametre örnekten tahmin edilirse, Kolmogorov- Smirnov testi için kullanılan kritik değer tablosu artık kullanılamaz. Bu çalışmaya dayanarak Lilliefors (1967), yığın ortalaması ve varyansı örnekten tahmin edildiğinde, Kolmogorov- Smirnov uyum iyiliği testi yerine kullanılabilecek bir test önermiştir. Ayrıca David ve Johnson (1948), tahmin edilen parametreler konum veya ölçüm parametreleri ise, belirli bir dağılım için Kolmogorov-Smirnov testi ile birlikte kullanılacak tabloların oluşturulmasının uygun olacağını ifade etmişlerdir. Bu amaçla Lilliefors (1967), ortalama ve varyans örnekten tahmin edildiğinde bir gözlem setinin normal dağılımdan gelip gelmediğini belirlemek amacıyla yapılacak testte Kolmogorov- Smirnov testi ile birlikte kullanılabilecek bir tabloyu Monte Carlo hesaplamalarından elde etmiştir.

Shapiro ve Wilk (1965), örneğe ait sıra istatistiklerinin uygun bir lineer bileşeninin karesinin, kareler toplamına oranıyla elde edilen bir test önermiştir. Örneklerin normal dağılıma uyumunu test etmek için Shapiro-Wilk test istatistiğinin hesaplanmasında lineer katsayılar tablosu mevcut olduğu için bu katsayılardan yararlanarak test istatistiğini hesaplamak çok kolaydır. Shapiro-Wilk testi $n<20$ küçük örnekler için bile normalliği test etmede kullanılan birçok alternatif testten daha hassastır. Shapiro-Wilk testi ile ilgili sakınca büyük örnekler için lineer katsayılar hesaplanması hem de büyük örnekler için dağılımın yüzde noktalarının belirlenmesi çok zordur.

Stephens (1974) yaptığı çalışmada, çeşitli örnek genişliklerinde ve çeşitli dağılımlar altında Kolmogorov-Smirnov, Anderson-Darling, Watson, Cramer-von Mises, Kuiper ve Shapiro-Wilk testlerinin güç performanslarını normallik bakımından karşılaştırmıştır. Bu karşılaştırma sonucunda deneysel dağılım fonksiyonu testlerinin Shapiro-Wilk testine karşı düşük güç değerlerine sahip olduğunu ancak Anderson- Darling testi ve Cramer-Von Mises testlerinin Shapiro-Wilk testine daha yakın güç değerleri verdiğini belirtmiştir.

Seier (2002), çeşitli örnek genişliklerinde ve çeşitli dağılımlar altında normallik testlerini deneysel I. tip hata ve güç bakımından karşılaştırmıştır. Yapılan bu çalışmada Anderson-Darling testinin Kolmogorov-Smirnov testinden daha güçlü olduğu sonucuna varılmıştır.

Bir rastgele değişkenin moment üreten fonksiyonu veya karakteristik fonksiyonu var olduğu zaman değişkenin dağılımı belirlenir. Bu fikirden yola çıkarak Epps ve diğerleri (1982), moment üreten fonksiyona dayanan bir test ve Epps ve Pulley' de (1983), karakteristik fonksiyonu temel alan bir başka test geliştirmişlerdir.

Zghoul tarafından Zghoul uyum iyiliği testi geliştirilmiştir. Bu test Epps ve Pulley (1983) tarafından önerilen testden farklı olarak teorik ve deneysel moment üreten fonksiyonları arasındaki sapmalara dayanırken, Epps-Pulley testi karakteristik fonksiyonun iki farklı hesaplaması arasındaki sapmayı temel almaktadır. [Zghoul, 2010].

Son zamanlarda da Vexler ve Gurevich (2010) entropiye dayalı uyum iyiliği testini parametrik olasılık birleşik yoğunluk fonksiyonlarını parametrik olmayan yöntemlerle tahmin ederek bulmuşlardır.

Bu çalışmada, yukarıda da bahsedilen bazı uyum iyiliği testlerinden Ki-Kare, Cramer-von Mises, Kolmogorov- Smirnov, Anderson-Darling, Watson, Shapiro-Wilk, Jarque Bera, Zhang, Esteban ve diğerleri tarafından önerilen testler incelenmiştir. Çalışmanın ikinci bölümünde bahsedilen bu test istatistikleri üzerinde durulmuş, üçüncü bölümde bu test istatistiklerinin kullanılışına ilişkin bir uygulama yapılmıştır. Dördüncü bölümde bu test istatistikleri simülasyon yoluyla karşılaştırılmıştır. Beşinci bölümde ele alınan testlerin birbirlerine göre hangi durumlarda daha iyi oldukları belirlenmeye çalışılmış ve bunlarla ilgili sonuç ve yorumlara yer verilmiştir.

2.TEST İSTATİSTİKLERİ

Bu bölümde Ki-Kare, Cramer-von Mises, Kolmogorov- Smirnov, Anderson-Darling, Watson, Shapiro-Wilk, Jarque Bera, Zhang, Esteban ve diğerleri testleri tanıtılacaktır. n hacimli bir örneğin belirtilen bir dağılımdan gelip gelmediğini belirlemek için yapılacak bir testte yokluk ve karşıt hipotezler;

H_0 : Örnek normal dağılıma sahip yığından seçilmiştir.

H_1 : Örnek normal dağılıma sahip yığından seçilmemiştir. (2.1)

biçiminde ifade edilir.

Eş 2.1' de verilen hipotez testi bu bölümde tanıtılacak olan uyum iyiliği testlerinde kullanılacaktır. Bölümde kullanılacak bazı ortak ifadeler aşağıdaki gibidir.

n birimlik $X_1, X_2, ..., X_n$ rastgele örneğinin geldiği dağılım fonksiyonu $F(x)$ ile gösterilir. Bu rastgele örneğin küçükten büyüğe doğru sıralanışı $X_{(1)}, X_{(2)}, ..., X_{(n)}$ şeklindedir. Ayrıca $F_0(x)$, H_0 yokluk hipotezinde belirtilen dağılım fonksiyonunu ve $F_n(x)$ deneysel dağılım fonksiyonunu ifade eder. Bu ifade $F_n(x)$; x değerine eşit ya da küçük değerli örnek birimlerinin sayısının örnek hacmine oranıdır.

2.1. Ki-Kare Uyum İyiliği Testi

Ki-Kare uyum iyiliği testi ilk olarak 1900' de Pearson tarafından ortaya atılmıştır.

Sınıflama ölçme düzeyinde ölçülmüş değişkenler için uyum iyiliği testi olan Ki-Kare testi için yokluk hipotezinin doğruluğu altında, yığından seçilen örneğin yığının karakteristiklerini taşıması yani sınıfların her birinde beklenen frekanslar ile gözlenen frekansların birbirine eşit ya da yakın olması beklenir. Buna göre Ki-Kare test istatistiği;

$$\chi_h^2 = \sum_{j=1}^{c} \frac{\left(G_j - B_j\right)^2}{B_j}$$

olarak tanımlanır. Burada c sınıf sayısı olmak üzere G_j , j' inci sınıftaki örnek birimlerin sayısı yani gözlenen frekans ve B_j beklenen frekans şeklindedir. Yokluk hipotezinin doğruluğu altında, rastgele seçilen herhangi bir birimin j' inci sınıfta olma olasılığı p_j ise beklenen frekans $B_j = np_j$ şeklinde olur.

Yokluk hipotezi doğru iken test istatistiği, yaklaşık olarak $c-1$ serbestlik dereceli Ki-Kare dağılımına sahiptir. α anlamlılık düzeyinde Ki-Kare dağılımından elde edilen değer $\chi^2_{c-1,\alpha}$ olmak üzere, $\chi^2_h \rangle \chi^2_{c-1,\alpha}$ ise H_O reddedilir.

Ki-Kare test istatistiğinin yaklaşık olarak $c-1$ serbestlik dereceli Ki-Kare dağılımına sahip olması için örnek hacminin büyük olması beklenir. Ayrıca hiçbir sınıfta beklenen frekansın 1 den küçük olmasına izin verilmez. Bu durum ortaya çıkarsa bazı sınıflar birleştirilebilir. Bu birleştirme işleminden sonra serbestlik derecesi değişmektedir.

2.2. Cramer -von Mises Testi

Cramer-von Mises uyum iyiliği testi Harald Cramer ve Richard Edler Mises (1928-1931) tarafından ortaya atılmıştır.

Cramer-von Mises test istatistiği W_n ;

$$W_n = \sum_{i=1}^{n} \left\{ F_0(x_{(i)}) - \frac{2i-1}{2n} \right\}^2 + \frac{1}{12n}$$

olarak tanımlanır.

Gözlem değerine bağlı olarak elde edilen test istatistiği değeri tablo değerinden daha büyük ise H_0 hipotezi reddedilir.

2.3. Kolmogorov-Smirnov Uyum İyiliği Testi

A.N. Kolmogorov (1933) ve N.V. Smirnov (1939) tarafından oranlama ya da eşit aralıklı düzeyde ölçülen değişkenler için geliştirilmiş uyum iyiliği testidir. Kolmogorov testi ve Smirnov testi benzerlik nedeniyle Kolmogorov-Smirnov uyum iyiliği testi olarak da bilinmektedir.

Kolmogorov-Smirnov testi yokluk hipotezinde belirtilen dağılım fonksiyonu $F_0(x)$ ile tüm x ' ler için örneğin dağılım fonksiyonu olan $F_n(x)$ arasındaki mutlak farklara dayanır. Buna göre Kolmogorov-Smirnov test istatistiği (D) ,

$$D = \sup_x |F_n(x) - F_0(x)|$$

şeklindedir.

Tek örnek için kullanılan D istatistiği için örnek dağılımından elde edilen kritik değerler Kolmogorov-Smirnov uyum iyiliği testine ilişkin tablodan n ve $1-\alpha$ değerlerine göre bulunan D_k değeri ve örnekten hesaplanan değer D_h olmak üzere, $D_h \geq D_k$ ise H_0 reddedilir.

Ayrıca tek örnek Kolmogorov-Smirnov uyum iyiliği test istatistiğinin değeri $F_0(x)$ ve $F_n(x)$ fonksiyonlarının grafikleri çizilerek de bulunabilir. Tek örnek

Kolmogorov-Smirnov uyum iyiliği test istatistiğinin değeri $F_0(x)$ ve $F_n(x)$ arasındaki en büyük dikey uzunluk hesaplanarak bulunur.

2.4. Anderson-Darling Uyum İyiliği Testi

Anderson ve Darling (1952), Kolmogorov- Smirnov testini uyarlayarak başka bir test istatistiği önermişlerdir.

Olasılık fonksiyonu ve bu olasılık fonksiyonunu tam olarak belirleyen parametre değerlerinin bilindiği bir yığından n birimlik bir rastgele örnek $\{ X_1, X_2, ..., X_n \}$ seçilir.

Anderson-Darling testi için yokluk hipotezi, örnek verilerin tüm parametre değerleri ile belirlenen dağılımdan geldiğidir. Eğer yokluk hipotezi test sonucu red edilirse, verilerin parametreler ile belirlenmiş dağılıma uymadığı sonucuna varılır. Bu test ilk olarak belli bir dağılım değil, sadece belli parametresi olan dağılım için oluşturulmuştur. Daha sonra parametrelerin bilinmediği durumlar içinde geliştirilmiştir.

Anderson-Darling test istatistiği (A^2);

$$A^2 = -\frac{2}{n}\sum_{i=1}^{n}\left[\left(i-\frac{1}{2}\right)\log\{F_0(x_{(i)})\}+\left(n-i+\frac{1}{2}\right)\log\{1-F_0(x_{(i)})\}\right]-n$$

şeklinde elde edilir.

Yokluk hipotezinde belirtilen olasılık dağılımına göre, elde edilen A^2 değerinin belirli bir sabitle çarpılması sonucu değiştirilmiş Anderson-Darling istatistiği A^{2*} test istatistiği bulunur. A^{2*} test istatistiği;

$$A^{2*} = A^2\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right)$$

şeklindedir.

Normal dağılıma uygunluğu test etmek için A^{2*} değeri kritik değerden büyük ise H_o yokluk hipotezi red edilir. Kritik değerler A^{2*} istatistik dağılımı bilinmediğinden simülasyon yoluyla elde edilir.

2.5. Watson Testi

Watson uyum iyiliği testi, Watson (1961-1962) tarafından ortaya atılmıştır.

Watson uyum iyiliği testinin en önemli özelliği, seçilen n hacimli örnekten elde edilen istatistiğin dağılımı bu örneğin geldiği $F(x)$ dağılım fonksiyonundan bağımsız olmasıdır.

Küçükten büyüğe doğru sıralanmış rastgele örnek $X_{(1)}, X_{(2)}, ..., X_{(n)}$ ve $y_i = F(X_{(i)})$ olsun.

$$\bar{y} = \left(\sum \frac{y_i}{n}\right)$$

olmak üzere; Watson test istatistiği (U_n);

$$U_n = \sum_{i=1}^{n}\left(y_i - \frac{2i-1}{2n}\right)^2 - n(\bar{y} - 1/2)^2 + \frac{1}{12n}$$

olarak tanımlanmıştır [Stephens, 1964].

U_n istatistiğinin oldukça küçük değerleri uyumun iyi olduğunu, tersine bu istatistiğin oldukça büyük değerleri de uyumun zayıf olduğunu ifade eder.

$n = 2,3,4,...,\infty$ iken U_n istatistiğinin tam olasılık dağılımlarını Watson (1961,1962) elde etmiş ve $P(U_n \rangle C_\alpha) = \alpha$ olmasını sağlayan C_α kritik değerlerini vermiştir.

n örnek hacmi büyük olduğu zaman U_n istatistiğinin tam olasılık dağılımının oluşturulması çok karmaşıktır. Bu nedenle Watson (1961,1962) U_n istatistiğinin dağılımlarını uygun Pearson eğrileri yöntemiyle yaklaşık olarak bulmuştur. Tam olasılık dağılımından elde edilen C_α değerleri ile Pearson eğrileri yaklaşımı ile bulunan C_α değerleri hemen hemen eşit çıkmıştır [Stephens, 1964].

n örnek hacmi için bu çizelge α anlamlılık düzeyinde $P(U_n \rangle C_\alpha) = \alpha$ olmasını sağlayan C_α kritik değerlerini verir.

U_n istatistiğinin örnekten hesaplanan değeri u_n ve α anlamlılık düzeyinde kritik değer çizelgesinden bulunan değer C_α olmak üzere,

$$u_n \geq C_\alpha$$

ise yokluk hipotezi α anlamlılık düzeyinde red edilir.

2.6. Shapiro-Wilk Testi

Shapiro-Wilk (1965) tarafından önerilen bu test, örneğe ait sıra istatistiklerinin uygun bir lineer bileşeninin karesinin, kareler toplamına bölümüyle elde edilir.

Shapiro-Wilk test istatistiği (W) ;

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

şeklindedir. Buradaki a' ,

$$a' = (a_1, a_2, ..., a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}}$$

olarak ifade edilir. Gerekli bazı tanımlamalar aşağıdaki gibidir.

$m' = (m_1, m_2, ..., m_n)$: standart normal dağılımda $N(0,1)$ ' da n adet sıra istatistiğinin beklenen değerlerinin vektörünü,

$V = (v_{ij})$: sıra istatistiklerinin varyans kovaryans matrisini

ifade eder.

Shapiro-Wilk test istatistiğinin hesaplanmasında a' katsayısının hesaplanmasındaki zorluklar nedeniyle 2' den 50' ye kadar a' katsayı değerleri tablo halinde elde edilmiştir. Bu tabloda a' değerlerini kullanarak b katsayısı ve W istatistiğinin hesaplanması aşağıda adımsal olarak gösterilmiştir.

i) $S_W^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2$ istatistiği hesaplanır.

ii) Eğer n ($n = 2k$) çift sayı ise b ifadesi;

$$b = \sum_{i=1}^{k} a_{n-i+1}(x_{n-i+1} - x_i)$$

Eğer n $(n = 2k + 1)$ tek sayı ise $a_{k+1} = 0$ olup b ifadesi

$$b = a_n(x_n - x_1) + ... + a_{k+2}(x_{k+2} - x_k)$$

şeklindedir. Burada x_{k+1} ifadesi b' nin hesaplanmasına katılmaz.

iii) W istatistiği ; $W = \dfrac{b^2}{S_W^2}$ eşitliği ile hesaplanır.

iv) Hesaplanan W istatistiği $W_{hesap} \langle W_{tablo}$ olduğunda hipotez reddedilir. [Shapiro-Wilk, 1965].

2.7. Jarque Bera Testi

Jarque- Bera uyum iyiliği testi, Jarque ve Bera (1980-1981) tarafından önerilmiştir.

Jarque-Bera testi, Lagrange çarpanı yöntemine dayalı bir testtir. Test istatistiği örnekten elde edilen basıklık ve çarpıklık ölçülerinin dönüşümlerinden elde edilmiştir. Burada yokluk hipotezi beklenen çarpıklığın ‘‘0’’ değerinde ve beklenen basıklığın ‘‘3’’ değerinde olmasıdır. O halde bir normal dağılım için bu değerler gereklidir.

Jarque-Bera test istatistiği (JB) ;

$$JB = \frac{n}{6}\left(C^2 + \frac{(K-3)^2}{4}\right)$$

şeklindedir.

Burada n örnek sayısı (veya genellikle serbestlik derecesi); C örnek çarpıklık ölçüsü, K örnek basıklık ölçüsü olmak üzere , C ve K değerleri

$$C = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}$$

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}}$$

elde edilir.

JB test istatistiği asimptotik olarak χ^2 dağılımına sahiptir. Örnek çarpıklığı '0' dan ve basıklığı '3' den sapma gösterdikçe, JB test istatistiği büyüme gösterir. Büyük örnekler için χ^2 yaklaşımı kullanılabiliyorken küçük örneklerde Monte Carlo yaklaşımı daha uygundur.

2.8. Zhang Testi

Zhang (2002), uyum iyiliği testlerine alternatif olarak uyum iyiliği testleri geliştirilmiştir.

Zhang (2002) tarafından önerilen üç test istatistiği aşağıdaki gibidir.

$$Z_K = \max_{1 \le i \le n}\left(\left(i - \frac{1}{2}\right)\log\left\{\frac{i - \frac{1}{2}}{nF_0(X_{(i)})}\right\} + \left(n - i + \frac{1}{2}\right)\log\left[\frac{n - i + \frac{1}{2}}{n\{1 - F_0(X_{(i)})\}}\right]\right)$$

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log\{F_0(X_{(i)})\}}{n-i+\frac{1}{2}} + \frac{\log\{1-F_0(X_{(i)})\}}{i-\frac{1}{2}}\right]$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{F_0(X_{(i)})^{-1}-1}{(n-\frac{1}{2})/(i-\frac{3}{4})-1}\right\}\right]^2$$

Z_K, Z_A, Z_C dağılımını oluşturmak zor olduğundan kritik değerler için Monte Carlo yöntemi kullanılmıştır.

2.9. Esteban ve Diğerleri Testi

Esteban ve diğerleri (2007) tarafından geliştirilen uyum iyiliği testidir.

Burada Cressie ve Read (1984-1988) tarafından elde edilen güç sapma istatistikleri,

$$T_{n,m}^{\lambda} = T_{n,m}^{\lambda}(p(Y_n), q) = \frac{2n}{\lambda(\lambda+1)}\sum_{i=1}^{m} p_i(Y_n)\left[\left(\frac{p_i(Y_n)}{q_i}\right)^{\lambda} - 1\right], -\infty\langle\lambda\langle\infty \qquad (2.2)$$

şeklindedir.

O halde $q_1 = \pi$ ve $p_1(Y_n) = F_0(F_n^{-1}(\pi))$ olmak üzere, Eş 2.2

$$T_{n,2,\pi}^{\lambda} = \frac{2n}{\lambda(\lambda+1)}\left[\frac{p_1(Y_n)^{\lambda+1}}{q_1^{\lambda}} + \frac{(1-p_1(Y_n))^{\lambda+1}}{(1-q_1)^{\lambda}} - 1\right]$$

şeklinde olur.

H_0 hipotezini test etmek için yeni istatistik formülleri:

$$S_{\max} = \sup_{\pi\in(0,1)} \left\{T^{\lambda}_{n,2,\pi}\omega(\pi)\right\} \tag{2.3}$$

$$S = \int_0^1 T^{\lambda}_{n,2,\pi} d\omega(\pi) \tag{2.4}$$

biçiminde ifade edilir. Burada [0,1] aralığında büyük değerleri yokluk hipotezini red eden S ve $S_{\max}$ kullanılarak $\omega(s)$ bazı ağırlık fonksiyonları elde edilmiştir. Burada $F_n(X_{(i)}) = \dfrac{i-1/2}{n}$ şeklinde tanımlanır. Elde edilen istatistikler aşağıdaki gibidir.

i) Eş 2.3' de $\omega(\pi) = \dfrac{1}{n}\pi(1-\pi)$ alınırsa

$$S_{n,A} = \max\left\{\frac{1}{n^2}, S^{1*}_{n,A}, (F_0(X_{(n)})-1)^2\right\}$$

elde edilir. Burada

$$S^{1*}_{n,A} = \max_{i=1,\dots,n-1}\left[\max\left\{\left(F_0(X_{(i)})-\frac{i}{n}\right)^2, \left(F_0(X_{(i)})-\frac{i+1}{n}\right)^2\right\}\right]$$

şeklindedir.

ii) Eş 2.4' de $d\omega(\pi) = d\pi$ alınırsa

$$S_{n,B} = n\sum_{i=1}^{n-1}\left[F_0^2(X_{(i)})\ln\frac{i+\frac{1}{2}}{i-\frac{1}{2}} - (1-F_0(X_{(i)}))^2\ln\frac{n-i-\frac{1}{2}}{n-i+\frac{1}{2}}\right]$$

olarak elde edilir.

iii) Eş 2.4’ de $d\omega(\pi) = \pi(1-\pi)$ alınırsa

$$S_{n,C} = \sum_{i=1}^{n-1} \left(F_0(X_{(i)}) - \frac{i + \frac{1}{2}}{n} \right)^2$$

ifadesi bulunur.

$S_{n,A}$, $S_{n,B}$, $S_{n,C}$ istatistiklerinin analitik dağılımını elde etmek zordur. Bu yüzden bu istatistiklerin dağılımını oluşturmak için Monte-Carlo yöntemi kullanılır.

3.UYGULAMA

Bu bölümde 2. Bölümde bahsedilen test istatistiklerinin uygulanışı gösterilecektir.

Bir meyve suyu firması üretim yapmakta kullandığı makineden 24 örnek alıp kaç ml dolum yaptığı ölçülmüş ve bu değerler Çizelge 3.1' de verilmiştir.

Çizelge 3.1. 24 örneğe ait dolum miktarları (ml olarak)

87,7	83	90,9	86,5
100,4	87,2	94,2	71,4
85,9	86,7	85,6	80,7
105,5	86	98,5	94,4
74,9	80,4	99,7	97,8
75,3	86,3	108,5	97,9

Çizelge 3.1.' de verilmiş olan verileri $\alpha = 0{,}05$ alınarak χ^2, W_n, D, A^2, U_n, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ testlerine uygulanmıştır. Yokluk ve karşıt hipotezler aşağıdaki gibidir.

H_0 : Örnek normal dağılıma sahip yığından seçilmiştir.

H_1 : Örnek normal dağılıma sahip yığından seçilmemiştir.

Test istatistiklerinde kullanılmak üzere $n=24$ için örnek ortalaması ve varyansı yani $\bar{x}$ ve s^2 değerleri;

$$\bar{x} = \frac{\sum_{i=1}^{24} x_i}{24} = 89{,}3916 \text{ ve } s^2 = \frac{1}{n-1}\sum_{i=1}^{24}(x_i - \bar{x})^2 = 92{,}5659, \quad s = 9{,}6211$$

olarak bulunur.

3.1.Ki-Kare Testinin Uygulanışı

Ki-Kare testi için sınıf sayısı $c=3$ olmak üzere, χ_h^2' de her bir sınıf için hesaplanan beklenen ve gözlenen frekans değerleri yerine yazıldığında;

$$\chi_h^2 = \sum_{j=1}^{c} \frac{\left(G_j - B_j\right)^2}{B_j}$$

$$\chi_h^2 = 2{,}9931$$

elde edilir.

Ki-Kare testi için serbestlik derecesi 2 ve $\alpha = 0{,}05$ anlamlılık düzeyinde tablo değeri de $\chi^2_{2,0,05} = 5{,}991$ olarak elde edilir.

Buradan $\chi_h^2 \langle \chi^2_{2,0.05}$ olduğundan yokluk hipotezi yani H_0 hipotezi red edilemez.

3.2.Cramer-von Mises Testinin Uygulanışı

Cramer-von Mises testi için, örneğin sıralı istatistiklerinin dağılım fonksiyonları bulunur. Her bir x değeri için hesaplanan dağılım fonksiyonları sırası ile W_n' de yerine yazılırsa;

$$W_n = \sum_{i=1}^{n} \left\{ F_0(x_{(i)}) - \frac{2i-1}{2n} \right\}^2 + \frac{1}{12n}$$

$$W_n = 0{,}0664$$

olarak elde edilir.

Cramer-von Mises testi için $n = 24$ ve $\alpha = 0{,}05$ anlamlılık düzeyinde tablo değeri de $W_k = 0{,}217$ olarak bulunur.

Buna göre $W_n \langle W_k$ olduğundan $\alpha = 0{,}05$ anlamlılık düzeyinde H_0 hipotezi red edilemez.

3.3.Kolmogorov- Smirnov Testinin Uygulanışı

Kolmogorov- Smirnov testi için $F_n(x)$ ve $F_0(x)$ dağılım fonksiyonları bulunup,

$$D = \sup_x \left| F_n(x) - F_0(x) \right|$$

D ifadesinde yerine yazıldığında; $n = 24$ ve $x_i = 14$ iken $D_h = 0{,}1508$ olarak bulunur.

Kolmogorov- Smirnov testi için $n = 24$ ve $\alpha = 0{,}05$ anlamlılık düzeyinde tablo değeri de $D_k = 0{,}269$ olarak elde edilir.

Buna göre $D_h \langle D_k$ olduğundan H_0 hipotezi red edilemez.

3.4. Anderson-Darling Testinin Uygulanışı

Anderson-Darling testi için sıralı istatistiklerinin dağılım fonksiyonları bulunup A^2 ifadesinde yerine yazıldığında

$$A^2 = -\frac{2}{n} \sum_{i=1}^{n} \left[\left(i - \frac{1}{2} \right) \log \left\{ F_0(x_{(i)}) \right\} + \left(n - i + \frac{1}{2} \right) \log \left\{ 1 - F_0(x_{(i)}) \right\} \right] - n$$

$$A^2 = -13{,}4303$$

olarak bulunur. Anderson-Darling testi için $n=24$ iken

$$A^{2*} = A^2\left(1+\frac{0.75}{n}+\frac{2.25}{n^2}\right) = -13{,}9017$$

elde edilir.

A^{2*} istatistik değeri tablo değeri olan 0,648 değerinden küçük olduğu için $\alpha = 0{,}05$ anlamlılık düzeyinde H_0 hipotezi red edilemez.

3.5. Watson Testinin Uygulanışı

Watson testinde $n=24$ için

$$\bar{y} = \left(\sum \frac{y_i}{n}\right) = 0{,}4958$$

olarak bulunur. Watson testi için her bir y değeri sırası ile U_n' de yerine yazıldığında;

$$U_n = \sum_{i=1}^{n}\left(y_i - \frac{2i-1}{2n}\right)^2 - n(\bar{y} - 1/2)^2 + \frac{1}{12n} = 0{,}0652$$

elde edilir. $n = 24$ ve $\alpha = 0{,}05$ anlamlılık düzeyinde Watson testi için tablo değeri $U_k = 0{,}185$ olarak bulunur.

Buna göre $U_h \langle U_k$ olduğundan H_0 hipotezi red edilemez.

3.6. Shapiro-Wilk Testinin Uygulanışı

Shapiro-Wilk testinde her bir x değeri sırası ile W ifadesinde yerine yazıldığında;

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0{,}9555$$

olarak bulunur.

Shapiro-Wilk testi için $n = 24$ ve $\alpha = 0{,}05$ anlamlılık düzeyinde tablo değeri de $W_k = 0{,}916$ olarak bulunur.

Buna göre $W_h \rangle W_k$ olduğundan H_0 hipotezi red edilemez.

3.7. Jarque-Bera Testinin Uygulanışı

Jarque-Bera testinde C ve K istatistikleri sırasıyla

$$C = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{3/2}} = 0{,}1126$$

ve

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{2}} = 2{,}3752$$

elde edilir. Bulunan C ve K değerleri sırası ile JB ifadesinde yerine yazıldığında;

$$JB = \frac{n}{6}\left(C^2 + \frac{(K-3)^2}{4}\right) = 0{,}4404$$

olarak bulunur.

Jarque-Bera testi için $n = 24$ ve $\alpha = 0{,}05$ anlamlılık düzeyinde tablo değeri de $JB_k = 5{,}991$ olarak bulunur.

Buna göre $JB_h \langle JB_k$ olduğundan H_0 hipotezi red edilemez.

3.8. Zhang Testinin Uygulanışı

Zhang testi için, örneğin sıralı istatistiklerinin dağılım fonksiyonları bulunur. Her bir x değeri için hesaplanan dağılım fonksiyonları sırası ile Z_K, Z_A ve Z_C ifadelerinde yerine yazıldığında;

$$Z_K = \max_{1\le i\le n}\left(\left(i-\frac{1}{2}\right)\log\left\{\frac{i-\frac{1}{2}}{nF_0(X_{(i)})}\right\} + \left(n-i+\frac{1}{2}\right)\log\left[\frac{n-i+\frac{1}{2}}{n\{1-F_0(X_{(i)})\}}\right]\right) = 0{,}5845$$

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log\{F_0(X_{(i)})\}}{n-i+\frac{1}{2}} + \frac{\log\{1-F_0(X_{(i)})\}}{i-\frac{1}{2}}\right] = 1{,}4298$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{F_0(X_{(i)})^{-1}-1}{(n-\frac{1}{2})/(i-\frac{3}{4})-1}\right\}\right]^2 = 0{,}6356$$

olarak bulunur.

Yapılan Monte Carlo çalışmasına göre bu istatistiklerin kritik değerleri $Z_K = 1{,}4376$, $Z_A = 3{,}4376$ ve $Z_C = 9{,}7507$ olarak elde edilir.

Buna göre, hesaplamalar sonucu elde edilen değerler, kritik değerlerden küçük olduğundan H_0 hipotezi red edilemez.

3.9. Esteban ve Diğerleri Testinin Uygulanışı

Esteban testinde, örneğin sıralı istatistiklerinin dağılım fonksiyonları bulunur. Her bir x değeri için hesaplanan dağılım fonksiyonları sırası ile $S_{n,A}$, $S_{n,B}$, $S_{n,C}$ ifadelerinde yerine yazılıp;

$$S_{n,A} = \max\left\{ \frac{1}{n^2}, S_{n,A}^{1*}, (F_0(X_{(n)}) - 1)^2 \right\} = 0{,}0370$$

$$S_{n,B} = n\sum_{i=1}^{n-1}\left[F_0^2(X_{(i)}) \ln\frac{i+\frac{1}{2}}{i-\frac{1}{2}} - (1 - F_0(X_{(i)}))^2 \ln\frac{n-i-\frac{1}{2}}{n-i+\frac{1}{2}} \right] = 17{,}2896$$

ve

$$S_{n,C} = \sum_{i=1}^{n-1}\left(F_0(X_{(i)}) - \frac{i+\frac{1}{2}}{n} \right)^2 = 0{,}1099$$

değerleri bulunur.

Yapılan Monte Carlo çalışmasına göre bu istatistiklerin kritik değerleri $S_{n,A} = 0{,}0413$, $S_{n,B} = 23{,}8304$ ve $S_{n,C} = 0{,}1734$ olarak elde edilir.

Buna göre, hesaplamalar sonucu elde edilen değerler, kritik değerlerden küçük olduğundan H_0 hipotezi red edilemez.

Genel olarak tüm analizler sonucunda bütün testlerin H_0 hipotezini red edemediği görülmektedir. Yani bütün testlere göre $\alpha = 0,05$ anlamlılık düzeyinde makineden alınan bu 24 dolum miktarının normal dağılıma sahip olduğunu söyleyebiliriz.

4.SİMÜLASYON ÇALIŞMASI

Çalışmanın bu bölümünde, Bölüm 2' de tanıtılan test istatistiklerinin yani χ^2, W_n, D, A^2, U_n, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ uyum iyiliği testlerinin deneysel 1.tip hata oranları ve güçleri bakımından karşılaştırılması simülasyon yoluyla yapılmıştır.

Her test için gerçekte doğru olan H_0 hipotezini red etme oranlarını yani deneysel 1.tip hata oranlarını bulmak amacıyla normal dağılımlı yığınlardan farklı örnek çaplarında örnekler üretilmiştir. Bu örneklerin her biri için H_0 hipotezini test etmek amacıyla χ^2, W_n, D, A^2, U_n, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ uyum iyiliği testlerinin test istatistiklerinin değerleri ve p-değerleri bulunmuştur. Bulunan p-değerleri α değerinden küçük ise yokluk hipotezi yani H_0 reddedilmiştir. Bu işlemler 5000 kez tekrarlanarak her bir test istatistiğinin H_0 hipotezini reddetme sayıları saptanmış ve bunlar 5000 tekrar sayısına bölünerek her bir test istatistiği için deneysel 1.tip hata oranları hesaplanmıştır. Bu yolla bulunan Ki-Kare, Cramer von Mises, Kolmogorov- Smirnov, Anderson-Darling, Watson, Shapiro Wilk, Jarque Bera, Zhang ve Esteban ve diğerleri testlerine ait deneysel 1.tip hata oranları anlamlılık düzeyleri $\alpha = 0{,}01$, $\alpha = 0{,}05$ ve $\alpha = 0{,}10$ iken elde edilmiş ve çizelge 4.1.' de verilmiştir.

Test istatistiklerinin güç bakımından karşılaştırması amacıyla normal dağılım varsayımı altında yığınlardan farklı örnek çaplarında örnekler üretilmiş, üretilen örneklerin her biri için gerçekte yanlış olan H_0 hipotezinin testi için χ^2, W_n, D, A^2, U_n, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ uyum iyiliği testlerinin test istatistiklerinin değerleri ve p-değerleri bulunmuştur. Bu işlemler 5000 kez tekrarlanarak her bir test istatistiğinin H_0 hipotezini reddetme sayıları saptanmış ve bunlar 5000 tekrar sayısına bölünerek her bir test istatistiği için deneysel güç değerleri hesaplanmıştır. Burada gamma, üstel, lognormal, tekdüze, beta, t ve uçdeğer dağılımları normal dağılım alınarak gerçekte yanlış olan H_0 hipotezini red etme oranları yani testlerin güç değerleri elde edilmiş ve sonuçlar çizelge 4.2.- 4.6.' da verilmiştir.

Çalışmada simülasyon çalışması için Matlab R2009A programı kullanılmıştır.

Çizelge 4.1. Normal dağılım altında testlerin deneysel 1. tip hata oranları

α	n	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
0,01	10	0,0000	0,0086	0,0090	0,0094	0,0006	0,0094	0,0116	0,0098	0,0106	0,0114	0,0154	0,0108	0,0134
	20	0,0000	0,0116	0,0130	0,0110	0,0010	0,0180	0,0090	0,0114	0,0124	0,0096	0,0142	0,0118	0,0138
	30	0,0092	0,0130	0,0130	0,0114	0,0000	0,0144	0,0108	0,0116	0,0132	0,0112	0,0124	0,0110	0,0114
	50	0,0122	0,0100	0,0102	0,0108	0,0004	0,0126	0,0094	0,0106	0,0114	0,0096	0,0124	0,0106	0,0118
	100	0,0092	0,0106	0,0106	0,0092	0,0000	0,0094	0,0106	0,0094	0,0094	0,0122	0,0100	0,0112	0,0090
0,05	10	0,0000	0,0484	0,0480	0,0498	0,0038	0,0604	0,0512	0,0480	0,0510	0,0512	0,0522	0,0448	0,0496
	20	0,0000	0,0524	0,0482	0,0508	0,0044	0,0694	0,0520	0,0492	0,0516	0,0482	0,0464	0,0488	0,0530
	30	0,0334	0,0452	0,0446	0,0440	0,0038	0,0516	0,0464	0,0428	0,0508	0,0480	0,0502	0,0502	0,0528
	50	0,0650	0,0560	0,0572	0,0538	0,0034	0,0560	0,0532	0,0532	0,0492	0,0574	0,0482	0,0566	0,0460
	100	0,0554	0,0466	0,0454	0,0438	0,0032	0,0426	0,0510	0,0466	0,0426	0,0500	0,0496	0,0442	0,0450
0,10	10	0,0000	0,1006	0,0972	0,1032	0,0132	0,1160	0,0976	0,0994	0,1026	0,1030	0,1050	0,1094	0,1044
	20	0,0000	0,1116	0,1056	0,1122	0,0160	0,1268	0,1046	0,1084	0,1010	0,1084	0,0992	0,1030	0,0982
	30	0,0540	0,1088	0,1064	0,1040	0,0170	0,1040	0,0928	0,1050	0,0936	0,0978	0,0962	0,1016	0,1044
	50	0,1192	0,0992	0,0918	0,0978	0,0180	0,1010	0,1028	0,1002	0,1040	0,1054	0,0980	0,1088	0,1070
	100	0,1176	0,1040	0,1044	0,1002	0,0146	0,0992	0,0998	0,1074	0,0958	0,1054	0,0974	0,0974	0,0976

Testlerin deneysel 1.tip hata oranlarını elde etmek için farklı örnek büyüklüklerinde yani n=10, 20, 30, 50 ve 100 için simülasyon çalışması yapılmıştır.

Çizelge 4.1.' de elde edilen sonuçlar incelendiğinde, Ki-kare testinin örnek çapı küçük iken deneysel 1. tip hata oranları verilen 1. tip hata oranlarından oldukça küçüktür. Örnek çapı arttığında yani 30 ve daha fazla olduğunda Ki-kare testinin deneysel 1.tip hata oranlarının verilen 1. tip hata oranlarına yakın değerler aldığı gözlemlenmiştir.

Watson testinin tüm örnek çaplarında deneysel 1. tip hata oranlarının verilen 1.tip hata oranlarından oldukça küçük olduğu görülmektedir.

W_n, D, A^2, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ testlerinin tüm örnek çaplarında testlerin deneysel 1. tip hata oranları verilen anlamlılık düzeyleri oldukça yakın sonuçlar vermiştir.

Testlerin güç oranlarını elde etmek için farklı örnek büyüklüklerinde yani n=10, 20, 30, 50 ve 100 için simülasyon çalışması yapılmıştır.

Testlerin güç karşılaştırması için normal dağılıma sahip olmayan yani gamma, üstel, lognormal, tekdüze, beta, *t* ve uçdeğer dağılımlarından elde edilen yığınlar için normallik testlerinin güç değerleri elde edilmiştir. Elde edilen sonuçlar çizelge 4.2.-4.6.' da verilmiştir.

Çizelge 4.2. (0,∞) aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Gamma (3,1)	10	0,0000	0,1416	0,1238	0,1500	0,0216	0,1286	0,1474	0,1230	0,0240	0,1630	0,1938	0,1950	0,2478
	20	0,0000	0,2934	0,2344	0,3316	0,0756	0,2974	0,3128	0,2908	0,0082	0,3790	0,3364	0,3596	0,4660
	30	0,0568	0,4082	0,3252	0,4596	0,1270	0,4432	0,4422	0,4456	0,0030	0,5516	0,4520	0,4718	0,6074
	50	0,2754	0,6646	0,5350	0,7334	0,2740	0,7384	0,6776	0,7610	0,0002	0,8284	0,6410	0,7516	0,7946
	100	0,5444	0,9314	0,8218	0,9660	0,6028	0,9854	0,9644	0,9882	0,0000	0,9934	0,8918	0,9770	0,9726
Gamma (1/3,1)	10	0,0000	0,8172	0,6974	0,8414	0,5344	0,7964	0,6354	0,8008	0,0004	0,8688	0,7802	0,8252	0,8802
	20	0,0000	0,9914	0,9688	0,9948	0,9364	0,9928	0,9314	0,9974	0,0000	0,9972	0,9772	0,9952	0,9956
	30	0,0120	0,9996	0,9994	1,0000	0,9944	1,0000	0,9912	1,0000	0,0000	1,0000	0,9994	1,0000	1,0000
	50	0,9010	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000
	100	0,9144	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000
Üstel(1)	10	0,0000	0,3968	0,3034	0,4244	0,1180	0,3624	0,3452	0,3416	0,0058	0,4620	0,4074	0,4422	0,5300
	20	0,0000	0,7182	0,5762	0,7732	0,3670	0,7450	0,6192	0,7936	0,0000	0,8378	0,7130	0,7948	0,8526
	30	0,1098	0,9012	0,7850	0,9384	0,6028	0,9374	0,8284	0,9682	0,0000	0,9700	0,8674	0,9496	0,9584
	50	0,4476	0,9878	0,9570	0,9960	0,8888	0,9976	0,9776	0,9994	0,0000	0,9984	0,9650	0,9980	0,9950
	100	0,8642	1,0000	1,0000	1,0000	0,9992	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000

Çizelge 4.2. (Devam) $(0,\infty)$ aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Lognormal (0,1)	10	0,0000	0,5636	0,4642	0,5888	0,2750	0,5392	0,5126	0,5014	0,0022	0,6190	0,5608	0,6154	0,6812
	20	0,0000	0,8784	0,7796	0,9010	0,6496	0,8852	0,8214	0,8974	0,0000	0,9250	0,8454	0,9060	0,9366
	30	0,2082	0,9742	0,9306	0,9852	0,8660	0,9842	0,9526	0,9918	0,0000	0,9918	0,9548	0,9876	0,9920
	50	0,7112	0,9992	0,9944	0,9998	0,9868	1,0000	0,9986	1,0000	0,0000	1,0000	0,9960	1,0000	0,9998
	100	0,8426	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000
Lognormal (0,2)	10	0,0000	0,8908	0,8196	0,9038	0,7108	0,8762	0,7952	0,8708	0,0004	0,9182	0,8702	0,9044	0,9356
	20	0,0000	0,9970	0,9910	0,9984	0,9786	0,9984	0,9814	0,9990	0,0000	0,9990	0,9934	0,9986	0,9984
	30	0,0052	1,0000	1,0000	1,0000	1,0000	1,0000	0,9994	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000
	50	0,4598	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000
	100	0,2932	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,0000	1,0000	1,0000	1,0000	1,0000

Çizelge 4.2.' yi incelediğimizde, gamma, üstel ve lognormal dağılımları için $(0,\infty)$ aralığında simetrik olmayan dağılımlara karşı testlerin güç değerlerinin elde edildiği görülmektedir.

Buna göre ilk önce $\alpha=3$, $r=1$ parametreli gamma dağılımına sahip yığınlar için sonuçları incelediğimizde, χ^2 testi örnek çapı 100 iken iyi sonuç vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin güç değerlerinin oldukça iyi olduğu görülmektedir. Özellikle örnek çapı 100 iken bu testler oldukça iyi sonuçlar vermişlerdir. Ayrıca daha ayrıntılı olarak incelediğimizde özellikle Z_A ve Z_K testleri diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Diğer testler içinde $\mathrm{S}_{\mathrm{n,C}}$ ve W testlerinin Z_A ve Z_K testlerine daha yakın sonuçlar verdiği görülmektedir.

Aynı şekilde Gamma dağılımı için $\alpha=1/3, r=1$ alındığında χ^2 testi örnek çapı 50 ve daha fazla iken iyi sonuçlar vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin özellikle örnek çapı 20 ve daha fazla iken oldukça iyi sonuçlar verdiği gözlenmektedir. Burada da $\alpha=3$, $r=1$ parametreli gamma dağılımına sahip yığınlarda elde edilen güç değerleri sonuçlarında olduğu gibi özellikle Z_A ve Z_K testleri diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Ayrıca diğer testler içinde $S_{n,B}$ ve $S_{n,C}$ testlerinin Z_A ve Z_K testlerine yakın güç değerleri verdiği gözlenmektedir.

Çarpık dağılım olan üstel dağılım için χ^2 testi örnek çapı 100 iken iyi sonuç vermiştir. Diğer testler içinde özellikle örnek çapı 30 ve daha fazla iken Z_C testi hariç tüm testler oldukça oldukça iyi sonuçlar vermişlerdir. Gamma dağılımında olduğu gibi özellikle Z_A ve Z_K testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğunu söylenebilir. Yine diğer testler içinde $S_{n,B}$, $S_{n,C}$, W ve A^2 testlerinin Z_A ve Z_K testlerine daha yakın sonuçlar verdiği gözlenmektedir.

Lognormal dağılımı için $\mu = 0$, $\sigma^2 = 1$ alındığında χ^2 testi örnek çapı 50 ve daha fazla iken iyi sonuçlar vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin oldukça yüksek güç değerlerine sahip olduğu görülmektedir. Özellikle örnek çapı 20 ve daha fazla iken $S_{n,C}$ ve Z_A testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir. Diğer testler içinde ise Z_K, $S_{n,B}$ ve A^2 testlerinin oldukça iyi sonuçlar verdiği söylenebilir.

Aynı şekilde $\mu = 0$, $\sigma^2 = 2$ parametreli lognormal dağılım için incelendiğinde χ^2 testi örnek çapı 50 ve daha fazla iken iyi sonuçlar vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin yine oldukça yüksek güç değerleri verdiği görülmektedir. Burada da özellikle örnek çapı 10 ve daha fazla iken Z_A ve Z_K testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu gözlenmektedir. Diğer testleri incelediğimizde $S_{n,B}$, $S_{n,C}$, A^2 ve W testlerinin de oldukça iyi sonuçlar verdiği söylenebilir.

Çizelge 4.3. (0,1) aralığında simetrik dağılımlara karşı testlerin güç değerleri

Dağılım	n	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Tekdüze (0,1)	10	0,0000	0,0698	0,0574	0,0756	0,0070	0,0412	0,0156	0,0684	0,0216	0,0864	0,0502	0,0262	0,0392
	20	0,0000	0,1436	0,0956	0,1698	0,0176	0,0978	0,0014	0,1256	0,0060	0,2280	0,0896	0,0690	0,0880
	30	0,0000	0,2234	0,1428	0,2952	0,0416	0,2226	0,0002	0,2064	0,0020	0,4380	0,1476	0,1780	0,1750
	50	0,1752	0,4412	0,2630	0,5670	0,1280	0,5764	0,0088	0,4748	0,0000	0,8264	0,2748	0,4746	0,4024
	100	0,4132	0,8416	0,5972	0,9508	0,5268	0,9878	0,7402	0,9740	0,0000	0,9992	0,5716	0,9536	0,8058
Beta (2,2)	10	0,0000	0,0430	0,0404	0,0412	0,0018	0,0410	0,0186	0,0434	0,0452	0,0380	0,0472	0,0228	0,0400
	20	0,0000	0,0600	0,0466	0,0580	0,0042	0,0352	0,0036	0,0486	0,0362	0,0514	0,0572	0,0296	0,0504
	30	0,0016	0,0710	0,0588	0,0762	0,0094	0,0426	0,0020	0,0604	0,0232	0,0834	0,0710	0,0404	0,0666
	50	0,0726	0,1120	0,0766	0,1328	0,0156	0,0744	0,0008	0,0836	0,0078	0,1778	0,0832	0,0950	0,0974
	100	0,1578	0,2414	0,1498	0,3114	0,0428	0,2808	0,0406	0,2042	0,0002	0,5382	0,1528	0,2820	0,2378
Beta (15,15)	10	0,0000	0,0470	0,0476	0,0460	0,0026	0,0528	0,0414	0,0462	0,0488	0,0426	0,0400	0,0360	0,0400
	20	0,0000	0,0452	0,0434	0,0430	0,0022	0,0546	0,0330	0,0462	0,0502	0,0406	0,0444	0,0408	0,0470
	30	0,0212	0,0510	0,0478	0,0502	0,0056	0,0494	0,0326	0,0468	0,0550	0,0444	0,0452	0,0454	0,0486
	50	0,0580	0,0474	0,0494	0,0486	0,0048	0,0410	0,0220	0,0452	0,0550	0,0382	0,0448	0,0456	0,0492
	100	0,0538	0,0418	0,0426	0,0408	0,0042	0,0344	0,0168	0,0380	0,0538	0,0334	0,0488	0,0414	0,0394

Çizelge 4.3.' ü incelediğimizde, tek düze ve beta dağılımları için $(0,1)$ aralığında simetrik dağılımlara karşı testlerin güç değerlerinin elde edildiği görülmektedir.

Buna göre tek düze dağılımlı yığınlardaki simülasyon çalışması sonucu χ^2 testi örnek çapı 100 iken iyi sonuç vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin örnek çapı arttıkça güç değerlerinin oldukça iyi olduğu görülmektedir. Özellikle örnek çapı 100 iken bu testler oldukça iyi sonuçlar vermişlerdir. Daha ayrıntılı olarak incelediğimizde özellikle Z_A ve W testleri diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Ayrıca diğer testler içinde $S_{n,B}$ ve Z_K testlerinin Z_A ve W testlerine yakın güç değerleri verdiği görülmektedir.

Beta dağılımı için $\alpha = 2$, $\beta = 2$ alındığında Z_C testi hariç tüm testlerin özellikle örnek çapı 100 iken iyi sonuçlar verdiği gözlenmektedir. Özellikle Z_A testi diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Diğer testler içinde de W testinin Z_A testine yakın değerler verdiği gözlenmektedir.

Genel olarak incelediğimizde beta dağılımında α ve β parametrelerinin değerleri arttıkça tüm testlerin güç değerlerinin azaldığı görülmektedir. Tüm testleri karşılaştırdığımızda ise Z_A testinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir.

Çizelge 4.4.(0,1) aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Beta (1,2)	10	0,0000	0,1164	0,0968	0,1218	0,0160	0,0866	0,0680	0,1014	0,0206	0,1380	0,0174	0,0162	0,0072
	20	0,0000	0,2202	0,1732	0,2582	0,0384	0,1856	0,0796	0,2410	0,0024	0,2990	0,0148	0,1102	0,0288
	30	0,0010	0,3564	0,2556	0,4160	0,0918	0,3424	0,0906	0,4742	0,0000	0,5082	0,0398	0,3020	0,1024
	50	0,1734	0,6082	0,4372	0,7164	0,2112	0,7116	0,1982	0,8734	0,0000	0,8502	0,1458	0,6974	0,3198
	100	0,4064	0,9370	0,8204	0,9820	0,6484	0,9950	0,8440	0,9998	0,0000	0,9994	0,6254	0,9930	0,8410
Beta (1,3)	10	0,0000	0,1690	0,1362	0,1826	0,0198	0,1370	0,1288	0,1468	0,0134	0,2056	0,2256	0,2062	0,2740
	20	0,0000	0,3562	0,2632	0,4060	0,0820	0,3370	0,2180	0,4036	0,0008	0,4802	0,3684	0,3838	0,4916
	30	0,0072	0,5484	0,4112	0,6268	0,1720	0,5896	0,3200	0,7174	0,0000	0,7278	0,5384	0,6346	0,6934
	50	0,2098	0,8072	0,6436	0,8884	0,4210	0,9076	0,5962	0,9790	0,0000	0,9630	0,7460	0,9234	0,9042
	100	0,5148	0,9912	0,9500	0,9990	0,8704	0,9996	0,9856	1,0000	0,0000	1,0000	0,9644	0,9998	0,9980
Beta (1,4)	10	0,0000	0,2148	0,1742	0,2286	0,0332	0,1760	0,1688	0,1818	0,0126	0,2548	0,2512	0,2490	0,3276
	20	0,0000	0,4478	0,3264	0,5052	0,1242	0,4504	0,3296	0,5102	0,0006	0,5818	0,4490	0,4942	0,6056
	30	0,0184	0,6484	0,5006	0,7302	0,2670	0,7130	0,4812	0,8190	0,0004	0,8250	0,6428	0,7582	0,8034
	50	0,2466	0,8930	0,7628	0,9450	0,5446	0,9584	0,7780	0,9922	0,0000	0,9852	0,8352	0,9646	0,9512
	100	0,5612	0,9988	0,9864	1,0000	0,9422	1,0000	0,9984	1,0000	0,0000	1,0000	0,9880	1,0000	0,9998

Çizelge 4.4.(Devam) (0,1) aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Beta (2,1)	10	0,0000	0,1184	0,1034	0,1238	0,0140	0,0816	0,0620	0,1108	0,0218	0,1284	0,1650	0,1162	0,1778
	20	0,0000	0,2220	0,1680	0,2532	0,0376	0,1750	0,0742	0,2372	0,0032	0,2950	0,2672	0,2068	0,3310
	30	0,0014	0,3598	0,2662	0,4226	0,0818	0,3526	0,0972	0,4768	0,0004	0,5096	0,3836	0,3934	0,4848
	50	0,1690	0,6080	0,4552	0,7130	0,2152	0,7062	0,2052	0,8694	0,0000	0,8474	0,5928	0,7342	0,7096
	100	0,4130	0,9352	0,8086	0,9822	0,6458	0,9950	0,8348	1,0000	0,0000	0,9994	0,8610	0,9926	0,9612
Beta (3,1)	10	0,0000	0,1646	0,1322	0,1816	0,0206	0,1338	0,1210	0,1382	0,0168	0,2002	0,0086	0,0230	0,0040
	20	0,0000	0,3632	0,2680	0,4120	0,0748	0,3414	0,2228	0,4062	0,0014	0,4850	0,0152	0,2154	0,0442
	30	0,0060	0,5480	0,3934	0,6260	0,1790	0,5810	0,3186	0,7072	0,0000	0,7252	0,0570	0,5226	0,1676
	50	0,2142	0,8068	0,6438	0,8862	0,4080	0,9008	0,5850	0,9746	0,0000	0,9578	0,2906	0,8876	0,5256
	100	0,5100	0,9932	0,9564	0,9984	0,8678	0,9998	0,9842	1,0000	0,0000	1,0000	0,8842	1,0000	0,9602
Beta (4,1)	10	0,0000	0,2258	0,1742	0,2410	0,0368	0,1824	0,1712	0,1872	0,0110	0,2618	0,0074	0,0394	0,0038
	20	0,0000	0,4464	0,3296	0,5042	0,1400	0,4384	0,3160	0,5076	0,0004	0,5810	0,0226	0,3164	0,0784
	30	0,0148	0,6540	0,5014	0,7278	0,2614	0,7138	0,4834	0,8182	0,0000	0,8314	0,1186	0,6492	0,2480
	50	0,2424	0,8920	0,7580	0,9476	0,5630	0,9606	0,7802	0,9928	0,0000	0,9834	0,4356	0,9490	0,6348
	100	0,5792	0,9986	0,9858	0,9998	0,9418	0,9998	0,9976	1,0000	0,0000	1,0000	0,9670	0,9998	0,9870

Çizelge 4.4.' ü incelediğimizde, beta dağılımı için $(0,1)$ aralığında simetrik olmayan dağılımlara karşı testlerin güç değerlerinin elde edildiği görülmektedir.

Beta dağılımı için $\alpha = 1$, $\beta = 2$ alındığında χ^2 testi örnek çapı 100 iken iyi sonuç vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin özellikle örnek çapı 100 iken oldukça iyi sonuçlar verdiği gözlenmektedir. Özellikle Z_A ve Z_K testleri diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Ayrıca diğer testler içinde de W , $S_{n,B}$ ve $S_{n,C}$ testlerinin Z_A ve Z_K testlerine yakın değerler verdiği gözlenmektedir.

Genel olarak, simülasyon çalışmasında alınan değerlere göre beta dağılımında α ve β parametreleri için örnek hacmi ne olursa olsun tüm testleri karşılaştırdığımızda ise, Z_A ve Z_K testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir. Ayrıca diğer testler içinde de W , $S_{n,B}$ ve $S_{n,C}$ testlerinin Z_A ve Z_K testlerine yakın değerler verdiği gözlenmektedir.

Çizelge 4.5. $(-\infty,\infty)$aralığında simetrik dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
t_1	10	0,0000	0,6142	0,5764	0,6130	0,4220	0,5814	0,5878	0,5786	0,0080	0,5764	0,4072	0,6042	0,4694
	20	0,0000	0,8812	0,8500	0,8846	0,7712	0,8688	0,8668	0,8610	0,0030	0,8504	0,7840	0,8900	0,8394
	30	0,7014	0,9620	0,9406	0,9620	0,9110	0,9530	0,9482	0,9498	0,0004	0,9376	0,9188	0,9638	0,9512
	50	0,7682	0,9972	0,9950	0,9968	0,9920	0,9962	0,9950	0,9940	0,0000	0,9930	0,9928	0,9976	0,9962
	100	0,6812	1,0000	1,0000	1,0000	1,0000	0,9998	1,0000	1,0000	0,0000	0,9998	1,0000	1,0000	1,0000
t_2	10	0,0000	0,3010	0,2712	0,3036	0,1332	0,2950	0,3192	0,2800	0,0286	0,2924	0,1984	0,3134	0,2272
	20	0,0000	0,5262	0,4686	0,5422	0,3266	0,5350	0,5766	0,5182	0,0110	0,5286	0,4022	0,5758	0,4662
	30	0,5396	0,6618	0,5918	0,6798	0,4762	0,6786	0,7140	0,6466	0,0070	0,6600	0,5284	0,7048	0,6178
	50	0,7102	0,8466	0,7834	0,8646	0,7022	0,8642	0,8886	0,8300	0,0026	0,8434	0,7524	0,8734	0,8394
	100	0,8740	0,9788	0,9568	0,9830	0,9452	0,9838	0,9866	0,9678	0,0000	0,9748	0,9534	0,9824	0,9764
t_3	10	0,0000	0,1840	0,1614	0,1904	0,0564	0,1892	0,2176	0,1694	0,0368	0,1904	0,1176	0,1860	0,1392
	20	0,0000	0,3144	0,2622	0,3328	0,1396	0,3446	0,4012	0,3072	0,0230	0,3454	0,2130	0,3690	0,2720
	30	0,3404	0,4144	0,3438	0,4404	0,2280	0,4580	0,5202	0,4212	0,0132	0,4564	0,2986	0,4656	0,3780
	50	0,4288	0,5700	0,4832	0,6006	0,3536	0,6324	0,6852	0,5790	0,0094	0,6140	0,4374	0,6162	0,5384
	100	0,6328	0,8200	0,7324	0,8510	0,6360	0,8740	0,9042	0,8142	0,0006	0,8522	0,7002	0,8596	0,8116

Çizelge 4.5.(Devam) (-∞,∞) aralığında simetrik dağılımlara karşı testlerin güç değerleri

Dağılım	n	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
t_{20}	10	0,0000	0,0632	0,0618	0,0618	0,0048	0,0672	0,0688	0,0610	0,0452	0,0634	0,0528	0,0562	0,0556
	20	0,0000	0,0604	0,0574	0,0674	0,0060	0,0786	0,0818	0,0640	0,0408	0,0736	0,0590	0,0746	0,0604
	30	0,0466	0,0606	0,0602	0,0650	0,0060	0,0796	0,0914	0,0682	0,0460	0,0808	0,0566	0,0622	0,0590
	50	0,0710	0,0694	0,0658	0,0742	0,0080	0,0836	0,1148	0,0802	0,0380	0,0954	0,0696	0,0832	0,0764
	100	0,0692	0,0780	0,0692	0,0842	0,0120	0,1078	0,1536	0,0982	0,0418	0,1164	0,0686	0,0882	0,0792
t_{30}	10	0,0000	0,0586	0,0510	0,0576	0,0034	0,0640	0,0586	0,0516	0,0500	0,0562	0,0480	0,0550	0,0574
	20	0,0000	0,0584	0,0524	0,0616	0,0066	0,0782	0,0736	0,0586	0,0462	0,0666	0,0496	0,0662	0,0560
	30	0,0430	0,0544	0,0526	0,0544	0,0050	0,0676	0,0722	0,0638	0,0468	0,0636	0,0550	0,0644	0,0576
	50	0,0692	0,0534	0,0546	0,0580	0,0052	0,0706	0,0900	0,0666	0,0408	0,0768	0,0564	0,0636	0,0522
	100	0,0642	0,0650	0,0560	0,0676	0,0052	0,0788	0,1140	0,0748	0,0462	0,0902	0,0632	0,0652	0,0714

Çizelge 4.5.' te t dağılımları altında $(-\infty,\infty)$ aralığında simetrik dağılımlara karşı testlerin güç değerleri elde edilmiştir.

Buna göre t_1 dağılımı alındığında χ^2 testi örnek çapı 30 ve daha fazla iken iyi sonuçlar vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin oldukça iyi sonuçlar verdiği söylenebilir. Özellikle örnek çapı 20 ve daha fazla iken $S_{n,B}$ ve W_n testleri diğer testlere göre daha yüksek güç değerlerine sahiptir. Diğer testler içinde de $S_{n,C}$ ve A^2 testlerinin oldukça iyi sonuçlar verdiği görülmektedir.

t_2 dağılımı alındığında χ^2 testi örnek çapı 30 ve daha fazla iken iyi sonuçlar vermiştir. Diğer testler içinde Z_C testi hariç tüm testlerin oldukça iyi sonuçlar verdiği gözlenmektedir. Özellikle örnek çapı 50 ve daha fazla iken JB ve W testleri oldukça yüksek güç değerlerine sahip olduğu görülmektedir. Ayrıca diğer testler içinde A^2 ve $S_{n,B}$ testlerinin oldukça iyi sonuçlar verdiği söylenebilir.

Genel olarak incelediğimizde de t dağılımında serbestlik derecesi arttıkça tüm testlerin güç değerlerinin azaldığı görülmektedir. Tüm testleri karşılaştırdığımızda ise JB testinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir.

Çizelge 4.6. (-∞,∞) aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri

Dağılım	*n*	χ^2	W_n	D	A^2	U_n	W	JB	Z_K	Z_C	Z_A	$S_{n,A}$	$S_{n,B}$	$S_{n,C}$
Uçdeğer (0,1)	10	0,0000	0,1328	0,1168	0,1370	0,0218	0,1224	0,1366	0,1174	0,0314	0,1428	0,0108	0,0358	0,0080
	20	0,0000	0,2472	0,1994	0,2714	0,0530	0,2526	0,2968	0,2362	0,0148	0,3072	0,0136	0,1852	0,0380
	30	0,0838	0,3544	0,2882	0,3952	0,1040	0,3814	0,4172	0,3594	0,0068	0,4528	0,0562	0,3552	0,1150
	50	0,2596	0,5362	0,4362	0,5958	0,1928	0,5894	0,6100	0,5732	0,0012	0,6710	0,1570	0,5956	0,2928
	100	0,5398	0,8470	0,7338	0,8924	0,5018	0,9154	0,9156	0,8790	0,0002	0,9400	0,5288	0,9104	0,7226
Uçdeğer (0,3)	10	0,0000	0,1326	0,1152	0,1392	0,0260	0,1260	0,1482	0,1182	0,0338	0,1492	0,0166	0,0450	0,0120
	20	0,0000	0,2524	0,1972	0,2760	0,0624	0,2594	0,2928	0,2348	0,0158	0,3158	0,0194	0,1890	0,0422
	30	0,0802	0,3628	0,2902	0,4014	0,1140	0,3902	0,4212	0,3624	0,0090	0,4682	0,0530	0,3516	0,1144
	50	0,2632	0,5352	0,4318	0,5914	0,1910	0,5988	0,6134	0,5640	0,0042	0,6732	0,1628	0,5930	0,2980
	100	0,5388	0,8510	0,7376	0,8888	0,4974	0,9076	0,9056	0,8726	0,0000	0,9308	0,5346	0,9076	0,7284

Uçdeğer dağılımları alındığında, $(-\infty,\infty)$ aralığında simetrik olmayan dağılımlara karşı testlerin güç değerleri çizelge 4.6.' da elde edilmiştir.

Buna göre, Uçdeğer dağılımları için χ^2 testi örnek çapı 100 iken iyi sonuç vermiştir. Diğer testler içinde örnek çapı arttıkça Z_C testi hariç tüm testlerin oldukça iyi sonuçlar verdiği söylenebilir. Özellikle Z_A testinin en iyi güç değerlerine sahip olduğu görülmektedir. Diğer testler arasında ise JB, W ve $S_{n,B}$ testlerinin oldukça iyi sonuçlar verdiği söylenebilir.

5.SONUÇ

Bu çalışmada normal dağılım için χ^2, W_n, D, A^2, U_n, W, JB, Z_K, Z_C, Z_A, $S_{n,A}$, $S_{n,B}$ ve $S_{n,C}$ uyum iyiliği testleri incelenmiş ve simülasyon yolu ile karşılaştırılmıştır. Testlerin deneysel 1. tip hata bakımından ve testin gücü bakımından karşılaştırılmaları yapılmıştır.

Deneysel 1.tip hata oranları bakımından Ki-Kare testinin örnek çapı küçük iken verilen 1.tip hata oranlarından oldukça küçük değerler verdiği, örnek çapı 30 ve daha fazla olduğunda ise verilen 1. tip hata oranlarına yakın değerler verdiği gözlemlenmiştir. Watson testinin tüm örnek çaplarında verilen 1. tip hata oranlarından oldukça küçük değerler verdiği görülmüştür. Ki-Kare ve Watson testleri dışındaki diğer testlerde ise tüm örnek çaplarında oldukça iyi sonuçlar verdiği gözlenmiştir.

Testin gücü bakımından değerlendirildiğinde gamma dağılımına sahip yığınlar için $(0,\infty)$ aralığında simetrik olmayan dağılımlara karşı küçük örnek çaplarında $S_{n,C}$, örnek çapı arttıkça Z_A ve Z_K testlerinin, ayrıca $S_{n,B}$ ve $S_{n,C}$ testlerinin de yüksek güç değerlerine sahip olduğu görülmektedir.

Çarpık dağılımlardan olan üstel dağılıma göre $(0,\infty)$ aralığında simetrik olmayan dağılımlara karşı özellikle örnek çapı 30 ve daha fazla iken tüm testler oldukça iyi sonuçlar vermişlerdir. Özellikle Z_K ve Z_A testleri diğer testlere göre çok yüksek güç değerlerine sahip olduğu gözlenmiştir. Bu testlerin ardından en iyi sonucu $S_{n,B}$, $S_{n,C}$, W ve A^2 testleri vermişlerdir.

Lognormal dağılım için ise $(0,\infty)$ aralığında simetrik olmayan dağılımlara karşı Z_A ve $S_{n,C}$ testlerinin diğer testlere göre en iyi güç değerlerine sahip olduğu gözlenmiştir. Diğer testler arasında da Z_K, $S_{n,B}$, A^2 ve W testleri için oldukça yüksek güç değerleri elde edilmiştir.

Tek düze dağılımına göre $(0,1)$ aralığında simetrik dağılımlara karşı özellikle Z_A ve W testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu söylenebilir. Diğer testler arasında da $S_{n,B}$ ve Z_K testleri için oldukça yüksek güç değerleri elde edilmiştir.

Beta dağılımı için $(0,1)$ aralığında simetrik dağılımlara karşı α ve β parametrelerinin değerleri arttıkça tüm testlerin güç değerlerinin azaldığı görülmüştür. Tüm testleri karşılaştırdığımızda Z_A testinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmüştür.

Beta dağılımı için $(0,1)$ aralığında simetrik olmayan dağılımlara karşı α ve β parametreleri için örnek hacmi ne olursa olsun tüm testleri karşılaştırdığımızda Z_A ve Z_K testlerinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu gözlenmiştir. Diğer testler arasında da W, $S_{n,B}$ ve $\mathrm{S}_{\mathrm{n,C}}$ testleri için oldukça yüksek güç değerleri elde edilmiştir.

t dağılımı için incelediğimizde ise $(-\infty,\infty)$ aralığında simetrik dağılımlara karşı serbestlik derecesi arttıkça tüm testlerin güç değerlerinde azalma olduğu görülmüştür. JB testinin güç değerlerinin diğer testlere göre oldukça yüksek olduğu gözlenmektedir.

Uçdeğer dağılımı için ise $(-\infty,\infty)$ aralığında simetrik olmayan dağılımlara karşı Z_A testi en iyi sonucu verirken diğer testler içinde de JB, $S_{n,B}$ ve W testlerinin en iyi güç değerlerine sahip olduğu gözlenmiştir.

KAYNAKLAR

Anderson, Jr.,T.W., Darling, D.A., “Asymptotic theory of certain ‘goodness-of-fit’ criteria based on stochastic processes”, ***Annals of Mathematical Statistics,*** 23, 193-212. (1952).

Anderson, Jr.,T.W., Darling, D.A., “A test of goodness-of-fit ”, ***Journal of the American Statistical Association,*** 49:765-769. (1954).

Anderson, TW., “On the Distribution of the Two-Sample Cramer-Von Mises Criterion”, ***The Annals of Mathematical Statistics (Institute of Mathematical Statistics),*** 33 (3): 1148–1159. (1962).

Baringhaus, L., Henze, N., “A consistent test for multivariate normality based on the empirical characteristic function.”, ***Metrika,*** 35:339-348. (1988).

Bera, A.K., Jarque, C.M., “Efficient tests for normality, homoscedasticity and serial independence of regression residuals.”, ***Economics Letters,*** 6 (3): 255–259. (1980).

Bera, A.K., Jarque, C.M., “Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence.”, ***Economics Letters***, 7 (4):313–318. (1981).

Bera, A.K., Bilias, Y., “The MM, ME, ML, EL, EF and GMM approaches to estimation: a synthesis.”, ***Journal of Econometrics***, 107:51-86. (2002).

Cox, D. R., “Tests of separate families of hypotheses.”, ***Proc. 4th Berkeley Symp.,*** 6, 105-23. (1961).

Cox, D. R., “Further results on tests of separate families of hypotheses.”, ***J. R. Statist. Soc. B,*** 24, 406-24. (1962).

Cramer, H., “On the composition of elementary errors.”, ***Skand. Aktuar.,*** 11:141-180. (1928).

Cressie, N., “Power results for tests based on high order gaps.”, ***Biometrika,*** 65:214-218. (1978).

Cressie, N., “An optimal statistic based on higher order gaps.”, ***Biometrika***, 66:619-627. (1979).

Cressie, N., Read, T.R.C., “Multinomial goodness-of-fit tests.”, ***Journal of the Royal Statistical Society. Series B,*** 46:440-464. (1984).

Cressie, N., Read, T.R.C., “Goodness-of-Fit Statistics for Discrete Multivariate Data.”, ***Springer, New York.,*** (1988).

David, F.N., Johnson, N.L., “The Probability Integral Transformation When *Parameters are Estimated from the sample ”,* ***Biometrika,*** *35:182-190 (1948).*

Dong, L.B., Giles, D.E.A., “An empirical likelihood ratio test for normality.”, ***Communications in Statistics. Simulation and Computation,*** 36:197-215 (2007).

Epps, T. W., Singleton, K. J., Pulley, L. B., “A test of separate families of distributions based on the empirical moment generating function.”, ***Biometrika,*** 69.2:391–399. (1982).

Epps, T.W., Pulley, L. B., “A test for normality based on the empirical characteristic function.”, ***Biometrika,*** 70(3):723–726. (1983).

Epps, T.W., “Tests for location-scale families based on the empirical characteristic function.”, ***Metrika,*** 62: 99-114. (2005).

Esteban,M.D., Castellanos, M.E., Morales, D., Vajda, I., “Monte Carlo comparison of four normality tests using different entropy estimates.”, ***Communications in Statistics-Simulation and Computation,*** 30:761-785. (2001).

Esteban,M.D., Marhuenda, Y.,Morales, D., Sanchez, A., “Goodness-of Fit Tests. New Goodness-of Fit Tests Based on Sample Quantiles.”, ***Communications in Statistics-Simulation and Computation,*** 36:631-642. (2007).

Fan, J., “Test of significance based on wavelet thresholding and Neyman’s truncation. ”, ***J.A.S.A,*** 91:674–688. (1996).

Fan, Y., “Goodness-of-fit tests for a multivariate distribution by the empirical characteristic function. ”, ***J. Multivariate Anal.***, 62:36-63. (1997).

Fan, Y., “Goodness-of-fit tests based on kernel density estimators with fixed smoothing parameters.”, ***Econom. Theory,*** 14: 604-621. (1998).

Klar, B., “Goodness-of-fit tests for the exponential and the normal distributions based on the integrated distribution function.”, **Ann. Inst. Statist. Math.,** 52(2):338–353. (2001).

Kolmogorov, A.N., “Sulla determinazione empirica di une legge di distribuzione.”, ***G. Ist. Attuari.,*** 83-91. (1933).

Kuiper, N.H., “Teste concerning random points on a circle.”, ***Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A.,*** Vol.63:38-47. (1962).

Ledwina, T., “Data-Driven Version of Neyman's Smooth Test of Fit.”, ***Journal of the American Statistical Association,*** 89:1000-1005. (1994).

Lilliefors, H.W., “On the Kolmogorov-Smirnov test for normality with mean and variance unknown.”, ***Journal of the American Statistical Association,*** 62(318):399-402. (1967).

Massey, F.J.Jr., “The Kolmogorov-Smirnov Test for Goodness of Fit”, ***Journal of the American Statistical Association,*** 46(253) : 68-78 (1951).

Matsui, M., Tamura, A., “Empirical characteristic function approach to goodness-of fit tests for the Cauchy distribution with parameters estimated by MLE or EISE.”, ***Ann. Inst. Statist. Math.,*** 57:183–199. (2005).

Owen, A.B., “Empirical Likelihood.”, ***New York: Chapman and Hall,*** (2001).

Pearson, K., “Skew variation in homogeneous material.”, ***Philosophical Transactions of the Royal Society of London. Series A,*** 186:343-414. (1895).

Pearson, K., “On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling.***”, Philosophical Magazine Series 5,*** 50: 157-175. (1900).

Pearson, K., “Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder ”, ***Biometrika,*** 4:169-212. (1905).

Pettitt, A.N., “Testing the Normality of Several Independent Samples Using the Anderson Darling Statistics.”, ***Applied Statistics,*** 26(2):156-161. (1977).

Romao, X.,Delgado, R.,Costa, A., “An empirical power comparison of univariate goodness-of –fit tests for normality.”, ***Journal of Statistical Computation and Simulation***, 0:1-47. (2009).

Seier, E., “Comparison of Tests for Univariate Normality”, ***http://interstat.statjournals.net/YEAR/2002/articles/0201001.pdf.,*** Erişim Tarihi: 11.12.2006 (2002).

Shan, G., Vexler, A., Wilding, G.E., Hutson, A.D., “Simple and Exact Empirical Likelihood Ratio Tests for Normality Based on Moment Relations”, ***Communications in Statistics-Simulation and Computation,*** 40:129-146. (2011).

Shapiro, S.S., Wilk, M.B., “An Analysis of Variance Test for Normality (Complete Samples) ”, ***Biometrika,*** 52, 3/4.591-611. (1965).

Shapiro, S.S., Francia, R.S., “An Approximate Analysis of Variance Test for. Normality.”, ***Journal of the American Statistical Association.,*** 67 (337):215-216. (1972).

Smirnov, N. V., “On the estimation of the discrepancy between empirical curves of distribution for two independent samples.”, ***(Russian) Bulletin of Moscow University,*** 2, 3-16. (1939).

Stephens, M.A., “The Distribution of The Goodness of Fit Statistic, U_n^2 -II”, ***Biometrika,*** 51, 3-4: 393-397 (1964).

Stephens, M.A., “EDF Statistics for Goodness of Fit and Some Comparisons.”, ***Journal of the America Statistical Association,*** No.347. pp.730-737. (1974).

Towhidi, M., Salmanpour, M., “A new goodness-of-fit test based on the empirical characteristic function.”, ***Communications in Statistics- Theory and Methods,*** 36:2777-2785. (2007).

Vasicek, O., “A test for normality based on sample entropy.”, ***Journal of Royal Statistical Society B,*** (38):54-59 (1976).

Vexler, A., Gurevich, G., “Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy.”, ***Computational Statistics and Data Analysis,*** 54:531-545 (2010).

von-Mises, R., “Wahrcheinlichkeitsrech nung und Ihre Anwendung in der Statistik und Theoretischen phisik. ”, ***Lepzig: Deutike.,*** (1931).

Zhang, J., “Powerpull goodness-of-fit tests based on the likelihood ratio”, ***Canada,*** (2002).

Zghoul, Ahmad A., “A goodness of fit test for normality based on the empirical moment generating function”, ***Communications in Statistics, Simulation and Computation,*** 39:1292-1304. (2010).

Watson, G.S., “Goodness of fit Tests On a Circle-I”, ***Biometrika,*** 48:109-114. (1961).

Watson, G.S., “Goodness of fit Tests On a Circle-II”, ***Biometrika,*** 49:57-63. (1962).

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : YILDIRIM, Nurcan

Uyruğu : T.C.

e-mail : nurcan_zeynep@mynet.com

Eğitim

Derece	***Eğitim Birimi***	***Mezuniyet tarihi***
Yüksek Lisans	Gazi Üniversitesi /Matematik Bölümü	2003
Lisans	Cumhuriyet Üniversitesi/ Matematik Bölümü	1998

İş Deneyimi

Yıl	**Yer**	**Görev**
1998-2013	M.E.B	Matematik Öğretmeni

Yabancı Dil

İngilizce

Yayınlar

1. Yıldırım, N., “1.Dereceden Lineer ve Yarı Lineer Denklem Sistemleri”, Gazi Üniversitesi Yüksek Lisans Tezi, (2003).

2. Yıldırım, N., Gökpınar, F., “Bazı Normallik Testlerinin 1. Tip Hataları ve Güçleri Bakımından Kıyaslanması”, SDÜ Fen Bilimleri Enstitü Dergisi, 16-1, 109-115 (2012).

Hobiler

Bilgisayar teknolojileri