KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

YÖNSEL YAKLAŞIM KULLANILARAK UYUM İYİLİĞİNİN OLABİLİRLİK İNDEKSİ

DOKTORA TEZİ

Özge TEZEL

AĞUSTOS 2022
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

YÖNSEL YAKLAŞIM KULLANILARAK UYUM İYİLİĞİNİN OLABİLİRLİK İNDEKSİ

Özge TEZEL
ORCID : - - -

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"DOKTOR (İSTATİSTİK)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 27 / 07 /2022
Tezin Savunma Tarihi : 23 / 08 /2022

Tez Danışmanı : Doç. Dr. Orhan KESEMEN
ORCID : - - -

Trabzon 2022

ÖNSÖZ

"Yönsel Yaklaşım Kullanılarak Uyum İyiliğinin Olabilirlik İndeksi" isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik ve Bilgisayar Bilimleri Anabilim Dalı, Doktora Programı'nda hazırlanmıştır.

Tez çalışmam boyunca hiçbir desteğini esirgemeyen ve bilim insanı olma yolundaki ilk adımlarımı atarken kıymetli bilgileriyle ve tecrübesiyle yolumu aydınlatan danışman hocam Sayın Doç. Dr. Orhan KESEMEN'e teşekkürü bir borç bilirim. Ayrıca, yapıcı eleştirileri ve önerileri ile tezime büyük katkıları bulunan saygıdeğer hocalarım Sayın Prof. Dr. Mualla YALÇINKAYA'ya ve Sayın Dr. Öğr. Üyesi Uğur ŞEVİK'e, eğitim öğretim hayatımda katkısı olan tüm hocalarıma ve ortak olarak çalışmalar yürüttüğümüz ve tez çalışması süresinde de hiçbir yardımdan kaçınmayan arkadaşlarım Arş. Gör. Dr. Eda ÖZKUL'a ve Arş. Gör. Dr. Buğra Kaan TİRYAKİ'ye teşekkür ederim.

Tez çalışması sürecinde bana desteklerini esirgemeyen İstatistik ve Bilgisayar Bilimleri Bölümü'ndeki tüm değerli hocalarıma ve asistan arkadaşlarıma teşekkürlerimi sunarım.

Son olarak, doğduğum günden beri elimi hiç bırakmayan, hayatım boyunca aldığım tüm kararlarda bilgi ve tecrübelerinden faydalandığım ve her kararımda arkamda olan, maddi ve manevi desteklerini hiçbir zaman esirgemeyen ve bugünlere gelmemde en büyük paya sahip olan annem ve babam Nuran-Hasan TEZEL'e, kardeşlerim Müge-Mehmet TEZEL'e teşekkür eder, minnettarlığımı sunarım.

Bu tez çalışmasının bundan sonraki çalışmalara katkı sağlamasını temenni ederim.

Özge TEZEL

Trabzon 2022

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduğum "Yönsel Yaklaşım Kullanılarak Uyum İyiliğinin Olabilirlik İndeksi" başlıklı bu çalışmayı baştan sona kadar danışmanım Doç. Dr. Orhan KESEMEN'in sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 23/08/2022

Özge TEZEL

İÇİNDEKİLER

Sayfa No

ÖNSÖZ III
TEZ ETİK BEYANNAMESİ IV
İÇİNDEKİLER V
ÖZET VII
SUMMARY VIII
ŞEKİLLER DİZİNİ IX
TABLOLAR DİZİNİ XI
SEMBOLLER DİZİNİ XIII
1. GENEL BİLGİLER 1
1.1. Gözlenen ve Beklenen Frekanslar Arasındaki Tutarsızlıklara Dayanan Uyum İyiliği Testleri 2
1.2. Deneysel Dağılım Fonksiyonuna Dayanan Uyum İyiliği Testleri 3
1.3. Momentlere Dayalı Uyum İyiliği Testleri 3
1.4. Regresyon ve Korelasyona Dayalı Uyum İyiliği Testleri 4
1.5. Entropi Kavramına Dayalı Uyum İyiliği Testleri 4
1.6. Çalışmada Kullanılacak Uyum İyiliği Testleri 8
1.6.1. Ki-Kare Uyum İyiliği Testi 8
1.6.2. Kolmogorov-Smirnov Uyum İyiliği Testi 9
1.6.3. Cramér-von Mises Uyum İyiliği Testi 11
1.6.4. Anderson-Darling Uyum İyiliği Testi 11
1.7. Çalışmada Kullanılacak Dağılımlar 12
1.7.1. Simetrik Dağılımlar 13
1.7.1.1. Normal Dağılım 13
1.7.1.2. Düzgün Dağılım 14
1.7.1.3. Laplace Dağılımı 15
1.7.1.4. Student-t Dağılım 16
1.7.2. Simetrik Olmayan Dağılımlar 17
1.7.2.1. Üstel Dağılım 18
1.7.2.2. Log-normal Dağılım 19
1.7.2.3. Weibull Dağılımı 20

1.7.2.4. Gamma Dağılımı 21
1.7.2.5. Ki-Kare Dağılımı 23
1.8. İstatistiksel Hipotez Testleri 24
2. YAPILAN ÇALIŞMALAR 25
2.1. Uyum İyiliğinin Olabilirlik İndeksi 25
2.2. Bağımsız Ki-Kare Uyum İyiliği Testi 28
2.2.1. Bileşke Vektörün Dağılımının Elde Edilmesi 29
2.2.2. Kritik Tablo Değerlerinin Elde Edilmesi 32
2.2.3. Bağımsız Ki-Kare Uyum İyiliği Test İstatistiğinin Hesaplanması 34
3. BULGULAR VE İRDELEME 38
3.1. Yönsel Olabilirlik İndeksinin Değerlendirilmesi 38
3.1.1. Merkezi Limit Teoremine Göre Olabilirlik İndeksinin Değerlendirilmesi 38
3.1.2. Büyük Sayılar Kanununa Göre Olabilirlik İndeksinin Değerlendirilmesi 44
3.1.3. Farklı Dağılımlara Göre Olabilirlik İndeksinin Hesaplanması 46
3.1.4. Deneysel Rastgele Sayıların Değiştirilmesi ile Olabilirlik İndeksinin Hesaplanması 48
3.2. Bağımsız Ki-Kare Testinin I. Tip Hata ve Güç Bakımından Karşılaştırılması 50
3.2.1. Simetrik Dağılımlarda I. Tip Hata ve Güç Değerlerinin Karşılaştırılması 50
3.2.2. Simetrik Olmayan Dağılımlarda I. Tip Hata ve Güç Değerlerinin Karşılaştırılması 59
3.3. Bağımsız Ki-Kare Testinin Gerçek Veri Seti Uygulamaları 67
3.3.1. Küresel Isınma Veri Seti 67
3.3.2. Katrina Kasırgası Veri Seti 70
3.3.3. Yağış Miktarı Veri Seti 73
3.3.4. Akciğer Kanseri Veri Seti 76
3.3.5. Uyku Süresi Veri Seti 79
3.4. Bağımsız Ki-Kare Testinin Değerlendirilmesi 81
3.4.1. Farklı Dağılımlara Göre Bağımsız Ki-Kare Testi 81
3.4.2. Deneysel Rastgele Sayıların Değiştirilmesi ile Bağımsız Ki-Kare Testinin Değerlendirilmesi 83
4. SONUÇLAR 85
5. ÖNERİLER 87
6. KAYNAKLAR 88
ÖZGEÇMİŞ

Doktora Tezi

ÖZET

YÖNSEL YAKLAŞIM KULLANILARAK UYUM İYİLİĞİNİN OLABİLİRLİK İNDEKSİ

Özge TEZEL

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik ve Bilgisayar Bilimleri Anabilim Dalı
Danışman: Doç. Dr. Orhan KESEMEN
2022, 96 Sayfa

Uyum iyiliği testlerinde yokluk hipotezinin kabul edilebilmesi için test istatistiğine ve test istatistiği kullanılarak hesaplanan p-değerine bakılmaktadır. Yani p-değeri birçok araştırmacı için kritik bir anlam taşımaktadır. Fakat p-değerinin küçük bir değişimi veya anlamlılık düzeyinin iyi seçilmemesi yokluk hipotezinin reddedilmesine veya kabul edilmesine sebep olmaktadır. Bu çalışmada öncelikle bu eksikliklerden kaçınmak için uyum iyiliği testlerinin güvenirlilik derecesi için ölçüt olarak kullanılabilecek bir olabilirlik indeksi geliştirilmesi amaçlanmıştır. Önerilen olabilirlik indeksi, yönsel yaklaşım temel alınarak geliştirilmiştir. Örneklemin test edilen dağılıma ne kadar iyi uyup uymadığının derecesini vermektedir. Ayrıca bu indeks değerinin dağılımı incelenerek bağımsız ki-kare uyum iyiliği testi önerilmiştir. Olabilirlik indeksinin ve bağımsız ki-kare uyum iyiliği testinin doğruluğunu, geçerliliğini ve performansını kapsamlı bir şekilde test etmek için farklı simülasyon çalışmaları yapılmıştır. Önerilen test yönteminin seçilen testlere göre birçok açıdan üstünlüğü ortaya konmuştur. Ayrıca bu test yöntemi, literatürdeki diğer testlere göre dağılımdan, serbestlik derecesinden ve sınıf sayısından bağımsız, hesaplaması ve kullanılması kolay bir yöntemdir. Son olarak, önerilen yöntem gerçek veri setlerine de uygulanmış ve başarımı gösterilmiştir.

Anahtar Kelimeler: Uyum iyiliği testi, p-değeri, Olabilirlik indeksi, Yönsel yaklaşım

PhD. Thesis

SUMMARY

A LIKELIHOOD INDEX OF GOODNESS OF FIT USING DIRECTIONAL APPROXIMATION

Özge TEZEL

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Statistics and Computer Sciences Graduate Program
Supervisor: Assoc. Prof. Orhan KESENEN
2022, 96 Pages

In goodness of fit tests, the test statistics and the p-value which is calculated by using test statistics are utilized for acceptance of the null hypothesis. That is, the p-value has a critical meaning for many researchers. However, a small change in the p-value or not well-chosen of significance level leads to the rejection or acceptance of the null hypothesis. In this study, it is aimed to develop a likelihood index that can be used as a criterion for the degree of reliability of the goodness-of-fit tests to avoid these deficiencies. The proposed likelihood index was developed based on the directional approximation and gives the degree to which the sample fits well to the distribution being tested. In addition, the distribution of this index value was examined, and a free chi-square goodness-of-fit test was proposed. Several simulation studies have been conducted to comprehensively test the accuracy, validity, and performance of the likelihood index and the free chi-square goodness-of-fit test. The proposed test method has demonstrated superiority in many aspects compared to the other selected test methods. Furthermore, the proposed test is distribution-free, degree of freedom-free, binning-free compared to other tests in the literature, it is easy to calculate and use. Finally, the proposed method has been applied to real data sets and its performance has been demonstrated.

Key Words: Goodness of fit tests, p-value, Likelihood index, Directional approximation

ŞEKİLLER DİZİNİ

Sayfa No

Şekil 1. Kolmogorov-Smirnov test istatistiğinin grafiksel gösterimi 10

Şekil 2. Normal dağılımın farklı parametreli olasılık yoğunluk fonksiyonları..... 14

Şekil 3. Düzgün dağılımın farklı parametreli olasılık yoğunluk fonksiyonları 15

Şekil 4. Laplace dağılımının farklı parametreli olasılık yoğunluk fonksiyonları 16

Şekil 5. Student-t dağılımının farklı parametreli olasılık yoğunluk fonksiyonları 17

Şekil 6. Üstel dağılımın farklı parametreli olasılık yoğunluk fonksiyonları 19

Şekil 7. Log-normal dağılımının farklı parametreli olasılık yoğunluk fonksiyonları..... 20

Şekil 8. Weibull dağılımının farklı parametreli olasılık yoğunluk fonksiyonları..... 21

Şekil 9. Gamma dağılımının farklı parametreli olasılık yoğunluk fonksiyonları 22

Şekil 10. Ki-kare dağılımının farklı parametreli olasılık yoğunluk fonksiyonları 23

Şekil 11. Önerilen olabilirlik indeks yönteminin grafiksel gösterimi (a) Oluşturulan rastgele veriler, (b) $0, 2\pi$ aralığına dönüştürülen veriler, (c) Dönüştürülen verilerin radyal sistemde gösterimi, (d) Verilerin bileşke kuvveti 27

Şekil 12. V_i değişkeninin ortalama yatay ve düşey bileşeninin histogramı (a) V_i değişkeninin ortalama yatay bileşeninin histogramı, (b) V_i değişkeninin ortalama düşey bileşeninin histogramı 29

Şekil 13. Bileşke vektörün karesinin histogramı 30

Şekil 14. Ki-kare uyum iyiliği testinin kabul ve kritik bölgesi..... 32

Şekil 15. R^2 ve χ^2 tablo değerlerinin grafiksel olarak karşılaştırılması..... 33

Şekil 16. C ve S değişkenlerinin Kolmogorov-Smirnov testi sonucundaki I. tip hataları 35

Şekil 17. Irwin-Hall dağılımının olasılık yoğunluk fonksiyonu 40

Şekil 18. Verilere bağlı olarak p ve Π değerlerinin değişimleri (a) Normal dağılıma göre elde edilen p ve Π değerleri, (b) Düzgün dağılıma göre elde edilen p ve Π değerleri 49

Şekil 19. Normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen I. tip hataları 51

Şekil 20. Düzgün dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri..... 53

Şekil 21. Laplace dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri..... 55

Şekil 22. Student-t dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri..... 57

Şekil 23. Üstel dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri 58
Şekil 24. Log-normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen I. tip hataları 60
Şekil 25. Normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri 62
Şekil 26. Weibull dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri 63
Şekil 27. Gamma dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri 65
Şekil 28. Ki-kare dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri 67
Şekil 29. CO^2 miktarı verilerinin olasılık yoğunluk fonksiyonu 70
Şekil 30. Basınç rüzgâr hızı verilerinin olasılık yoğunluk fonksiyonu 73
Şekil 31. Yağış miktarı verilerinin olasılık yoğunluk fonksiyonu 76
Şekil 32. Kötü huylu tümör verilerinin olasılık yoğunluk fonksiyonu 78
Şekil 33. Ortalama uyku süresi verilerinin olasılık yoğunluk fonksiyonu 81
Şekil 34. Normal ve düzgün dağılıma göre KS uygulanarak elde edilen p-değerlerinin değişimi 83
Şekil 35. Normal ve düzgün dağılıma göre FCS uygulanarak elde edilen p-değerlerinin değişimi 84

TABLOLAR DİZİNİ

Sayfa No

Tablo 1. İstatistiksel hata türleri ve olasılıkları 24

Tablo 2. Farklı α değerlerine göre R^2 ve χ^2 tablo değerleri 33

Tablo 3. Farklı α değerlerine göre bağımsız ki-kare testinin kritik tablo değerleri 36

Tablo 4. Bağımsız ki-kare testinin uygulaması 37

Tablo 5. Düzgün dağılımdan elde edilen örneklemin KS testi sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi 41

Tablo 6. Düzgün dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri, Π ve p-değerleri arasındaki korelasyon katsayısı 42

Tablo 7. Üstel dağılımdan elde edilen örneklemin KS testin sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi 43

Tablo 8. Üstel dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri, Π ve p-değeri arasındaki korelasyon katsayısı 44

Tablo 9. Normal dağılımdan elde edilen örneklemin KS testi sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi 45

Tablo 10. Normal dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri ve Π ve p-değerleri arasındaki korelasyon katsayısı 46

Tablo 11. Dağılımların yönsel olabilirlik indeks değerleri Π 47

Tablo 12. KS uyum iyiliği testi sonucunda elde edilen p-değerleri 47

Tablo 13. KS testi sonucundaki p-değerlerine göre yokluk hipotezlerinin kabul yüzdeleri 48

Tablo 14. Normal dağılım $\mu = 0, \sigma^2 = 1$ için uyum iyiliği testlerinin I. tip hataları 51

Tablo 15. Düzgün dağılım için uyum iyiliği testlerinin güç karşılaştırmaları 52

Tablo 16. Laplace dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları 54

Tablo 17. Student-t dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları 56

Tablo 18. Üstel dağılım için uyum iyiliği testlerinin güç karşılaştırmaları 58

Tablo 19. Log-normal dağılım için uyum iyiliği testlerinin I. tip hataları 59

Tablo 20. Normal dağılım için uyum iyiliği testlerinin güç karşılaştırmaları 61

Tablo 21. Weibull dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları 63

Tablo 22. Gamma dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları 64

Tablo 23. Ki-kare dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları 66

Tablo 24. Küresel ısınma veri seti 68

Tablo 25. Küresel ısınma veri setinin uyum iyiliği testi sonuçları 69

Tablo 26. Katrina kasırgasına ait veri seti 71
Tablo 27. Katrina kasırgası veri setinin uyum iyiliği testi sonuçları 72
Tablo 28. Yağış miktarı veri seti 74
Tablo 29. Yağış miktarı veri setinin uyum iyiliği testi sonuçları 75
Tablo 30. Akciğer kanseri veri seti 77
Tablo 31. Akciğer kanser veri setinin uyum iyiliği testi sonuçları 77
Tablo 32. Uyku süresi veri seti 79
Tablo 33. Ortalama uyku süresi veri setinin uyum iyiliği testi sonuçları 80
Tablo 34. KS testi sonucundaki yokluk hipotezinin kabul yüzdeleri 82
Tablo 35. FCS testi sonucundaki yokluk hipotezinin kabul yüzdeleri 82

SEMBOLLER DİZİNİ

AD : Anderson-Darling testi

BCS : Ki-kare testi

CVM : Cramer-von Mises testi

FCS : Önerilen bağımsız ki-kare testi

KS : Kolmogorov-Smirnov testi

R^2 : Bağımsız ki-kare test istatistiği

Π : Olabilirlik indeks değeri

α : I. Tip hata (anlamlılık düzeyi)

$1-\alpha$: Güven düzeyi

β : II. Tip hata

$1-\beta$: Testin gücü

n : Örneklem büyüklüğü

$F(x)$: Birikimli dağılım fonksiyonu

$F_0(x)$: Hipotezdeki birikimli dağılım fonksiyonu

1. GENEL BİLGİLER

İstatistikte yapılan bir araştırma sonucunda elde edilen gözlem değerlerinin dağılımının bilinmesi, analiz aşamasında uygulanacak yöntemin belirlenmesi açısından oldukça önemlidir. Dağılımın gerçeği yansıtmadığı durumlarda yanlış istatistiksel analiz yöntemleri kullanılabilecek bu durumda araştırmacı, hatalı sonuçlar elde etmesinin yanı sıra bilgi ve zaman kaybı gibi ciddi sorunlarla karşılaşacaktır. Bazı istatistiksel çıkarım yöntemleri farklı dağılım varsayımlarına özellikle normal dağılım varsayımına dayanmaktadır. Bunun temel nedeni, parametrik istatistiksel tekniklerin çoğunun normal dağılıma dayalı olarak geliştirilmesidir. Örneğin; regresyon ve varyans analizinde, hataların normal dağılıma sahip olduğu varsayılır. Bu sebeple, normallik varsayımının geçerliliğini test etmek, bulguları doğrulamak için gereklidir (Wijekularathna vd., 2019). Örneklemin bu varsayımları sağlayıp sağlamadığı grafiksel olarak veya uyum iyiliği testleri kullanılarak kontrol edilebilir. Q-Q grafikleri, P-P grafikleri ve histogramlar verinin normal dağılıma uygunluğu hakkında bilgi veren grafiksel yöntemlerdir. Ancak bir grafik veya histogram incelenerek normallik varsayımının sağlanıp sağlanmadığına karar vermek öznel bir değerlendirme olacaktır ve dolayısıyla grafiksel incelemeler zaman zaman yanıltıcı sonuçlar doğurabilmektedir. Bu nedenle, normallik varsayımının incelenmesinde uyum iyiliği testlerinin kullanılması daha doğru olacaktır. Hem yapay verilerin hem de gerçek hayattan toplanan gözlem değerlerinin verilen bir dağılımdan gelip gelmediğini araştırmak için kullanılan tüm testlere uyum iyiliği testleri denir. Rastgele sayıların veya gözlem değerlerinin uyum iyiliği testleri, veri analizinde ve çıkarımsal istatistikte en temel problemlerden biridir (Saksena, 1980; D'Agostino ve Stephens, 1986; Wang, 2012). Uyum iyiliğinde yokluk hipotezi ve alternatif hipotez;

$$H_0: F(x) = F_0(x)$$
$$H_1: F(x) \neq F_0(x)$$

biçiminde verilir. Uyum iyiliği testleri, ilk olarak Karl Pearson tarafından 1900 yılında ortaya atılan, günümüzde de popülaritesini kaybetmeyen konulardan biridir. Uyum iyiliği testleriyle ilgili literatürde farklı test istatistikleri önerilmiştir. Bu çalışmalardan bazıları Pearson (1900), Cramer-von Mises (1928), Kolmogorov-Smirnov (1933; 1939), Anderson-

Darling (1952; 1954; 1962), Cox (1961; 1962), Watson (1961; 1962), Kuiper (1960), Lilliefors (1967), Shapiro-Wilk (1965; 1968), Hegazy ve Green (1975), Vasicek (1976), Pettitt (1977), Jarque-Bera (1980; 1981), Epps (1982; 1983; 2005), Baringhaus ve Henze (1988), Ledwina (1994), Fan (1996; 1997; 1998), del Barrio ve diğerleri (1999), Zhang (1999), Owen (2001), Song (2002), Zhang (2002), Fortiana ve Grané (2003), Matsui ve Tamura (2005), Zhang ve Wu (2005), Esteban ve diğerleri (2001; 2007), Dong ve Giles (2007), Towhidi ve Salmanpour (2007), Brys ve diğerleri (2008), Romao vd. (2010), Shan vd. (2010), Grané (2012), Torabi ve diğerleri (2016) tarafından yapılmıştır. Literatürdeki uyum iyiliği testleri farklı yaklaşımlardan yararlanılarak geliştirilmiştir ve çeşitli sınıflandırmalar yapılmıştır. Bu sınıflandırmalar göz önünde bulundurularak uyum iyiliği testleri,

- Gözlenen ve beklenen frekanslar arasındaki tutarsızlıklara dayanan uyum iyiliği testleri,
- Deneysel dağılım fonksiyonuna dayanan uyum iyiliği testleri,
- Momentlere dayalı uyum iyiliği testleri,
- Regresyon ve korelasyona dayalı uyum iyiliği testleri,
- Entropi kavramına dayanan uyum iyiliği testleri,

biçiminde beş ana kategoriye ayrılabilir.

1.1. Gözlenen ve Beklenen Frekanslar Arasındaki Tutarsızlıklara Dayanan Uyum İyiliği Testleri

Olabilirlik oranı ve ki-kare test istatistikleri, gözlenen ve beklenen frekanslar arasındaki tutarsızlıklara dayanan uyum iyiliği testleri kategorisinde yer almaktadır. Literatürde gözlenen ve beklenen frekansları temel alan birçok uyum iyiliği testi geliştirilmiştir. Bu kategoride yer alan testlerden ki-kare test istatistikleri yaygın bir şekilde kullanılmaktadır. Pearson 1900'deki makalesinde, ki-kare test istatistiğini önermiştir (Pearson, 1900). Pearson makalesinde önerdiği test istatistiğinin pratikte kullanımını göstermek için zar atma, rulet ve atış gibi gerçek yaşam örneklerinden yararlanmıştır. Pearson ki-kare testi, Cochran tarafından modern istatistiğin temellerinden biri olarak görülmüştür (Cochran, 1952). Neyman (1949) tarafından değiştirilmiş ki-kare testi önerilmiştir ve bu test yöntemi de Pearson ki-kare testi gibi küçük gözlenen frekans

sayısından etkilenmektedir. Freeman ve Tukey (1950) tarafından önerilen test istatistiği gözlenen ve beklenen frekansların kareköklerinin farklılıklarına dayanmaktadır. Wilks (1935) frekansların oranına dayanan logaritmik olabilirlik oran test istatistiğini önermiştir. Literatürde ki-kare ve olabilirlik oran testlerinin birçok modifikasyonu bulunmaktadır (Balakrishnan vd., 2013).

1.2. Deneysel Dağılım Fonksiyonuna Dayanan Uyum İyiliği Testleri

Literatürde, ki-kare uyum iyiliği testine alternatif olarak deneysel dağılım fonksiyonuna dayalı ve daha duyarlı olan birçok uyum iyiliği testi geliştirilmiştir. Bu testlerdeki temel ilke, örneklemin birikimli dağılım fonksiyonu ile yokluk hipotezi altında deneysel dağılım fonksiyonunu karşılaştırarak, aralarında iyi bir uyum olup olmadığını kontrol etmektir. Deneysel dağılım fonksiyonuna dayanan uyum iyiliği testleri bireysel gözlem değerlerini doğrudan kullandıkları için ki-kare uyum iyiliği testine göre daha tutarlı sonuçlar vermektedir (Bain ve Engelhardt, 1987). Deneysel dağılım fonksiyonuna dayanan uyum iyiliği testleri, yokluk hipotezi altında deneysel dağılım fonksiyonu ve birimli dağılım fonksiyonu arasındaki farkı ölçmeye yöneliktir. Cramér-Von Mises (1928), Kolmogorov-Smirnov (1933), Anderson-Darling (1952), Lilliefors (1967), Hegazy Green (1975), Frosini (1987), Glen-Leemis-Barr (2001), Zhang-Wu (2005) uyum iyiliği testleri bu kategoride yer alan testlerdendir. Bunlardan Kolmogorov-Smirnov, Cramér-Von Mises ve Anderson-Darling testleri yaygın bir şekilde kullanılmaktadır (Arshad vd., 2003).

1.3. Momentlere Dayalı Uyum İyiliği Testleri

İstatistiksel momentler, merkezi eğilim ve dağılım ölçülerinin yanı sıra bir dağılımın özellikleri hakkında daha fazla bilgi vermektedir. Momentler bir dağılımı ölçmek için kullanılan bir dizi istatistiksel parametrelerdir. Birinci moment ortalamayı, ikinci merkezi moment varyansı, üçüncü ve dördüncü merkezi momentler ise sırasıyla çarpıklık ve basıklığı vermektedir. İki dağılım aynı ortalama ve standart sapmaya sahip olabilir; ancak dağılım şekilleri çarpıklık ve basıklıkla açıklanmaktadır. Literatürde momentlere dayalı çarpıklık ve basıklığı temel alan birçok uyum iyiliği testi geliştirilmiştir. D'Agostino–Pearson K^2 (1973), Jarque-Bera (1980), Hosking (1990), Bonett-Seier (2002), Cabaña-Cabaña (2003),

Bontemps-Meddahi (2005), Brys–Hubert–Struyf MC–LR (2008), Doornik–Hansen (2008), Gel–Gastwirth (2008) testleri momentlere dayalı uyum iyiliği testlerindendir.

1.4. Regresyon ve Korelasyona Dayalı Uyum İyiliği Testleri

Regresyon ve korelasyona dayalı uyum iyiliği testleri, sıra istatistiklerinden elde edilen ağırlıklı en küçük kareler tahmininin oranına dayanmaktadır (Uyanto, 2022). Shapiro-Wilk testi (1965), D'Agostino D testi (1971), Shapiro-Francia testi (1972), Filliben korelasyon testi (1975), Chen–Shapiro testi (1995), Rahman-Govindarajulu testi (1997), Zhang-Q testleri (1999), Barrio-Cuesta-Matran-Rodriguez testi (1999), Coin testi (2008) regresyon ve korelasyona dayalı uyum iyiliği testlerindendir.

1.5. Entropi Kavramına Dayalı Uyum İyiliği Testleri

Rastgele değişkenin entropisi, bilgi ve belirsizliğinin bir ölçüsü olarak tanımlanmaktadır ve ilk olarak C. E. Shannon (1948) tarafından ortaya atılmıştır. Ayrıca istatistiksel bir deney gerçekleştirilmeden önceki belirsizliğin ölçüsü olarak da ifade edilmektedir (Khinchin, 1957). Bir dağılımın bir diğerine uygunluğunun belirlenmesinde de entropiye dayalı ölçümler kullanılmaktadır (Evren ve Tuna, 2012). Shannon entropisi bazı uyum iyiliği testlerinin temelini oluşturmaktadır. Gokhale (1983) tarafından maksimum entropi dağılımlarının aileleri (normal, üstel, laplace, gamma ve beta dağılımı) için uyum iyiliği testlerinin genel yapısı elde edilmiştir. Ayrıca üstel ve çift üstel dağılımları için simülasyonlar yapılarak önerilen yöntem diğer uyum iyiliği testleriyle karşılaştırılmıştır. Vasicek (1976) ve Arizono ve Ohta (1989) tarafından entropiye dayalı normallik testleri önerilirken, Dudewicz vd. (1995) örneklem entropisine dayalı düzgün dağılım için bir test geliştirmişlerdir. Ayrıca bu test rastgele sayı üreteçlerinin değerlendirilmesinde kullanılmıştır. Crzcgorzewski ve Wirczorkowski (1999), Shannon entropisine dayalı üstel dağılım için uyum iyiliği testleri ve Park (1999) sıra istatistiklerinin örneklem entropisi üzerine çalışmalar yapmışlardır. Entropiye dayalı uyum ölçülerinden en yaygın olarak kullanılanı Kullback-Leibler (1951) sapmasıdır. Arizono ve Ohta (1989) tarafından yapılan çalışmada da Kullback-Leibler bilgisinden faydalanılarak geniş kapsamlı bir uyum iyiliği testi önerilmiştir.

Literatürde, önerilen uyum iyiliği test istatistiklerine ilaveten farklı dağılımlar ve varsayımlar göz önüne alınarak uyum iyiliği testlerinin güç değerlerini karşılaştıran birçok çalışma bulunmaktadır. Yapılan karşılaştırma çalışmaları, farklı dağılımlar ve farklı varsayımlara göre hangi uyum iyiliği test yönteminin tercih edilebileceğini ortaya koymak için gerçekleştirilmiştir. Stephens (1974) yaptığı çalışmada, çeşitli örnek genişliklerinde Kolmogorov-Smirnov, Anderson-Darling, Watson, Cramér-von Mises, Kuiper ve Shapiro-Wilk testlerinin güç performanslarını karşılaştırmıştır. Bu karşılaştırma sonucunda, deneysel dağılım fonksiyonuna bağlı uyum iyiliği testlerinin Shapiro-Wilk testine karşı zayıf performans gösterdiği elde edilmiştir. Ancak Anderson-Darling ve Cramér-von Mises testlerinin deneysel dağılım fonksiyonuna bağlı testler içerisinde en iyi performansı gösterdiği belirtilmiştir. Seier (2002) yaptığı çalışmada, çeşitli örneklem büyüklükleri kullanılarak normallik testlerinin performansını simülasyonlarla karşılaştırmıştır. Bu karşılaştırmalar sonucunda, Anderson-Darling testinin Kolmogorov-Smirnov testine göre daha güçlü bir test istatistiği olduğu sonucuna varılmıştır. Bunun yanı sıra, normallik testlerinin yalnızca güç açısından değil aynı zamanda basitlik, hesaplama zamanı ve kritik değere duyulan ihtiyaç bakımından da farklılıklar gösterdiği belirtilmiştir. Mendes ve Pala (2003), Shapiro-Wilk, Lilliefors ve Kolmogorov-Smirnov testlerinin I. tip hata ve güçlerini karşılaştırmışlardır. Bu karşılaştırma sonucunda, Shapiro-Wilk testinin en güçlü test olduğu, en zayıf testin ise Kolmogorov-Smirnov testi olduğu belirtilmiştir. Yazıcı ve Yolacan (2007) tarafından yapılan çalışmada, normallik testlerinin güçleri çeşitli örneklem büyüklükleri ve çeşitli dağılımlar kullanılarak karşılaştırılmıştır. Bu çalışma sonucunda Kuiper, Vasicek ve Jarque-Bera testlerinin diğer uyum iyiliği testlerine göre daha az güçlü olduğu elde edilmiştir ve küçük örneklem büyüklüğüne sahip simetrik dağılımlar için Kolmogorov-Smirnov ve Anderson-Darling testlerinin kullanılabilir olduğu belirtilmiştir. Chaichatschwal ve Budsaba (2007), olabilirlik oranına bağlı olan Z_A, Z_C, Z_K uyum iyiliği testleriyle Anderson-Darling, Shapiro-Wilk, Shapiro-Francia testlerini karşılaştırmıştır. Simülasyonda kullanılan dağılımlar normale yakın, simetrik uzun kuyruklu, simetrik kısa kuyruklu, asimetrik uzun kuyruklu ve asimetrik kısa kuyruklu dağlımlar olarak sınıflandırılmıştır. Çalışma sonucunda, Z_A ve Z_C istatistiklerinin güç değerlerinin Z_K ve Anderson-Darling test istatistiklerine göre daha yüksek olduğu ortaya koyulmuştur. Nakas (2007) yaptığı çalışmada, Reschenhofer ve Bomze (1991) tarafından önerilen P-P grafiğinin uzunluğuna dayanan uyum iyiliği testi ile Cramér-von Mises ve Kolmogorov-Smirnov testlerini karşılaştırmıştır. Karşılaştırma sonucunda, uzunluk testinin diğer testlere göre

yüksek güce sahip olduğu ve pratikte tercih edilebilecek bir test yöntemi olabileceği gösterilmiştir. Romao, Delgado ve Costa (2010), farklı örneklem büyüklükleri için 33 normallik testinin performansını çeşitli anlamlılık seviyelerini göz önüne alarak karşılaştırmışlardır. Bu çalışmada, dağılımın simetrik, asimetrik olması ve dağılımın yapısı hakkında bilgi olmaması durumunda kullanılabilecek en iyi uyum iyiliği testleri belirtilmiştir. Noughabi ve Arghami (2011), Kolmogorov-Smirnov, Anderson-Darling, Kuiper, Jarque-Bera, Cramér-von Mises, Shapiro-Wilk ve Vasicek testlerini farklı dağılımlar altında testin gücü bakımından karşılaştırmışlardır. Farklı sınırlar içerisinde tanımlı olan çeşitli dağılımlar için hangi uyum iyiliği testlerinin daha güçlü olabileceği belirlenmiştir. Bu çalışma sonucunda, $(-\infty, \infty)$ aralığında tanımlı, simetrik dağılımlar için Jarque-Bera testinin, asimetrik dağılımlar için Shapiro-Wilk testinin daha yüksek güce sahip olduğu gösterilmiştir. Ayrıca $(0, 1)$ aralığında tanımlı dağılımlar için Vasicek testi, $(0, \infty)$ aralığında tanımlı dağılımlar için de Vasicek ve Shapiro-Wilk testleri daha yüksek güç değerleri göstermişlerdir. Yap ve Sim (2011) tarafından yapılan çalışmada, simetrik kısa kuyruklu, simetrik uzun kuyruklu ve asimetrik dağılımlar kullanılarak farklı uyum iyiliği testleri için güç karşılaştırmaları yapılmıştır. Simülasyonlar sonucunda, simetrik kısa kuyruklu dağılımlar için D'Agostino ve Shapiro-Wilk testlerinin kullanılabileceği, simetrik uzun kuyruklu dağılımlar için Jarque-Bera, D'Agostino testlerinin güçlerinin Shapiro-Wilk testi ile oldukça benzer olduğu ve asimetrik dağılımlar için de Shapiro-Wilk ve Anderson-Darling testinin en güçlü testler olduğu belirtilmiştir. Razali ve Yap (2011) tarafından Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors ve Anderson-Darling testleri farklı dağılımlar altında testin gücü bakımından karşılaştırılmıştır. Sonuç olarak, Shapiro-Wilk testinin tüm dağılım türleri ve örneklem büyüklükleri için en güçlü test olduğu, Kolmogorov-Smirnov testinin ise karşılaştırılan testler içinde en az güce sahip olduğu sonucuna varılmıştır. Van Zyl (2018) tarafından deneysel karakteristik fonksiyona dayalı yeni bir test istatistiği önerilmiş ve bu test yöntemi Lilliefors, Jarque-Bera, Shapiro-Wilk, Anderson-Darling ve D'Agostino-Pearson gibi normallik testleriyle karşılaştırılmıştır. Simülasyon çalışması sonucunda, önerilen test yönteminin büyük örneklemlerde diğer test yöntemlerine göre daha iyi bir performansa sahip olduğu görülmüştür. Wijekularathna vd. (2019), Monte-Carlo simülasyonunu kullanarak normallik için geliştirilen 12 test yöntemini karşılaştırmışlardır. Karşılaştırma aşamasında kullanılan dağılımlar simetrik kısa kuyruklu, simetrik uzun kuyruklu ve asimetrik olarak kategorize edilmiştir. Simülasyonlar sonucunda, farklı örneklem büyüklükleri ve dağılımlar ele alındığında her durum için farklı test

yönteminin başarılı sonuçlar verdiği görülmüştür. Siraj-Ud-Doulah (2019), farklı örneklem büyüklükleri altında simetrik kısa ve uzun kuyruklu, asimetrik kısa ve uzun kuyruklu dağılımlar kullanarak, Monte Carlo simülasyonu yoluyla 27 normallik testinin gücünü incelemiş ve karşılaştırmıştır. Shapiro-Wilk testinin hemen hemen tüm durumlarda iyi performans gösterdiği ve daha iyi sonuçlar verdiği görülmüştür. Domanski ve Szczepocki (2020), moment ölçümlerine dayanan 9 normallik testinin güç değerlerini Monte Carlo simülasyonundan faydalanarak karşılaştırmışlardır. Bu testlerin farklı örneklem büyüklüğünde, farklı anlamlılık düzeylerinde ve alternatif dağılımlar üzerindeki etkinliği araştırılmıştır. Simülasyonlar sonucunda, incelenen testlerden tüm senaryolarda en yüksek güç değerine sahip olan bir test yöntemine rastlanmamıştır. Bayond (2021), deneysel ve teorik dağılım eğrileri altındaki ortak alana dayalı normallik için yeni bir uyum iyiliği testi önermiştir ve önerilen test yöntemi farklı test yöntemleriyle çeşitli senaryolar göz önüne alınarak karşılaştırılmıştır. Simülasyon çalışması sonucunda, önerilen test yönteminin tutarlı olduğu ve çeşitli senaryolarda iyi performans gösterdiği sonucuna varılmıştır. Ancak örneklem, $(-\infty, \infty)$ aralığında ve simetrik olduğunda robust Jargue-Bera yönteminin, $(0, \infty)$ veya $(-\infty, \infty)$ aralığında asimetrik olduğunda büyük örneklemlerde Shapiro-Wilk yönteminin, küçük örneklemlerde önerilen yöntemin, $(0, 1)$ aralığında ise entropiye dayalı Vasicek ve Park-Park yönteminin en güçlü olduğu görülmüştür. Uhm ve Yi (2021) tarafından yapılan çalışmada bazı normallik testleri ele alınarak alternatif hipotezler altında değişen örneklem büyüklüğü, anlamlılık düzeyi ve dağılımlar kullanılarak güç karşılaştırması yapılmıştır. Ayrıca test istatistiklerinin p-değerlerinin beklenen değeri ve medyanı kullanılarak da performansları karşılaştırılmıştır. Batsidis, Economou ve Bar-Lev (2022), Laplace dağılımı için yeni bir uyum iyiliği testi geliştirerek, bu testin karşılaştırma çalışmalarını yapmışlardır. Simülasyonlarda 27 farklı dağılım ve 22 farklı uyum iyiliği testi kullanılmıştır. Karşılaştırmalar sonucunda, her durum için diğerlerinden daha iyi performans gösteren bir test yöntemi bulunmamıştır. Uyanto (2022), literatürde yer alan 50 normallik testi için farklı örneklem büyüklüklerinde simetrik ve simetrik olmayan dağılımları kullanarak performans karşılaştırması yapmıştır. $(-\infty, \infty)$ aralığında tanımlanan, simetrik dağılımlar için robust Jarque-Bera ve Gel-Miao-Gastwirth testlerinin asimetrik dağılımlar için 1. Cabana-Cabana ve 2. Zhang-Wu testlerinin en güçlü testler olduğu gösterilmiştir.

1.6. Çalışmada Kullanılacak Uyum İyiliği Testleri

Literatürde farklı kategorilerde değerlendirilen birçok uyum iyiliği test yöntemi olmasına rağmen bu tezde simülasyon çalışmalarında, gözlenen ve beklenen frekanslar arasındaki tutarsızlıkları temel alan ki-kare testi ve deneysel dağılım fonksiyonuna dayanan ve yaygın bir şekilde kullanılan Kolmogorov-Smirnov, Cramér-von Mises ve Anderson-Darling uyum iyiliği testlerinden faydalanılmıştır. Simülasyon çalışmalarında kullanılan bu test yöntemlerinden ayrıntılı bir şekilde aşağıda bahsedilmiştir.

1.6.1. Ki-Kare Uyum İyiliği Testi

Ki-kare uyum iyiliği testi, gözlenen verilerin seçilen bir dağılıma uyup uymadığını test etmek için uyum iyiliği testleri içerisinde en yaygın kullanılan yöntemlerden biridir. Ki-kare testinin arkasındaki temel fikir veriyi sınıflara bölmektir. Bu sınıflardaki gözlenen ve beklenen frekanslar arasındaki tutarsızlıklar ki-kare test istatistiğini ortaya koymaktadır. Dolayısıyla ki-kare testi, gözlenen frekanslar ile beklenen frekanslar arasındaki farkın istatistiksel olarak anlamlı olup olmadığı temeline dayanır. Eğer veriler test edilen dağılımdan geliyorsa, her bir sınıfa düşen gözlem sayısı o sınıf için beklenen gözlem sayısına yakın olacaktır.

$\mathbf{x} = \{x_{j,}, (j = 1,2, \dots, n)\}$ bağımsız özdeş dağılımdan bir örneklem olsun. Ki-kare uyum iyiliği testinde veri aralığı k tane sınıfa bölünürse test istatistiği,

$$\chi_h^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

biçiminde hesaplanır. Burada, O_i, i-nci sınıfın gözlenen frekansını; E_i ise i-nci sınıfın beklenen frekansını temsil etmektedir. Beklenen frekans,

$$E_i = n \cdot \big(F_0(U_i) - F_0(L_i)\big), \qquad (i = 1,2, \dots, k) \tag{2}$$

biçiminde hesaplanırken, burada, F_0 test edilen dağılımın birikimli dağılım fonksiyonu; L_i, i-nci sınıfın alt sınırını; U_i, i-nci sınıfın üst sınırını göstermektedir. Gözlenen frekans ise,

$$O_i = \sum_{x_j \in [L_i, U_i]} 1, \quad (j = 1,2,\dots,n; \;\; i = 1,2,\dots,k) \tag{3}$$

şeklinde hesaplanmaktadır. Test istatistiği $(k-1)$ serbestlik dereceli ki-kare dağılımına sahiptir ve α anlamlılık düzeyinde tablo değeri $\chi^2_{k-1,a}$ olmak üzere, $\chi^2_h > \chi^2_{k-1,a}$ ise H_0 hipotezi reddedilir.

Ki-kare uyum iyiliği testi sınıf sayısına duyarlıdır dolayısıyla ki-kare testinin uygulanabilmesi için her sınıftaki beklenen frekansların 5'ten büyük olması istenir. Beklenen frekansların 5'ten büyük olması için ya örneklemin büyük hacimli olması gerekir ya da komşu sınıflar birleştirilerek beklenen frekansların 5'ten büyük olması sağlanır. Ki-kare testindeki en büyük zorluk optimal sınıf sayısının belirlenmesidir. Bu testin bir diğer dezavantajı ise sınıflandırmadan kaynaklı olarak bilgi kaybının meydana gelmesidir (Zghoul, 2010). Ancak ki-kare testi ayrık veya sürekli, tek değişkenli veya çok değişkenli veriler için kullanılabilir. Bu nedenle, uyum iyiliği testlerinden en çok kullanılanlardan biridir (D'Agostino ve Stephens, 1986).

1.6.2. Kolmogorov-Smirnov Uyum İyiliği Testi

Kolmogorov-Smirnov testi, ilk olarak Kolmogorov (1933) tarafından tek örneklem için uyum iyiliği testi olarak geliştirilmiştir (von Plato, 2005). 1939 yılında ise Smirnov, iki bağımsız örneklem için uyum iyiliği testini önermiştir (Smirnov, 1939). Bu sebeple literatürde Kolmogorov-Smirnov uyum iyiliği testleri olarak geçmektedir. Kolmogorov-Smirnov testi iki kümülatif dağılım fonksiyonunun incelenmesini temel alır. Birincisi yokluk hipotezinde verilen $F_0(x)$, uygunluğu istenen teorik birikimli dağılım fonksiyonudur. İkincisi ise, örneklemden hesaplanan $F_n(x)$, deneysel dağılım fonksiyonudur. Kolmogorov-Smirnov testi bu iki dağılım fonksiyonu arasındaki maksimum farka dayanmaktadır (Massey Jr, 1951). $x = \{x_1, x_2, \dots, x_n\}$ bağımsız özdeş dağılımdan oluşan bir örneklem olsun. Yokluk hipotezi $H_0: F(x) = F_0(x)$ ve alternatif hipotez de $H_1: F(x) \neq F_0(x)$ biçiminde verilsin bu durumda Kolmogorov-Smirnov test istatistiği (D_n),

$$D_n = \sup_x |F_n(x) - F_0(x)| \tag{4}$$

şeklinde hesaplanmaktadır. Burada, örneklemden hesaplanan deneysel dağılım fonksiyonu olan $F_n(x)$, x değerinden küçük veya ona eşit olan örnek gözlemlerin oranıdır $(x_i \leq x)$ ve

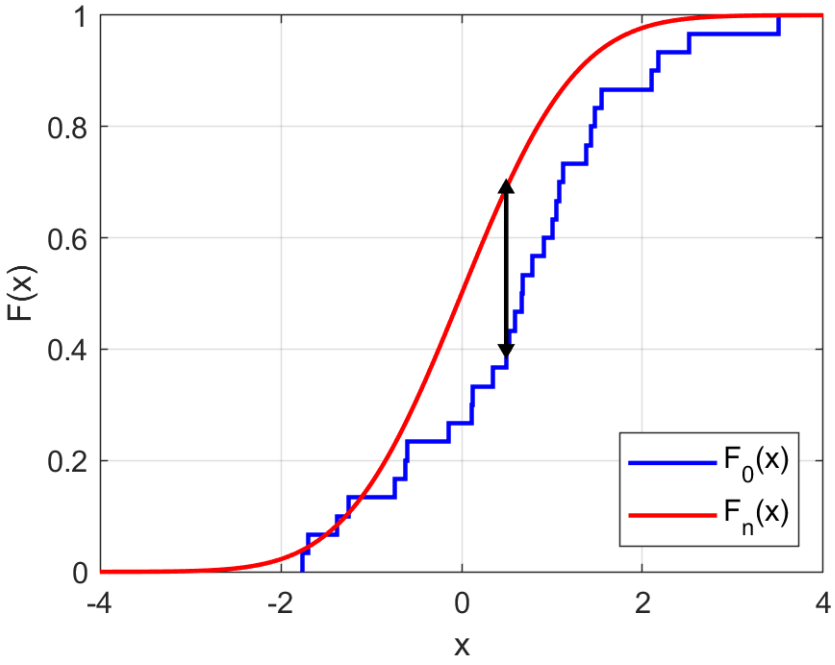
$$F_n(x_i) = \frac{1}{n}\sum_{j=1}^{n} I(x_j, x_i), \qquad (i = 1, 2, \dots, n) \tag{5}$$

biçiminde hesaplanmaktadır. Burada, $I(x)$ gösterge fonksiyonudur ve

$$I(a, b) = \begin{cases} 1, & a \leq b \\ 0, & diğer\ durumlarda \end{cases} \tag{6}$$

şeklinde ifade edilmektedir. Yani, deneysel dağılım fonksiyonu her bir gözlemde yüksekliği $1/n$ artan bir basamak fonksiyonudur.

Kolmogorov-Smirnov uyum iyiliğine ilişkin tablodan n ve $1-\alpha$ değerlerine göre bulunan $D_{n,1-\alpha}$ değeri ve örneklemden hesaplanan değer D_n olmak üzere $D_n > D_{n,1-\alpha}$ ise H_0 hipotezi reddedilir. Ayrıca tek örnek Kolmogorov-Smirnov uyum iyiliği test istatistiğinin değeri grafiksel olarak Şekil 1'deki gibi gösterilir.



Şekil 1. Kolmogorov-Smirnov test istatistiğinin grafiksel gösterimi

Kolmogorov-Smirnov testinde, öncelikle $F_0(x)$ ve $F_n(x)$ fonksiyonlarının grafikleri çizilir ve test istatistiği bu iki eğri arasındaki en büyük dikey uzaklık ile hesaplanır.

1.6.3. Cramér-von Mises Uyum İyiliği Testi

Cramér-von Mises uyum iyiliği testi, Harold Cramér ve Richard Edler von Mises tarafından 1928-1930 yıllarında ortaya atılmıştır (Cramér, 1928). Bu test yöntemi, Kolmogorov-Smirnov testi gibi birikimli dağılım fonksiyonu ile deneysel dağılım fonksiyonunun karşılaştırılmasına dayanmaktadır. Cramér-von Mises testinde bu iki dağılım fonksiyonu arasındaki uyum, karesel farkın integrali kullanılarak,

$$Q^2 = n\int_{-\infty}^{\infty}[F_n(x) - F_0(x)]^2\psi(x)dF(x) \tag{7}$$

biçiminde tanımlanmıştır. Burada, $\psi(x)$ ağırlık fonksiyonudur ve $\psi(x) = 1$ olduğunda Cramér-von Mises istatistiği elde edilir (Laio, 2004). Dolayısıyla Cramér-von Mises test istatistiği,

$$W_n = \frac{1}{12n} + \sum_{i=1}^{n}\left[F_0(x_i) - \frac{2i-1}{2n}\right]^2 \tag{8}$$

biçiminde hesaplanmaktadır. W_n test istatistiğinin değeri kritik değer tablosundaki değerinden büyükse $H_0: F(x) = F_0(x)$ hipotezi reddedilir.

1.6.4. Anderson-Darling Uyum İyiliği Testi

Deneysel dağılım fonksiyonuna dayanan diğer bir uyum iyiliği testi de Anderson-Darling testidir ve ilk olarak Theodore W. Anderson ile Donald A. Darling tarafından 1952 yılında geliştirilmiştir (Anderson ve Darling, 1952). Bu test istatistiği Cramér-von Mises testinden yararlanılarak ortaya atılmıştır (Yap ve Sim, 2011). Anderson-Darling test istatistiği Eşitlik (7)'deki ağırlık fonksiyonu $\psi(x) = \left[F_0(x)\left(1 - F_0(x)\right)\right]^{-1}$ alındığında elde edilir. Bu durumda Anderson-Darling test istatistiği,

$$A_n = n\int_{-\infty}^{\infty} \frac{[F_n(x) - F_0(x)]^2}{F_0(x)\big(1 - F_0(x)\big)}\, dF_0(x) \tag{9}$$

biçiminde tanımlanır. Anderson-Darling testinde $F_0(x)$ veya $1 - F_0(x)$ değeri küçük olduğunda $\psi(x)$ ağırlık fonksiyonunun değeri daha da yüksek olacaktır. Dolayısıyla, bu yöntem Cramér-von Mises testiyle karşılaştırıldığında dağılımın kuyruklarındaki gözlemlere daha fazla ağırlık vermektedir (Tolikas ve Heravi, 2008). Anderson-Darling test istatistiği,

$$A_n = -n - \frac{1}{n}\sum_{i=1}^{n}\Big[(2i-1)\ln F_0\big(x_{(i)}\big) + (2n+1-2i)\ln\Big(1 - F_0\big(x_{(i)}\big)\Big)\Big] \tag{10}$$

biçiminde hesaplanır. A_n test istatistiğinin değeri, kritik değer tablosundaki değerinden büyükse $H_0: F(x) = F_0(x)$ hipotezi reddedilir.

Yokluk hipotezinde belirtilen olasılık dağılımına göre elde edilen A_n değerinin belirli bir sabitle çarpılması sonucunda değiştirilmiş Anderson-Darling test istatistiği (A_n^*) elde edilir (Yıldırım ve Gökpınar, 2012). A_n^* test istatistiği,

$$A_n^* = A_n\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right) \tag{11}$$

şeklinde hesaplanmaktadır. A_n^* test istatistiğinin kritik değerleri, dağılımı bilinmediği için simülasyonlar sonucunda belirlenir ve hesaplanan A_n^* değeri kritik değerden büyükse $H_0: F(x) = F_0(x)$ hipotezi reddedilir. Literatürdeki diğer uyum iyiliği testlerine göre Anderson-Darling testi daha tutarlı olmasına rağmen kritik değerler, hesaplanması bakımından dağılıma bağımlıdır ve bu nedenle her dağılım için kullanılamamaktadır.

1.7. Çalışmada Kullanılacak Dağılımlar

İstatistikte dağılımlar grafiksel olarak incelendiğinde çeşitli şekiller alabilirler. Bir dağılımın şekliyle ilgili özellikler, basit nicel tanımlayıcı istatistiklerin ve histogram gibi çizim tekniklerinin kullanıldığı istatistiksel veri analizi yöntemleriyle ortaya çıkmaktadır. İstatistiksel dağılımlar, simetrik ve simetrik olmayan dağılımlar olarak iki grupta sınıflandırılabilmektedir. Genel olarak, dengeli bir düzenden gelen frekans dağılımlarına

simetrik dağılımlar, dengesiz bir düzene sahip olanlara ise çarpık ya da simetrik olmayan dağılımlar denir.

1.7.1. Simetrik Dağılımlar

Simetrik bir dağılım, dağılımın sol ve sağ taraflarının ortalama etrafında kabaca eşit olarak dengelendiği bir dağılımdır. Simetrik dağılımlarda ortalama, ortanca ve mod dağılımın merkezinde yer almaktadır. Dağılım simetriden uzaklaştıkça bu üç değer birbirinden uzaklaşır. Simetrik dağılımlara, normal dağılım, düzgün dağılım, Laplace dağılımı ve student-t dağılımı örnek olarak verilebilir ve bu çalışmada da simülasyonlarda bu simetrik dağılımlardan faydalanılacaktır.

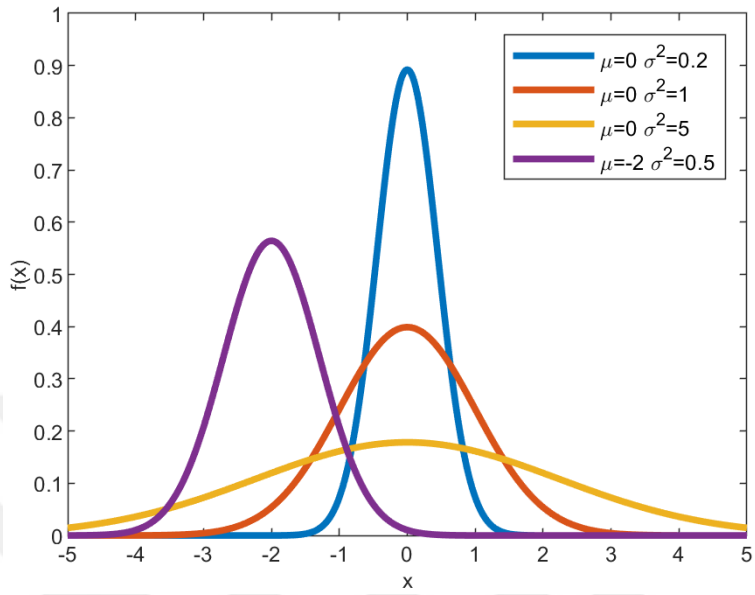
1.7.1.1. Normal Dağılım

Normal dağılım ilk olarak 1733 yılında Fransız matematikçi de Moivre tarafından binom olasılıklarına yaklaşan yardımcı bir fonksiyon olarak tanımlanmıştır (Balakrishnan ve Nevzorov, 2004). 19. yüzyılın başlarında, Laplace ve Gauss'un çalışmalarıyla normal dağılımın teorik önemi daha geniş bir şekilde ortaya koyulmuştur. Bu nedenle, normal dağılım bazen Gauss dağılımı, Gauss-Laplace dağılımı olarak da adlandırılır. Dağılımın gelişimi genellikle teoriyi gök cisimlerinin hareketlerine uygulayan Gauss'a atfedilir (Forbes vd., 2011). Normal dağılım, farklı uygulama alanlarında birçok istatistiksel çalışmanın temelini oluşturmaktadır. Bu dağılım, özellikle olasılık teorisinde benzersiz bir konuma sahiptir ve diğer dağılımlara bir yaklaşım olarak kullanılabilir. Aynı zamanda, çok çeşitli koşullar altında rastgele değişkenlerin toplamının asimptotik şeklidir. Normal dağılımın iki parametresi vardır. Bunlar, konum parametresi olan ortalama (μ) ve ölçek parametresi olan standart sapmadır (σ). Normal dağılımın olasılık yoğunluk fonksiyonu,

$$f(x;\mu,\sigma) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right), & -\infty < x < \infty \\ 0, & d.d \end{cases} \tag{12}$$

biçiminde ifade edilir. Bununla birlikte, normal dağılım ortalama ve standart sapma parametrelerinin değişimi sonucu birbirinden farklı yapılar göstermektedir. Normal

dağılımının farklı parametrelerine ait olasılık yoğunluk fonksiyonlarının grafiği Şekil 2'de verilmiştir.



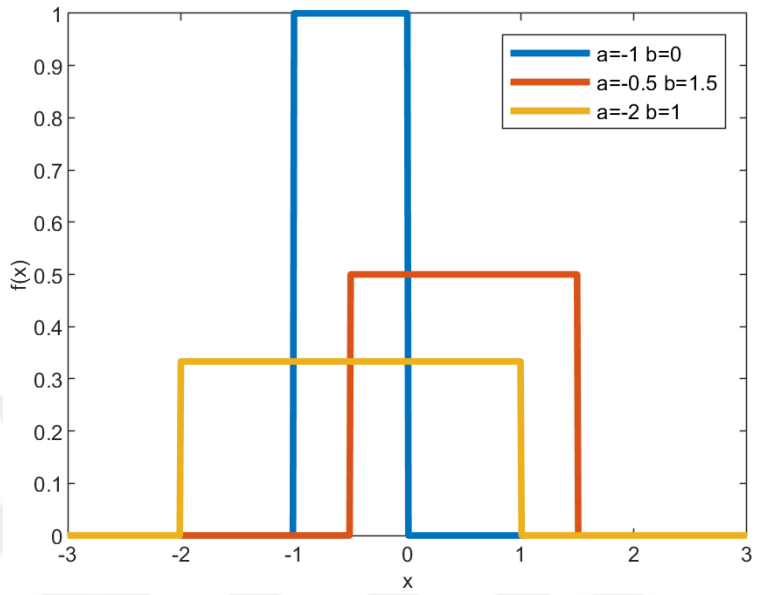
Şekil 2. Normal dağılımın farklı parametreli olasılık yoğunluk fonksiyonları

1.7.1.2. Düzgün Dağılım

Olasılık teorisinde, düzgün dağılım veya tekdüze dağılım bir simetrik olasılık dağılımı ailesidir. Düzgün dağılım, en basit olasılık dağılımıdır; ancak rastgele değişkenlerin modellenmesinde çok fazla kullanıldığı için istatistik teorisinde önemli bir rol almaktadır. Sürekli düzgün dağılım, belirli sınırlar arasında her sürekli değer için sabit olasılık değerine sahip deneyleri temsil eden istatistiksel bir dağılımdır (Dekking vd., 2005). Bu sınırlar dağılımın parametrelerini oluşturmaktadır. Bu parametreler minimum değer (a) ve maksimum değerdir (b). Düzgün dağılımın olasılık yoğunluk fonksiyonu,

$$f(x;a,b) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & d.d \end{cases} \tag{13}$$

biçiminde tanımlanır. Şekil 3'te, düzgün dağılımın farklı minimum ve maksimum değerlerine ait olasılık yoğunluk fonksiyonlarının grafiği verilmiştir.



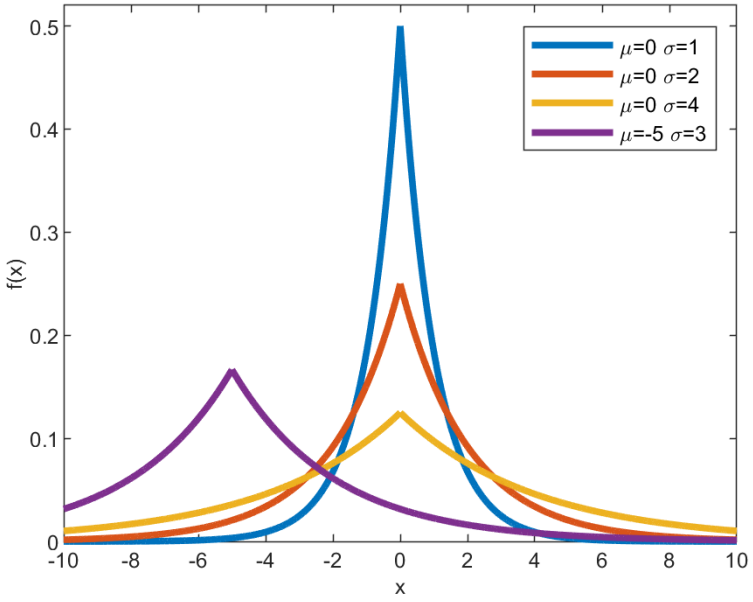
Şekil 3. Düzgün dağılımın farklı parametreli olasılık yoğunluk fonksiyonları

1.7.1.3. Laplace Dağılımı

En eski olasılık dağılımlarından biri olan Laplace dağılımı, adını aldığı Fransız matematikçi Pierre Simon Laplace tarafından 1774 yılında tanıtılmıştır (Johnson vd., 1995). Laplace dağılımı, normal dağılım gibi tek tepeli (tek modlu) ve aynı zamanda simetrik bir dağılımdır. Ancak normal dağılıma göre daha keskin bir tepe noktasına sahiptir. Genellikle yoğun kuyruklu veya verilerin normal dağılımdan daha yüksek bir tepe noktasına sahip olduğu deneyleri modellemek için kullanılmaktadır (Forbes vd., 2011). Laplace dağılımı, konum parametresi ve biri pozitif biri negatif olmak üzere birleştirilmiş iki üstel dağılımdan oluştuğu için çift üstel dağılım olarak da adlandırılmaktadır. Ayrıca bu dağılım, Laplece'ın birinci yasası olarak da bilinmektedir. Laplace dağılımının konum parametresi (μ) ve ölçek parametresi $(\sigma > 0)$ olmak üzere iki parametresi bulunmaktadır. Laplace dağılımının olasılık yoğunluk fonksiyonu,

$$f(x;\mu,\sigma) = \begin{cases} \frac{1}{2\sigma}\exp\left(-\frac{|x-\mu|}{\sigma}\right), & -\infty < x < \infty \\ 0, & d.d \end{cases} \quad (14)$$

biçiminde tanımlanmaktadır. Laplace dağılımının farklı konum ve ölçek parametreleri kullanılarak oluşturulan olasılık yoğunluk fonksiyonlarının grafiği Şekil 4'te verilmiştir.



Şekil 4. Laplace dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

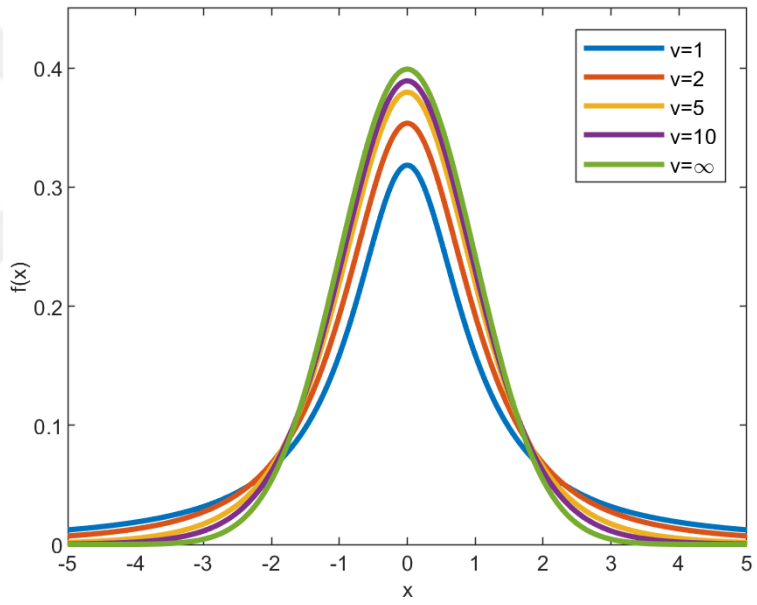
1.7.1.4. Student-t Dağılım

Student-t dağılımı ya da kısaca t-dağılımı, "Öğrenci" takma adıyla bilinen İngiliz istatistikçi William Sealy Gosset tarafından geliştirilmiştir (Student, 1908). Student-t dağılımı, normal dağılımın da içerisinde bulunduğu sürekli olasılık dağılımları ailesinde yer alan dağılımlardan biridir. Bu dağılım, iki örneklem ortalaması arasındaki farkın istatistiksel olarak anlamlı olup olmadığını test etmek için kullanılır (Forbes vd., 2011). Student-t dağılımı normal dağılım gibi simetrik ve çan eğrisi şeklindedir. Bununla birlikte, t-dağılımı normal dağılıma göre daha ağır kuyruklara sahiptir, yani ortalamadan uzak değerler üretmeye daha yatkındır. Student-t dağılımının şekil parametresi (v) olmak üzere tek bir

parametresi vardır. Bu parametre pozitif değerler almaktadır ve dağılımın serbestlik derecesini temsil etmektedir. Student-t dağılımının olasılık yoğunluk fonksiyonu,

$$f(x;v)=\begin{cases}\dfrac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{v\pi}}\left(1+\dfrac{x^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, & -\infty<x<\infty \\ 0, & d.d\end{cases} \tag{15}$$

biçiminde verilmektedir. Student-t dağılımının farklı şekil parametrelerine ait olasılık yoğunluk fonksiyonlarının grafiği Şekil 5'te gösterilmiştir ayrıca Student-t dağılımının serbestlik derecesi büyüdükçe normal dağılıma yaklaşmaktadır.



Şekil 5. Student-t dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

1.7.2. Simetrik Olmayan Dağılımlar

Simetrik olmayan dağılım, değişken değerlerinin düzensiz frekanslarda, ortalama, ortanca ve modun farklı noktalarda meydana gelmesi durumudur. Bir dağılım simetrik değilse, çarpık veya asimetrik dağılım olarak adlandırılır. Çarpıklık bir dağılımın simetriden ayrılma eğilimini göstermektedir. Simetrik olmayan bir dağılım, bir tarafta diğerine göre daha fazla yayılma eğiliminde olan değerlere sahiptir. Dolayısıyla, simetrik olmayan dağılım

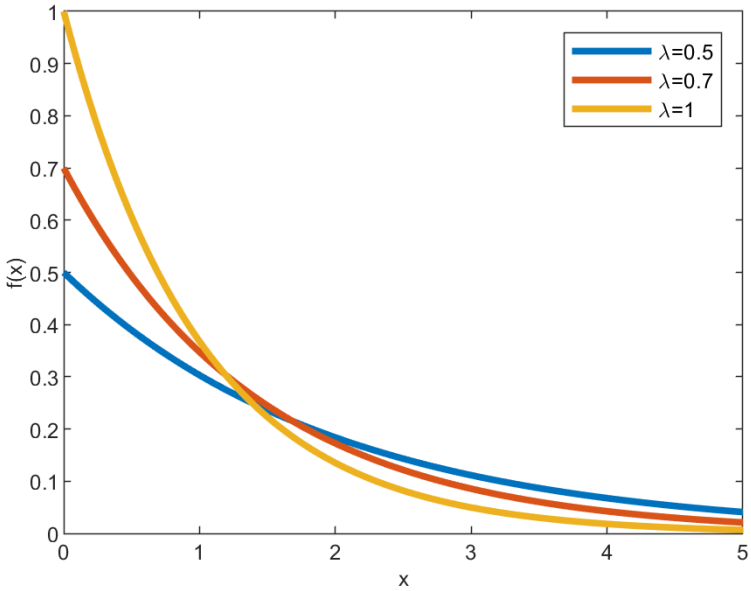
bir yönde veya diğer yönde belirgin bir kuyruğa sahip olan bir dağılımdır. Simetrik olmayan dağılımlar sola veya sağa çarpık olmaktadır. Üstel dağılım, log-normal dağılım, Weibull dağılımı, gamma dağılımı ve ki-kare dağılımı simetrik olmayan dağılımlara örnek olarak verilebilir ve bu çalışmada da simülasyonlarda bu simetrik olmayan dağılımlar kullanılmıştır.

1.7.2.1. Üstel Dağılım

Olasılık teorisi ve istatistikte, üstel dağılım genellikle belirli bir olay gerçekleşene kadar geçen süreyle ilgilenen sürekli, simetrik olmayan bir olasılık dağılımıdır. Örneğin, bir radyoaktif atomun bozunma süresi veya sabit arıza oranlarına sahip bileşenlerin arızalanma süresi üstel dağılım göstermektedir (Forbes vd., 2011). Sağ kalım analizi ve güvenilirlik gibi çeşitli alanlarda da yaygın olarak kullanılan üstel dağılım, 1985 yılında Karl Pearson tarafından tartışılan gamma dağılımının özel bir durumu olmasına rağmen, istatistiksel literatürde dağılımın kendi başına ortaya çıkması biraz zaman almıştır (Johnson vd., 2019). Bununla birlikte, üstel dağılımın diğer dağılımlardan farklı olarak unutkanlık özelliği de vardır. Üstel dağılımın λ ile gösterilen tek bir parametresi vardır ve bu parametre pozitif değerler almaktadır. Üstel dağılımın olasılık yoğunluk fonksiyonu,

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ \\ 0, & d.d \end{cases} \tag{16}$$

biçiminde tanımlanmaktadır. Şekil 6'da üstel dağılımının farklı parametre değerlerine ait olasılık yoğunluk fonksiyonlarının grafiği verilmiştir.



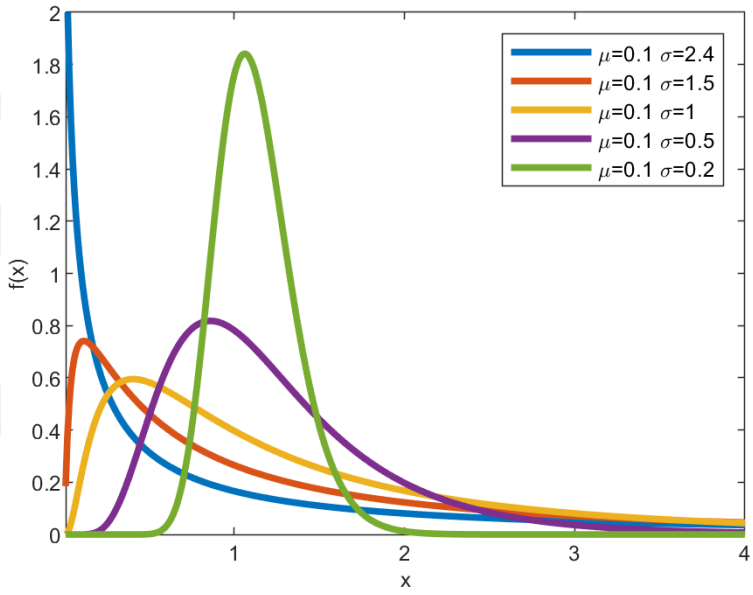
Şekil 6. Üstel dağılımın farklı parametreli olasılık yoğunluk fonksiyonları

1.7.2.2. Log-normal Dağılım

Log-normal dağılım, logaritması normal dağılım gösteren rastgele bir değişkenin olasılık dağılımıdır. Eğer X rastgele değişkeninin logaritması normal dağılım gösteriyorsa yani $\log X \sim N(\mu, \sigma)$ ise bu rastgele değişken bir log-normal dağılıma sahiptir. Bu rastgele değişkenin olasılık yoğunluk fonksiyonu normal dağılımın olasılık yoğunluk fonksiyonundaki değişkene logaritmik dönüşüm uygulanarak elde edilebilir. Normal dağılım simetrik bir dağılım olmasına rağmen log-normal dağılım simetrik olmayan ve pozitif çarpık bir dağılımdır. Log-normal olarak dağılmış bir rastgele değişken sadece pozitif değerler alır. Verilere logaritmik bir dönüşümün uygulanması, verilerin simetrik normal dağılıma yaklaşmasına olanak sağlayabilir; ancak log-normal dağılıma sahip rastgele değişkenlerin negatif değerler almaması bu prosedürün geçerliliğini sınırlayabilir (Forbes vd., 2011). Tıp, mühendislik, biyoloji, finans gibi farklı çalışma alanlarında log-normal dağılımın kullanımları mevcuttur. Özellikle, finans alanında çoğunlukla hisse senedi fiyatlarını, endeks değerlerini, varlık getirilerini ve ayrıca döviz kurlarını modellemek için log-normal dağılımdan yararlanılır. Log-normal dağılımın konum (μ) ve ölçek (σ) olmak üzere iki parametresi vardır ve olasılık yoğunluk fonksiyonu,

$$f(x,\mu,\sigma) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}, & x \geq 0 \\ 0, & d.d \end{cases} \quad (17)$$

biçiminde tanımlanmaktadır. Parametrelerinin değişimi sonucu log-normal dağılımın her biri birbirinden farklı yapılar göstermektedir. Log-normal dağılımın, farklı parametrelerine ait olasılık yoğunluk fonksiyonlarının grafiği Şekil 7'de gösterilmiştir.



Şekil 7. Log-normal dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

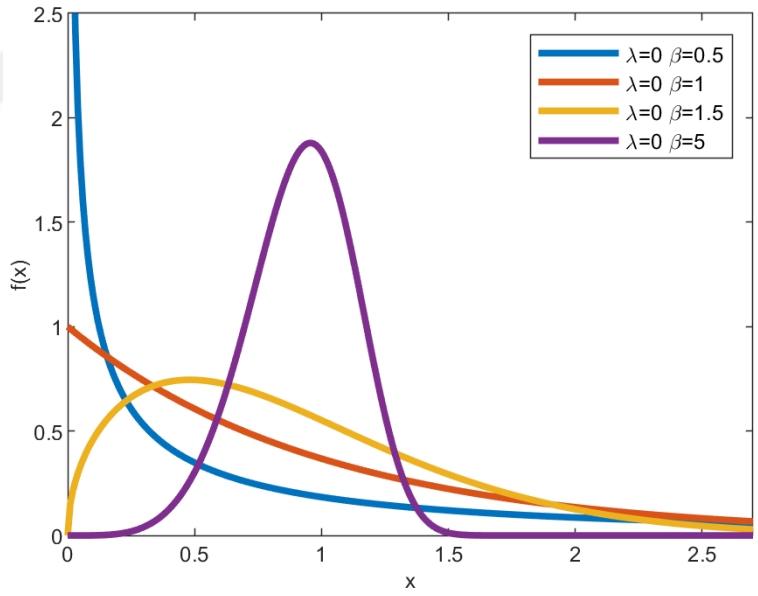
1.7.2.3. Weibull Dağılımı

Weibull dağılımı sürekli ve pozitif rastgele değişkenli bir olasılık dağılımıdır. Bu dağılım adını 1939 yılında malzeme mukavemetinin dağılımını temsil etmek ve 1951 yılında da çeşitli uygulamalar için kullanan İsveçli fizikçi Waloddi Weibull'dan almıştır (Weibull, 1939; Weibull, 1951). Buna rağmen, Weibull dağılımı, ilk olarak 1933 yılında Rosin ve Rammler tarafından bir parçacık boyutunun dağılımını tanımlamak için kullanılmıştır (Rosin ve Rammler, 1933). Güvenilirlik, sağ kalım, rüzgâr hızları gibi çeşitli verilerin modellenmesi için bu dağılımdan yararlanılmaktadır. Weibull dağılımının çok çeşitli kullanım alanlarının bulunmasının sebebi esnek bir yapıya sahip olmasıdır. Çünkü parametre

değişimi ile normal ve üstel dağılım gibi uygulamalarda daha yaygın olarak kullanılan farklı dağılımlar elde edilebilmektedir. Weibull dağılımının ölçek parametresi (λ) ve şekil parametresi (β) olmak üzere iki parametresi vardır. Bu dağılımın olasılık yoğunluk fonksiyonu,

$$f(x;\beta,\lambda) = \begin{cases} \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^{\beta}}, & x \geq 0 \\ 0, & d.d \end{cases} \tag{18}$$

biçimindedir. Eğer $\beta = 3.4$ alınırsa Weibull dağılımı normal dağılıma, $\beta = 2$ alınırsa Raiyleigh dağılımına, $\beta = 1$ alınırsa da üstel dağılıma dönüşmektedir (Altın Yavuz, 2021). Weibull dağılımının farklı parametre değerlerine ait olasılık yoğunluk fonksiyonlarının grafiği Şekil 8'de verilmiştir.



Şekil 8. Weibull dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

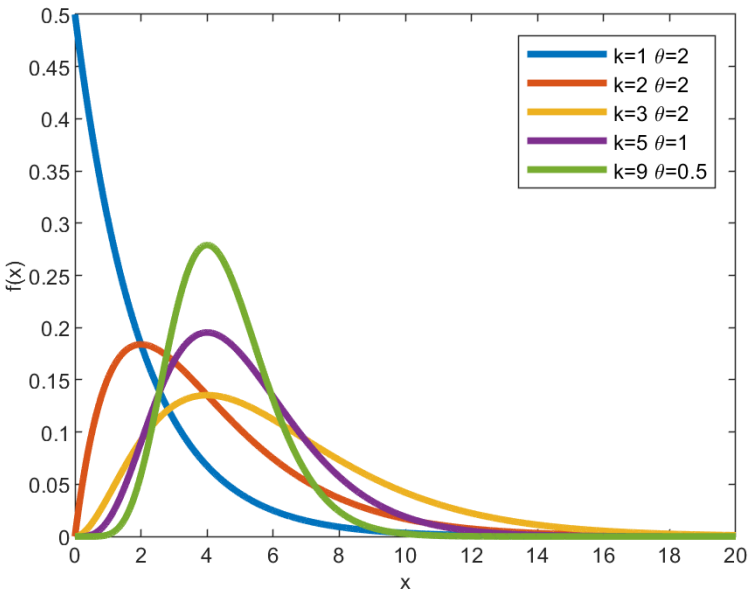
1.7.2.4. Gamma Dağılımı

Gamma dağılımı, her zaman pozitif değerler alan sürekli değişkenleri modellemek için işletme, finans ve mühendislik gibi alanlarda yaygın olarak kullanılan bir olasılık

dağılımıdır. Orijinden başlayan ve esnek bir yapıya sahip olan gamma dağılımı, ki-kare, Erlang ve üstel dağılımlarını özel bir durum olarak içermektedir (Forbes vd., 2011). Ortalama sabit bir sürede meydana gelen bağımsız olaylar arasındaki süreyi modellemek için gamma dağılımı kullanılmaktadır. Güvenilirlik analistleri de hata sürelerini modellemek için genellikle gamma dağılımından yararlanmaktadır. Gamma dağılımı ölçek parametresi (θ) ve şekil parametresi (k) olmak üzere iki parametreye sahiptir ve olasılık yoğunluk fonksiyonu,

$$f(x;k,\theta)=\begin{cases}\dfrac{1}{\Gamma(k)\theta^{k}}x^{k-1}e^{-\frac{x}{\theta}}, & x>0\\ 0, & d.d\end{cases} \tag{19}$$

biçiminde tanımlanmaktadır. Bu olasılık yoğunluk fonksiyonunda şekil parametresi olan k tamsayı değeri almışsa, gamma dağılımı Erlang dağılımına, $k=1$ olduğunda ise üstel dağılıma dönüşmektedir. Gamma dağılımının farklı parametreler kullanılarak elde edilen olasılık yoğunluk fonksiyonlarının grafiği Şekil 9'da gösterilmiştir.



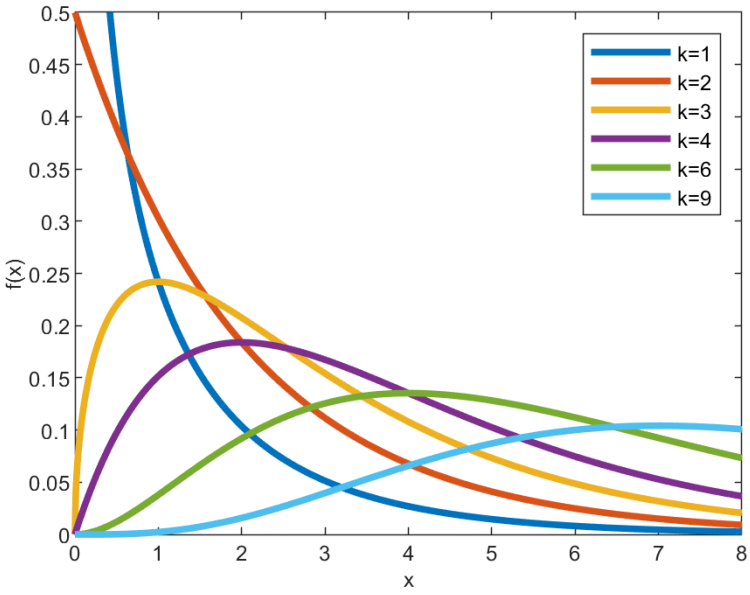
Şekil 9. Gamma dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

1.7.2.5. Ki-Kare Dağılımı

Ki-kare dağılımı, birçok hipotez testinde kullanılan k serbestlik derecesine sahip sürekli bir olasılık dağılımıdır ve gamma dağılımının özel bir halidir. Her biri standart normal dağılıma sahip k tane bağımsız rastgele değişkenin karelerinin toplamının olasılık dağılımı k serbestlik derecesine sahip ki-kare dağılımıdır (χ_k^2). Ki-kare dağılımının şekil parametresi (k) olmak üzere tek bir parametresi vardır. Bu parametre pozitif değerler almaktadır ve dağılımın serbestlik derecesini temsil etmektedir. Ki-kare dağılımının olasılık yoğunluk fonksiyonu,

$$f(x;k) = \begin{cases} \dfrac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & d.d \end{cases} \tag{20}$$

biçimindedir. Dağılımın şekli parametrenin farklı değerlerine göre değişkenlik göstermektedir ve farklı parametrelerine ait olasılık yoğunluk fonksiyonlarının grafiği Şekil 10'da verilmiştir.



Şekil 10. Ki-kare dağılımının farklı parametreli olasılık yoğunluk fonksiyonları

1.8. İstatistiksel Hipotez Testleri

İstatistiksel hipotez testi, elde edilen (gözlenen) verilere dayanarak ortaya atılan bir iddianın (hipotez) yeterince desteklenip desteklenmediğinin araştırılması için kullanılan istatistiksel bir yöntemdir. Hipotezlerin reddedilmesi veya gözlemlenen örneklerin beklenen sonuçlardan önemli ölçüde farklılık gösterip göstermediğine karar verilmesini sağlayan prosedürlere hipotez testleri, önem testleri veya karar kuralları denir (Ramachandran ve Tsokos, 2021). Hipotez testleri kusursuz değildir yani bir hipotezin doğruluğu kesin olarak kanıtlanamaz. Bir hipotez testinde birbiriyle karşılaştırılan iki hipotez bulunmaktadır. Bunlar yokluk hipotezi (H_0) ve alternatif hipotezdir (H_1). Hipotezler ifade edildikten sonra, yokluk hipotezinin reddedilip reddedilmeyeceğini belirlemek için uygun istatistiksel testler kullanılır. İstatistiksel testler sonucunda dört olası durum ve bu durumlar sonucunda da iki farklı türde hata meydana gelmektedir. Hataların her biri de farklı olasılıklarla oluşmaktadır. Bu hatalardan biri aslında doğru olduğu halde H_0 hipotezini reddetmektir. Buna I. tip hata denir ve bu hatanın meydana gelme olasılığı α ile temsil edilir. Yani, $P(H_0\ red \mid H_0\ doğru) = \alpha$'dır. Bununla birlikte, $1-\alpha$ olasılığı da güven düzeyi olarak adlandırılır. Hatalardan diğeri ise yanlış olduğu halde H_0 hipotezini kabul etmektir. Bu hata türüne de II. tip hata denir ve hatanın meydana gelme olasılığı β ile temsil edilir. Yani, $P(H_0\ kabul \mid H_0\ yanlış) = \beta$'dır. $1-\beta$ olasılığı da testin gücünü ifade etmektedir. İstatistiksel hata türleri ve olasılıkları Tablo 1'de verilmiştir.

Tablo 1. İstatistiksel hata türleri ve olasılıkları

		İstatistiksel Karar	
		H_0 Kabul edildi.	**H_0 Reddedildi.**
Gerçek	**H_0 Doğru**	Doğru Karar $(1-\alpha)$ →Güven Düzeyi	I. Tip Hata (α)
	H_0 Yanlış	II. Tip Hata (β)	Doğru Karar $(1-\beta)$ →Testin Gücü

2. YAPILAN ÇALIŞMALAR

Uyum iyiliği testlerinde yokluk hipotezinin kabul edilip edilmeyeceğine karar vermek için test istatistiğine veya test istatistiği kullanılarak hesaplanan p-değerine bakılmaktadır. Bu nedenle araştırmalarda p-değeri çok önemli bir parametredir. Fakat p-değerinin küçük bir değişimi veya anlamlılık düzeyinin iyi seçilmemesi yokluk hipotezinin doğru olduğu halde reddedilmesine veya yanlış olduğu halde kabul edilmesine sebep olabilmektedir. p-değeri genellikle uyum iyiliği testlerinin sonuçlarının güvenirliliğini ortaya koymaktadır. Ancak literatürde yapılan çalışmalarda p-değerinin stabilitesindeki sorunlar her zaman güvenilir olmadığını göstermiştir (Abelson, 1997; Hubbard ve Lindsay, 2008; Sullivan ve Feinn, 2012). Bu nedenle, bu eksikliklerden kaçınmak için bu çalışmada, literatürdeki uyum iyiliği testlerine alternatif olarak yeni bir uyum iyiliği test yöntemi ve uyum iyiliği testlerinin güvenilirlik derecesi için ölçüt olarak kullanılabilecek yeni bir olabilirlik indeksi geliştirilmiştir. Önerilen yöntemde, yönsel yaklaşım temel alınarak bir olabilirlik indeksi hesaplanacak ve bu değere göre örneklemin iddia edilen dağılıma ne kadarlık bir uyum gösterdiği yorumlanabilecektir. Böylece, $p < \alpha$ olması durumunda kurulan yokluk hipotezini doğrudan reddetmek yerine, örneklemin geldiği düşünülen dağılımı ne kadar yansıttığı hesaplanacak ve ne kadarlık bir uyumun çalışma için kabul edilebileceği araştırmacıya bırakılacaktır. Geliştirilen olabilirlik indeks değerine ek olarak bu çalışmada, indeks değerinin dağılımı incelenerek bağımsız ki-kare uyum iyiliği testi önerilmiştir.

2.1. Uyum İyiliğinin Olabilirlik İndeksi

İstatistik biliminde, olasılıksal integral dönüşümüne göre keyfi olasılık dağılımı gösteren bir rastgele değişken, $[0, 1)$ aralığında tanımlı düzgün dağılımdan bir rastgele değişkene dönüştürülebilir (Angus, 1994). Yani, X sürekli rastgele değişkeninin birikimli dağılım fonksiyonunun $F_X(x)$ olduğunu varsayılsın $\left(X \sim F_X(x)\right)$. Bu durumda, $U = F_X(X)$ eşitliğiyle düzgün dağılıma sahip U rastgele değişkeni elde edilebilir (Dodge, 2006).

Uyum iyiliği testlerinde yokluk hipotezi ve alternatif hipotez;

$H_0: F(x) = F_0(x)$
$H_1: F(x) \neq F_0(x)$

biçiminde verilir. Eğer x değerleri $F_0(x)$ dağılımdan geliyorsa,

$$U_i = F_0(X_i) \tag{21}$$

eşitliği yazıldığında $U_i \sim Düzgün\ (0,1)$ tanımlaması geçerli olur. Daha sonra,

$$V_i = 2\pi U_i \tag{22}$$

dönüşümü yapılarak V_i değişkeni $(0, 2\pi)$ aralığında düzgün dağılıma dönüştürülür. Yani, $V_i \sim Düzgün\ (0, 2\pi)$ olur. Bu durumda, V_i değişkeni birim çember üzerinde tanımlanmış verilere dönüşür. Bu veriler birer birim vektör kabul edilerek ortalama bileşke vektör bulunmak istenirse, bu bileşke vektörün ortalama yatay bileşeni,

$$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} \cos(V_i) \tag{23}$$

ve ortalama düşey bileşeni,

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} \sin(V_i) \tag{24}$$

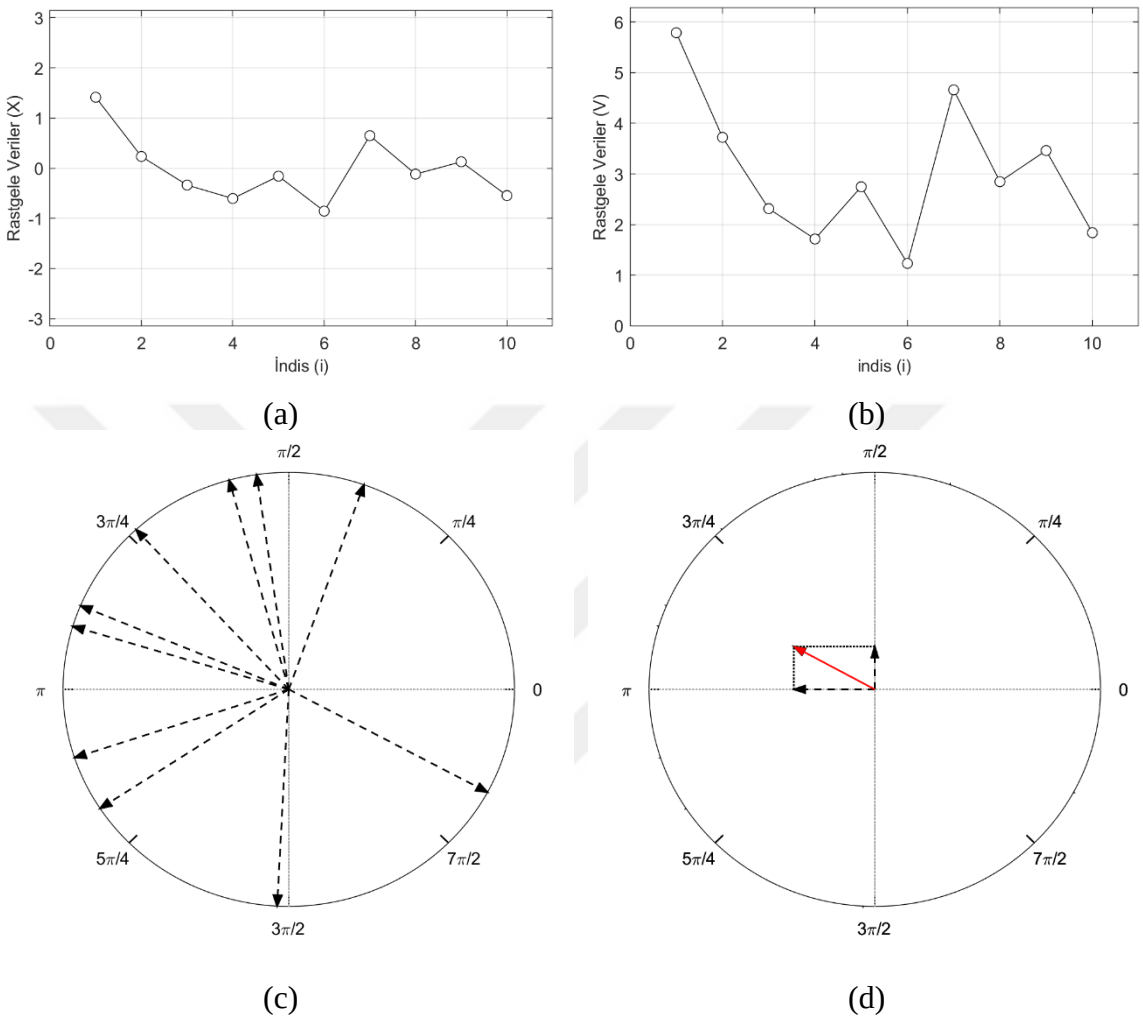
biçiminde tanımlanmaktadır. Birim çember üzerinde vektör şeklinde gösterilen dairesel verilerin, bileşke vektör uzunluğu (R),

$$R = \sqrt{\bar{C}^2 + \bar{S}^2} \tag{25}$$

şeklinde hesaplanmaktadır. Bileşke vektör uzunluğu $0 \leq R \leq 1$ aralığında değer almaktadır. Bileşke vektör uzunluğu çember üzerindeki verinin yoğunlaşmasının bir ölçüsüdür (Mardia, 1972). Bileşke vektör ile yönsel yaklaşım kullanılarak önerilen olabilirlik indeks değeri,

$$\Pi = 1 - R \tag{26}$$

biçiminde belirlenir. Önerilen olabilirlik indeks değerinin grafiksel gösterimi Şekil 11'de verilmiştir.



Şekil 11. Önerilen olabilirlik indeks yönteminin grafiksel gösterimi (a) Oluşturulan rastgele veriler, (b) $(0, 2\pi)$ aralığına dönüştürülen veriler, (c) Dönüştürülen verilerin radyal sistemde gösterimi, (d) Verilerin bileşke kuvveti

Şekil 11(a)'da standart normal dağılımdan rastgele üretilmiş on tane örnek, Şekil 11(b)'de ise bu verilerin yönsel verilere $\left(Düzgün\ (0, 2\pi)\right)$ dönüştürülmüş değerleri görülmektedir. Şekil 11(c)-(d)'de de yönsel verilerin birim vektör gösterimi ve bu verilerin bileşke vektörü yer almaktadır.

Bileşke vektör uzunluğunun sıfıra yaklaşması $(R \to 0)$ durumunda olabilirlik indeks değeri bire yaklaşacaktır $(\Pi \to 1)$. Bu durum, verinin birim çember üzerinde herhangi bir noktada yoğunlaşma göstermediğinin tüm çembere eşit şekilde dağıldığını, yani düzgün

dağılıma sahip olduğunun göstergesidir. Öte yandan, bileşke vektör uzunluğunun bire yaklaşması $(R \to 1)$ durumunda olabilirlik indeks değeri sıfıra yaklaşacaktır $(\Pi \to 0)$. Bu ise verilerin birim çember üzerinde herhangi bir noktada yoğunlaşma gösterdiğinin tüm çembere eşit dağılmadığını, yani düzgün dağılıma sahip olmadığını gösterir. Algoritma 1'de önerilen olabilirlik indeks yönteminin algoritması verilmiştir.

Algoritma 1. Önerilen olabilirlik indeksinin hesaplanması

Adım 1. İddia edilen teorik dağılım fonksiyonunu belirle,

Adım 2. Örneklem değerlerini bu özel dağılımda yerine koyarak (Eşitlik (21)), $[0, 1)$ aralığında U_i değerlerini elde et,

Adım 3. $[0, 1)$ aralığındaki U_i değerlerini, 2π ile çarparak (Eşitlik (22)) yönsel V_i değerlerine çevir,

Adım 4. V_i yönsel değerlerinin, Eşitlik (23)'ü kullanarak ortalama düşey bileşenini $(\bar{S})$ ve Eşitlik (24)'ü kullanarak ortalama yatay bileşenini $(\bar{C})$ hesapla,

Adım 5. Eşitlik (25)'teki gibi düşey bileşen ile yatay bileşenin bileşkesini alarak bileşke vektörü (R) hesapla,

Adım 6. Eşitlik (26) yardımıyla, birden ortalama bileşke vektörünün değerini çıkararak olabilirlik indeksini (Π) hesapla.

Yönsel yaklaşım kullanılarak hesaplanacak olan olabilirlik indeks değeri örneklemin iddia edilen dağılımı ne derece yansıttığını ortaya koyabilecektir. Böylece, önerilen olabilirlik indeksi, uyum iyiliği testleri için sayısal bir ölçüt olarak kullanılabilecektir.

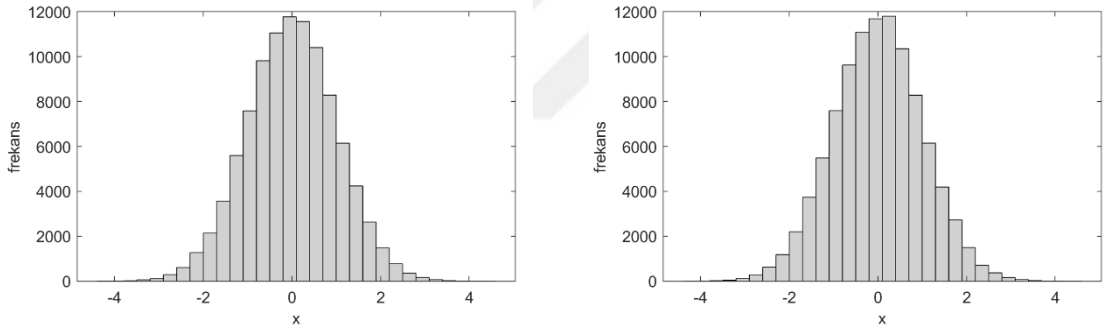
2.2. Bağımsız Ki-Kare Uyum İyiliği Testi

Bu çalışmada, önerilen olabilirlik indeks değerine ek olarak, bir çember üzerinde tanımlanmış verilere dönüştürülen verilerin yoğunlaşmasının bir ölçüsü olarak tanımlanan bileşke vektörün (R) dağılımından yola çıkılarak yeni bir uyum iyiliği test yöntemi geliştirilmiştir.

2.2.1. Bileşke Vektörün Dağılımının Elde Edilmesi

Önerilen olabilirlik indeks değeri hesaplanırken dairesel verilerin istatistiklerinden faydalanılmıştır. Öncelikle, $U_i = F_0(X_i)$ durumunda $U_i \sim Düzgün(0,1)$ tanımlaması geçerli olacağından ve $V_i = 2\pi U_i$ dönüşümü uygulandığında $V_i \sim Düzgün(0, 2\pi)$ olacaktır. Böylelikle, V_i değişkeni birim çember üzerinde tanımlanmış verilere dönüşür. V_i değişkeninin sinüslerinin ve kosinüslerinin ortalamaları alınarak ortalama yatay ve düşey bileşeni bulunur. Daha sonra da bu bileşenler kullanılarak birim çember üzerindeki verinin yoğunlaşmasının bir ölçüsü olan bileşke vektör hesaplanır. Bileşke vektörün dağılımı incelendiğinde, bu değişkenin ki-kare dağılımına yakınsadığı görülmektedir.

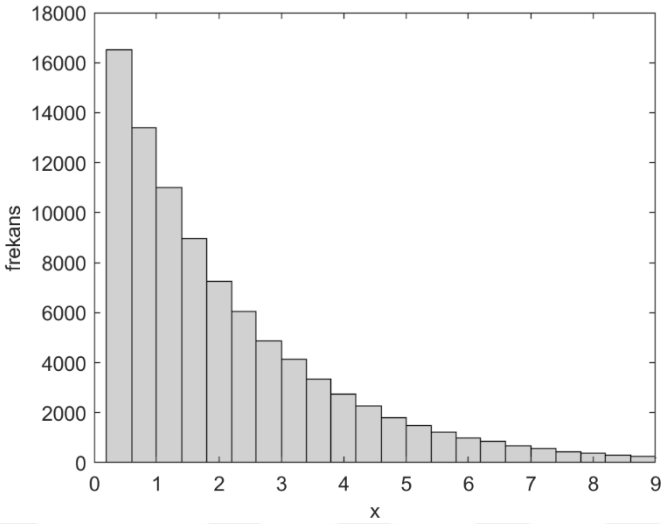
V_i değişkenleri, $(0, 2\pi)$ aralığında düzgün dağılıma sahip olduğundan sinüslerinin ve kosinüslerinin ortalaması alındığında elde edilen iki rastgele değişken $(\bar{C}, \bar{S})$ normal dağılıma sahip olmaktadır. V_i değişkeninin ortalama yatay ve düşey bileşeninin histogram grafiği Şekil 12'de verilmiştir.



(a) (b)

Şekil 12. V_i değişkeninin ortalama yatay ve düşey bileşeninin histogramı (a) V_i değişkeninin ortalama yatay bileşeninin histogramı, (b) V_i değişkeninin ortalama düşey bileşeninin histogramı

Bileşke vektörün karesi, Eşitlik (25)'teki gibi bu iki rastgele değişkenin karelerinin toplamı kullanılarak hesaplanmaktadır. Her biri standart normal dağılıma sahip k tane bağımsız rastgele değişkenin karelerinin toplamının olasılık dağılımı k serbestlik derecesine sahip χ_k^2 dağılımıdır (Balakrishnan ve Nevzorov, 2004). Bu bilgiden yararlanılarak, yatay ve düşey bileşenler standart normal dağılım göstermedikleri için öncelikle standartlaştırma işlemi uygulanmıştır. Standartlaştırma uygulanmış bileşenler kullanılarak elde edilen bileşke vektörün karesinin histogramı Şekil 13'te gösterilmiştir.



Şekil 13. Bileşke vektörün karesinin histogramı

Şekil 13'te de görüldüğü gibi bileşke vektörün karesinin (R^2) dağılımı, iki rastgele değişkenin karelerinin toplamı şeklinde olduğundan, $k = 2$ serbestlik derecesine sahip χ^2 dağılımına sahiptir.

İspat 1:

Ortalama yatay ve düşey bileşen sırasıyla

$$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} \cos(V_i), \tag{27}$$

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} \sin(V_i) \tag{28}$$

biçiminde hesaplanır. Burada $(V_1, V_2, \dots, V_n) \sim Düzgün(0, 2\pi)$, $C = \cos V$ ve $S = \sin V$'dir. Daha sonra bu bileşenlerin birinci ve ikinci momentleri hesaplandığında

$$E(C) = \int_0^{2\pi} \cos V \frac{1}{2\pi} dV = 0, \tag{29}$$

$$E(S) = \int_0^{2\pi} \sin V \frac{1}{2\pi} dV = 0 \tag{30}$$

elde edilir ve benzer olarak,

$$E(C^2) = E(S^2) = \frac{1}{2} \tag{31}$$

$$E(C \cdot S) = 0 \tag{32}$$

biçiminde hesaplanır. Bu durumda, $(0, 2\pi)$ aralığında düzgün dağılıma sahip olan rastgele değişkenler, sinüs ve kosinüslerinin ortalaması alınarak $N(\mu, \Sigma)$ dağılımına dönüştürülmüş olur. Böylece,

$$\vec{\mu} = 0, \quad \Sigma = \begin{bmatrix} \frac{1}{2n} & 0 \\ 0 & \frac{1}{2n} \end{bmatrix} \tag{33}$$

olur. Eğer $(\bar{C}, \bar{S})$ rastgele değişkenleri standartlaştırılırsa, $\left(\sqrt{2n}\bar{C}\right)$ ve $\left(\sqrt{2n}\bar{S}\right)$ rastgele değişkenleri $Normal(0, 1)$ dağılımından gelmektedir. Yani,

$$\begin{aligned} \sqrt{2n}\bar{C} &= \sqrt{2n}\frac{1}{n}\sum_{i=1}^{n} \cos(V_i) \\ &= \sqrt{\frac{2}{n}}\sum_{i=1}^{n} \cos\left(2\pi F_0(x_i)\right), \end{aligned} \tag{34}$$

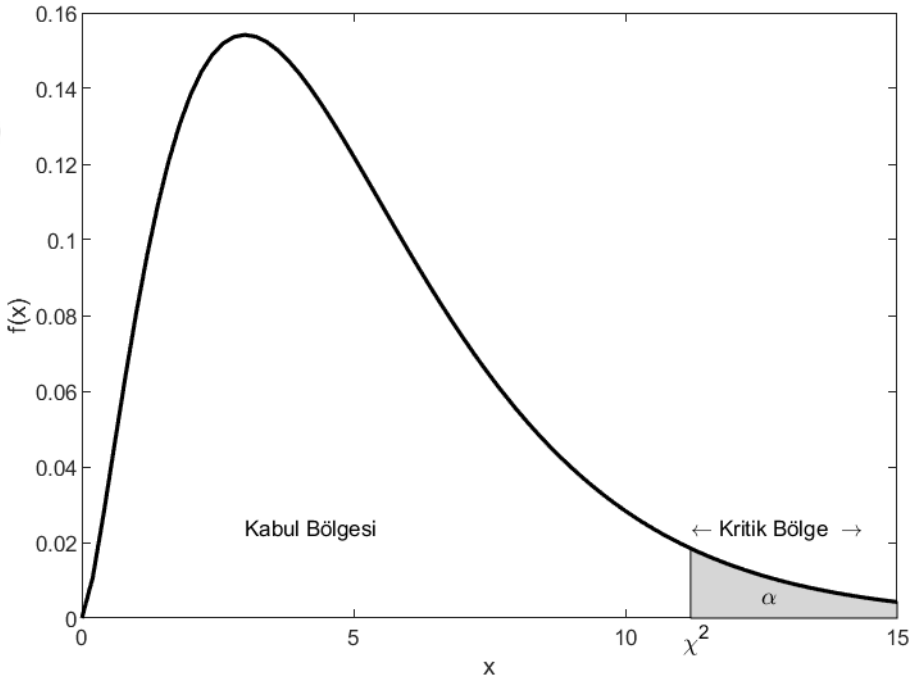
ve benzer şekilde,

$$\begin{aligned} \sqrt{2n}\bar{S} &= \sqrt{2n}\frac{1}{n}\sum_{i=1}^{n} \sin(V_i) \\ &= \sqrt{\frac{2}{n}}\sum_{i=1}^{n} \sin\left(2\pi F_0(x_i)\right) \end{aligned} \tag{35}$$

olur. $\left(\sqrt{2n}\bar{C}\right)$ ve $\left(\sqrt{2n}\bar{S}\right)$ değişkenleri $Normal(0,1)$ dağılım gösterdiğinden $\left(\sqrt{2n}\bar{C}\right)^2$ ve $\left(\sqrt{2n}\bar{S}\right)^2$ değişkenleri $k=1$ serbestlik dereceli ki-kare dağılımına sahiptirler ve bu değişkenlerin toplamı olan R^2 değişkeni de $k=2$ serbestlik dereceli ki-kare dağılımına sahiptir (χ_2^2).

2.2.2. Kritik Tablo Değerlerinin Elde Edilmesi

Ki-kare uyum iyiliği testi sağ kuyruk testidir ve bu test uygulanırken, hesaplanan test istatistiğinin ki-kare dağılımının ana kütlesinde ya da kritik bölgesinde olup olmadığı bilinmelidir. Kritik bölgenin sınırı kritik değer olarak bilinir. Kritik değer de testin anlamlılık düzeyine bağlıdır. Anlamlılık düzeyi genellikle 0.05 veya 0.01 olarak alınır. Kritik değerler de ki-kare test tablosundan bulunur. Eğer test istatistiği kritik bölge içindeyse H_0 hipotezi reddedilir. Şekil 14'te, ki-kare uyum iyiliği testinin kabul ve kritik bölgesi verilmiştir.



Şekil 14. Ki-kare uyum iyiliği testinin kabul ve kritik bölgesi

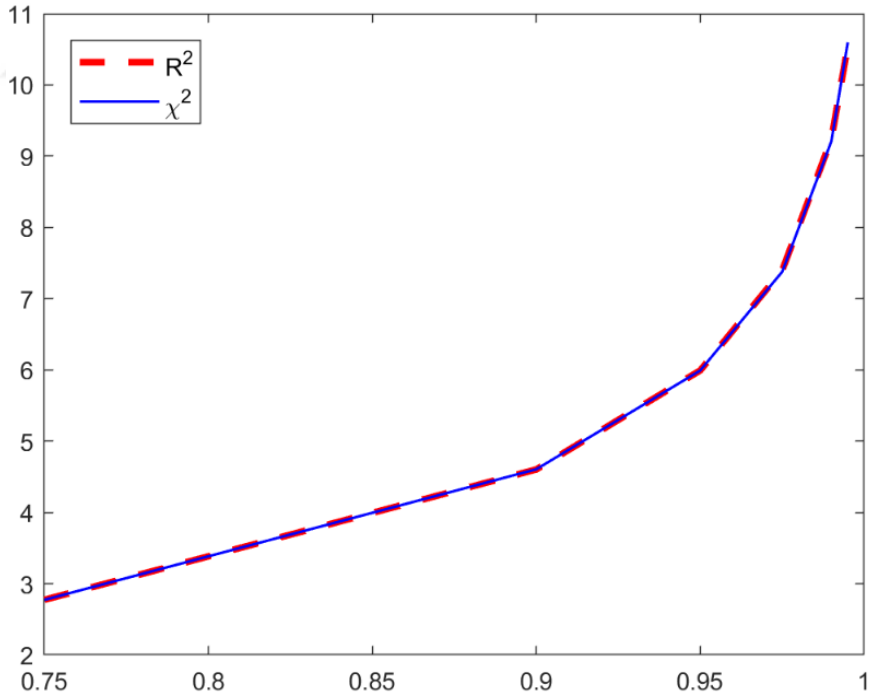
Bileşke vektörün dağılımından yararlanılarak geliştirilen bağımsız ki-kare testine ait kritik tablo değerlerini elde etmek için kitle olarak normal dağılım seçilmiştir. Normal dağılımdan örneklem büyüklüğü $n=200$ alınarak rastgele veriler üretilmiştir. Daha sonra,

bu veriler önerilen yöntemdeki adımlar kullanılarak çember üzerinde tanımlanmış yönsel verilere dönüştürülmüştür ve verilerin bileşke vektörünün karesi (R^2) hesaplanmıştır. Bu işlem 10000 kez tekrarlanmıştır ve elde edilen R^2 ve χ^2 tablo değerleri Tablo 2'de verilmiştir.

Tablo 2. Farklı α değerlerine göre R^2 ve χ^2 tablo değerleri

$\boldsymbol{\alpha}$	**0.25**	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
$\boldsymbol{R^2}$	2.775757	4.605028	5.987124	7.372935	9.180856	10.55664
$\boldsymbol{\chi^2}$	2.772589	4.60517	5.991465	7.377759	9.21034	10.59663

Tablo 2 incelendiğinde, R^2 ve χ^2 tablo değerlerinin birbirine oldukça yakın çıktığı görülmektedir. Bu değerlerin grafiksel olarak karşılaştırılması da Şekil 15'te gösterilmiştir.



Şekil 15. R^2 ve χ^2 tablo değerlerinin grafiksel olarak karşılaştırılması

Şekil 15 incelendiğinde, R^2 ve χ^2 tablo değerlerinin tamamen birbiriyle örtüştüğü görülmektedir. Bu durum, bileşke vektör kullanılarak önerilen bağımsız ki-kare uyum iyiliği

testinin sürekli dağılımlara da kolaylıkla uygulanabileceğini göstermektedir. Önerilen bağımsız ki-kare testi sınıf sayısından bağımsız olacağından, literatürdeki ki-kare uyum iyiliği testinin en önemli eksikliklerinden biri olan sınıflandırmadan kaynaklı meydana gelebilecek bilgi kaybını ortadan kaldırılabilecektir.

2.2.3. Bağımsız Ki-Kare Uyum İyiliği Test İstatistiğinin Hesaplanması

Bir çember üzerinde tanımlanmış yönsel verilere dönüştürülen verilerin yoğunlaşmasının bir ölçüsü olan ortalama bileşke vektörün dağılımı incelendiğinde, bu değişkenin ki-kare dağılıma sahip olduğu görülmüştür. Bu yaklaşımdan yararlanılarak literatürdeki ki-kare uyum iyiliği testinin aksine sınıf sayısından bağımsız ve sürekli dağılımlar için de kullanılabilecek olan bağımsız ki-kare uyum iyiliği testi önerilmiştir. Ayrıca farklı anlamlılık düzeyleri için R^2 ve χ^2 tablo değerlerinin örtüştüğü Tablo 2'de görülmüştür. Önerilen bağımsız ki-kare uyum iyiliği testinde öncelikle her veri birer birim vektör kabul edilerek bileşke vektör bulunmak istenirse, bu bileşke vektörün standartlaştırılmış ortalama yatay bileşeni,

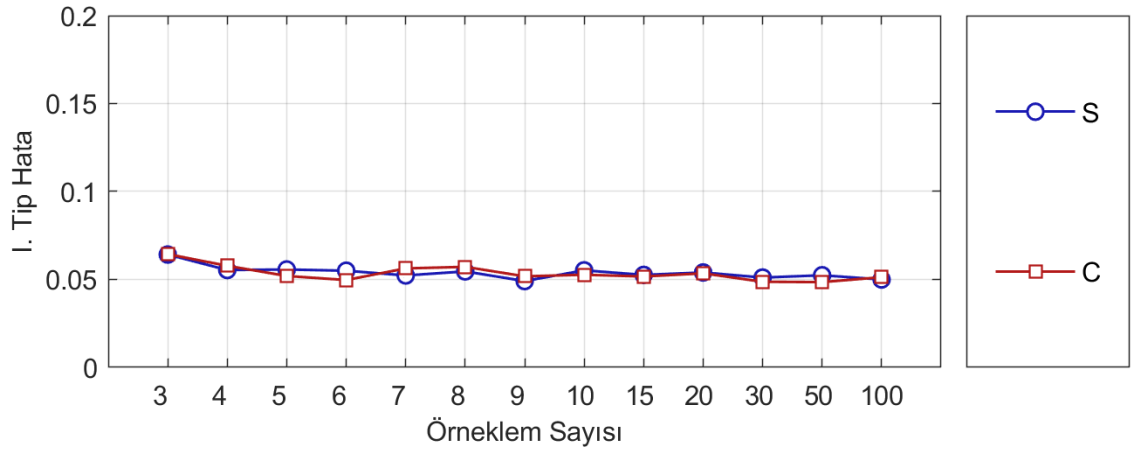
$$C = \sqrt{\frac{2}{n}} \sum_{i=1}^{n} \cos\left(2\pi F_0(x_i)\right) \tag{36}$$

ve standartlaştırılmış ortalama düşey bileşeni,

$$S = \sqrt{\frac{2}{n}} \sum_{i=1}^{n} \sin\left(2\pi F_0(x_i)\right) \tag{37}$$

biçiminde tanımlanır. Her yatay ve düşey bileşen rastgele değişkenlerinden yeteri kadar örnek alınırsa, merkezi limit teoremine göre standart normal dağılıma yakınsar. Bu yeterli örnek sayısının belirlenmesi için yapılan simülasyon çalışmasında $[0, 2\pi]$ aralığında düzgün dağılımdan üretilen farklı büyüklükteki rastgele örneklerin (C, S) değerleri hesaplanmıştır. Her C ve S rastgele değerlerinden 100'er tane oluşturularak bir veri seti elde edilmiş ve standart normal dağılımdan gelip gelmediği Kolmogorov-Smirnov uyum iyiliği testi ile test edilmiş ve I. tip hataları hesaplanmıştır. Bu işlem 10000 kez tekrar edilmiştir. C ve S

değişkenlerinin Kolmogorov-Smirnov testi sonucundaki I. tip hataları Şekil 16'da verilmiştir.



Şekil 16. C ve S değişkenlerinin Kolmogorov-Smirnov testi sonucundaki I. tip hataları

Şekil 16'da görüldüğü gibi örneklem büyüklüğü 4'ten küçük olduğunda I. tip hata yükselmektedir yani normallik varsayımı geçersiz olmaktadır. Buna göre, önerilen yöntemin sadece 4 ve 4'ten büyük örneklemler için kullanılabilir olduğu söylenebilir.

Birim çember üzerinde vektör şeklinde gösterilen yönsel verilerin, standart bileşke-kare vektörü,

$$R^2 = C^2 + S^2 \tag{38}$$

şeklinde tanımlanır. Bu bileşke-kare vektör genel gösterim olarak,

$$R^2 = \frac{2}{n}\left[\left(\sum_{i=1}^{n}\cos\left(2\pi F_0(X_i)\right)\right)^2 + \left(\sum_{i=1}^{n}\sin\left(2\pi F_0(X_i)\right)\right)^2\right] \tag{39}$$

biçiminde verilebilir. Elde edilen R^2, önerilen bağımsız ki-kare uyum iyiliği yönteminin test istatistiğini vermektedir. Bu bileşke-kare vektör, standart normal dağılımdan gelen iki rastgele değişkenin karelerin toplamı olan 2 serbestlik dereceli ki-kare dağılımına sahiptir. Önerilen bağımsız ki-kare uyum iyiliği testinin algoritması, Algoritma 2'de verilmiştir.

Algoritma 2. Bağımsız ki-kare uyum iyiliği testi

Adım 1. Yokluk hipotezi için teorik dağılım fonksiyonunu ve parametreleri tanımla,

Adım 2. Test istatistiği R^2'yi hesapla,

Adım 3. Seçilen α değerine göre 2 serbestlik dereceli χ^2 tablosundan kritik değeri belirle,

Adım 4. Test istatistiği R^2 kritik değerden $\chi^2_{2,1-\alpha}$ büyükse yokluk hipotezini reddet.

Önerilen bağımsız ki-kare uyum iyiliği testinin kritik değerleri, tablo değerine gerek duymadan sadece tek satırlı χ^2_2 tablosu kullanılarak verilebilir.

Tablo 3. Farklı α değerlerine göre bağımsız ki-kare testinin kritik tablo değerleri

$\boldsymbol{\alpha}$	$\mathbf{0.25}$	$\mathbf{0.1}$	$\mathbf{0.05}$	$\mathbf{0.025}$	$\mathbf{0.01}$	$\mathbf{0.005}$
$\boldsymbol{\chi^2_{2,1-\alpha}}$	2.772589	4.60517	5.991465	7.377759	9.21034	10.59663

Önerilen bağımsız ki-kare uyum iyiliği testi, sınıf sayısı ve dağılımdan bağımsız olmasının yanı sıra sadece iki serbestlik derecesine sahip χ^2_2 kritik değerlerini kullandığı için serbestlik derecesinden de bağımsızdır. Ayrıca kullanım olarak da basit ve kolay hesaplanabilir bir yöntemdir. Tablo 4'te verilen örnekte, önerilen yöntemin basit bir şekilde uygulanabilir olduğunu gösterilmiştir. Örnekte verilen 10 büyüklüğündeki örneklem için yokluk hipotezi, örneklemin $(\mu = 10, \sigma^2 = 4)$ parametreli normal dağılıma sahip olduğu şeklinde verilmiştir. Bu yokluk hipotezi $\alpha = 0.05$ anlamlılık düzeyinde test edilmiştir.

Tablo 4. Bağımsız ki-kare testinin uygulaması

i	x_i	$V_i = 2\pi F_0(x_i)$	$C = \cos V_i$	$S = \sin V_i$
1	13.5142	6.0353	0.9694	-0.2453
2	10.2069	3.4004	-0.9667	-0.2559
3	14.9308	6.2402	0.9991	-0.0430
4	11.5484	4.9046	0.1910	-0.9816
5	11.1112	4.4658	-0.2441	-0.9698
6	11.8232	5.1460	0.4202	-0.9074
7	9.6953	2.7612	-0.9285	0.3713
8	8.8668	1.7938	-0.2212	0.9752
9	11.9410	5.2408	0.5041	-0.8636
10	8.0430	1.0299	0.5149	0.8573
		Σ	1.2383	-2.0628
		$(\Sigma)^2$	1.5334	4.2553
		R^2	**1.1577**	
		Sonuç $(\alpha = 0.05)$	**1.1577<5.9915 : H_0 Reddedilemez**	

Yukarıdaki tabloda yapılan hesaplamalar yokluk hipotezinin reddedilemeyeceğini göstermektedir. Tüm hesaplamalar basit bir hesap makinesi ile kolaylıkla yapılabilmektedir.

3. BULGULAR VE İRDELEME

Bu bölümde, önerilen olabilirlik indeksinin ve bağımsız ki-kare uyum iyiliği testinin doğruluğunu, geçerliliğini ve performansını kapsamlı bir şekilde test etmek için farklı simülasyon çalışmaları yapılmıştır. Simülasyon çalışmalarının birinci kısmında, uyum iyiliğinin değerlendirilmesinde bir ölçüt olarak önerilen olabilirlik indeksinin doğruluğunun ve geçerliliğinin belirlenmesi amaçlanmıştır. Simülasyon çalışmalarının ikinci kısmında da önerilen bağımsız ki-kare uyum iyiliği testi Monte Carlo yaklaşımı kullanılarak literatürde bilinen ve yaygın bir şekilde kullanılan uyum iyiliği test yöntemleriyle I. tip hata ve testin gücü bakımından karşılaştırılarak, önerilen test yönteminin performansı ortaya koyulmaya çalışılmıştır. Ayrıca önerilen bağımsız ki-kare uyum iyiliği testi beş farklı gerçek hayat veri setine uygulanmıştır. Simülasyon çalışmasının son kısmında ise önerilen uyum iyiliği yönteminin başarısını göstermek için farklı simülasyonlar gerçekleştirilmiştir. Tüm işlemler, Intel® Core (TM) i7-7500 CPU 2.70 GHz işlemcili, 8 GB RAM özellikli bir bilgisayar ortamında ve MATLAB® R2020b yazılımı kullanılarak gerçekleştirilmiştir.

3.1. Yönsel Olabilirlik İndeksinin Değerlendirilmesi

Uyum iyiliği testlerinin güvenilirlik derecesi için bir ölçüt olması amacıyla önerilen olabilirlik indeksinin, doğruluğunun ve geçerliliğinin belirlenmesi amaçlanmıştır. Bu amaç doğrultusunda simülasyon çalışmasının bu kısmında, istatistikte en çok bilinen ve yaygın bir şekilde kullanılan merkezi limit teoremi ve büyük sayılar kanunundan faydalanılmıştır.

3.1.1. Merkezi Limit Teoremine Göre Olabilirlik İndeksinin Değerlendirilmesi

Merkezi limit teoremi, istatistikteki en temel ve en önemli teoremlerden biridir. Bu teoreme göre, dağılımı bilinen herhangi bir kitleden elde edilen örneklemin ortalamasının dağılımı, örnek boyutu yeteri derecede büyükse $(n > 25)$ yaklaşık olarak normal dağılır (Rosenblatt, 1956; Marques de Sá, 2007). Bununla birlikte, örneklem boyutu ikiden başlayarak birer birer arttığında örneklemin aritmetik ortalamasının dağılımının giderek normal dağılıma yaklaşması beklenir. Dağılımın hangi oranda normale yaklaştığını belirlemek için boyutu n olan örneklemden m tane üretilerek her birinden elde edilen m tane

aritmetik ortalama değerinin oluşturduğu yeni örneklemin normal dağılıma hangi oranda yaklaştığının belirlenmesi uygun bir yaklaşım olacaktır. Bu amaçla, her bir örneklem (x_j),

$$x_1 = \{x_{11}, x_{12}, \dots, x_{1n}\}$$
$$x_2 = \{x_{21}, x_{22}, \dots, x_{2n}\}$$
$$\vdots$$
$$x_m = \{x_{m1}, x_{m2}, \dots, x_{mn}\}$$

biçiminde tanımlanmaktadır. Bu örneklemlerin aritmetik ortalaması,

$$\bar{x_j} = \frac{1}{n}\sum_{i=1}^{n} x_{ji} \qquad (j = 1, 2, \dots, m) \tag{40}$$

şeklinde hesaplanmaktadır. Örneklemlerin aritmetik ortalamalarını oluşturduğu küme ise,

$$\Omega = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m\} \tag{41}$$

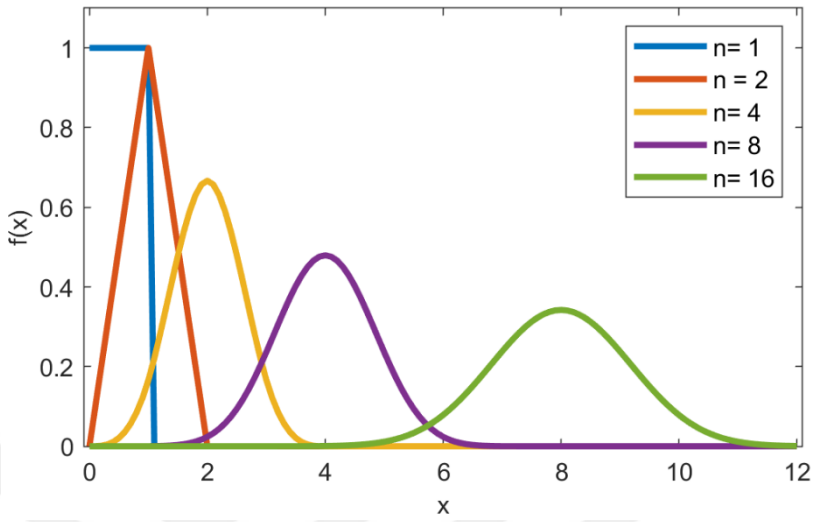
biçiminde verilebilir. Bu kümedeki verilerin normal dağılıma sahip olup olmadığını belirlemek için Kolmogorov-Smirnov (KS) uyum iyiliği testi uygulanmıştır ve uyum iyiliği testi için hipotezler,

$$H_0: \quad F(x) = F_0(x)$$
$$H_1: \quad F(x) \neq F_0(x)$$

biçiminde kurulmuştur ve hipotezlere göre, $F_0(x)$ fonksiyonu normal dağılım fonksiyonu olarak alınmıştır. Bu uyum iyiliği testi sonucunda tek bir p-değeri elde edileceği için p-değerinin değişimi rastgelelikten dolayı sağlıklı bir şekilde gözlemlenemeyecektir. Dolayısıyla, bu sorunu çözmek için çoklu örneklemler K kez tekrarlanmıştır ve K tane p-değeri elde edilerek yorumlamalar yapılmıştır.

İlk olarak, kitle olarak düzgün dağılım seçilmiştir $\left(X_n \sim Düzgün(a = 0, b = 1)\right)$ ve örneklem boyutu $n = 2, 3, \dots, 12$; $m = 1000$ ve $K = 10000$ olacak şekilde rastgele veriler üretilmiştir. Her biri düzgün dağılıma sahip n tane bağımsız rastgele değişkenin toplamı

olarak tanımlanan rastgele bir değişkenin dağılımı Irwin-Hall dağılımıdır. Şekil 17'de, Irwin-Hall dağılımının olasılık yoğunluk fonksiyonunun grafiği verilmiştir.



Şekil 17. Irwin-Hall dağılımının olasılık yoğunluk fonksiyonu

Irwin-Hall dağılımının ortalaması $n/2$ ve varyansı da $n/12$'dir. Bu durumda,

$$Z_n = \frac{\left(X_n - \frac{n}{2}\right)}{\sqrt{\frac{n}{12}}} \tag{42}$$

dönüşümü uygulandığında Z rastgele değişkeninin beklenen değeri 0 ve varyansı da 1 olmaktadır. Bu parametreler yardımıyla Z_n verilerinin normal dağılımdan gelip gelmediği KS uyum iyiliği testi ile sınanabilir. Elde edilen p-değerlerinin değişimini gözlemlemek amacıyla çeyreklikleri, yokluk hipotezinin $\alpha = 0.05$ anlamlılık düzeyi ile kabul yüzdeleri Tablo 5'te verilmiştir. Aynı simülasyon değerleri kullanılarak önerilen olabilirlik indeks değerleri (Π) elde edilmiş ve p-değerleri ile arasındaki korelasyon hesaplanarak sonuçları Tablo 6'da gösterilmiştir.

Tablo 5. Düzgün dağılımdan elde edilen örneklemin KS testi sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi

n	$Q_0\,(min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	$\%H_0$
2	0.0000	0.0906	0.2223	0.4325	0.9999	85.000
3	0.0000	0.1407	0.3342	0.5939	1.0000	89.694
4	0.0000	0.1824	0.4038	0.6642	1.0000	92.554
5	0.0000	0.2050	0.4368	0.7005	1.0000	93.181
6	0.0000	0.2164	0.4560	0.7163	1.0000	93.824
7	0.0000	0.2245	0.4677	0.7242	1.0000	94.109
8	0.0000	0.2320	0.4767	0.7315	1.0000	94.338
9	0.0000	0.2347	0.4791	0.7352	1.0000	94.469
10	0.0000	0.2380	0.4848	0.7406	1.0000	94.645
11	0.0000	0.2385	0.4854	0.7398	1.0000	94.654
12	0.0000	0.2418	0.4888	0.7444	1.0000	94.682

Tablo 5'te, KS testi kullanılarak elde edilen p-değerlerinin değişimini görmek için bu değerlerin çeyreklikleri verilmiştir. Ortanca değerler, n arttıkça artmasına rağmen değişim aralığı hala 1 olarak kalmaktadır. Yani, 12 tane düzgün dağılım gösteren rastgele değişken toplanmasına rağmen bazı p-değerleri hala sıfır değeri alabilmektedir. Bu ise p-değerlerinin güvenirliliğini sorgulatmaktadır.

Tablo 6. Düzgün dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri, Π ve p-değerleri arasındaki korelasyon katsayısı

n	$Q_0\ (min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	r
2	0.8626	0.9400	0.9544	0.9683	1.0000	0.693
3	0.8722	0.9479	0.9616	0.9743	0.9997	0.718
4	0.8811	0.9545	0.9673	0.9786	0.9999	0.705
5	0.8695	0.9573	0.9697	0.9803	0.9999	0.704
6	0.8710	0.9591	0.9710	0.9812	0.9999	0.700
7	0.8914	0.9601	0.9717	0.9818	0.9999	0.699
8	0.8847	0.9608	0.9722	0.9821	0.9999	0.698
9	0.8782	0.9611	0.9725	0.9823	1.0000	0.698
10	0.8908	0.9615	0.9728	0.9826	0.9999	0.698
11	0.8905	0.9617	0.9729	0.9825	1.0000	0.696
12	0.8783	0.9619	0.9732	0.9826	1.0000	0.694

Tablo 6'daki olabilirlik indeks değerlerinde ortanca değerler, n arttıkça artış göstermekte ve aynı zamanda değişim aralığı da sabit kalmaktadır. Bu sonuçlar olabilirlik indeksinin p-değeri ile benzer davranışlar sergilediğini göstermektedir. Bu durum ise, p-değerleri ile olabilirlik indeks değerleri arasındaki korelasyon katsayısının 0.7 civarında olması ile açıklanabilir.

İkinci olarak, kitle olarak üstel dağılım seçilmiştir $\left(X_n \sim Üstel(\theta = 1)\right)$ ve örneklem boyutu $n = 2, 3, \dots, 12;\ m = 1000$ ve $K = 10000$ olacak şekilde rastgele veriler üretilmiştir. n tane üstel dağılım gösteren rastgele değişkenlerin toplamı gamma dağılımına sahiptir.

Rastgele değişkenlerin her birinin üstel dağılım için parametresi $1/\theta$'dır. Bu durumda, gamma dağılımının ortalaması $n\theta \cdot \left(\frac{1}{\theta}\right) = n$ ve varyansı da $n\theta^2 \cdot \left(\frac{1}{\theta^2}\right) = n$ olur. Dolayısıyla,

$$Z_n = \frac{(X_n - n)}{\sqrt{n}} \tag{43}$$

dönüşümü uygulandığında Z rastgele değişkeninin beklenen değeri 0 ve varyansı da 1 olmaktadır. Bu parametreler yardımıyla Z_n verilerinin normal dağılıma sahip olup olmadığı KS uyum iyiliği testi ile kolaylıkla sınanabilir. Elde edilen p-değerlerinin çeyreklikleri ve yokluk hipotezinin $\alpha = 0.05$ anlamlılık düzeyi ile kabul yüzdeleri Tablo 7'de verilmiştir.

Tablo 7. Üstel dağılımdan elde edilen örneklemin KS testin sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi

n	$Q_0\,(min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	$\%H_0$
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.000
3	0.0000	0.0000	0.0000	0.0000	0.0073	0.000
4	0.0000	0.0000	0.0000	0.0004	0.0723	0.005
5	0.0000	0.0000	0.0002	0.0020	0.2065	0.847
6	0.0000	0.0000	0.0006	0.0060	0.4104	4.527
7	0.0000	0.0001	0.0015	0.0129	0.5465	9.842
8	0.0000	0.0003	0.0032	0.0232	0.6474	15.228
9	0.0000	0.0005	0.0055	0.0362	0.8183	20.581
10	0.0000	0.0009	0.0088	0.0516	0.8413	25.467
11	0.0000	0.0014	0.0126	0.0686	0.8477	29.978
12	0.0000	0.0021	0.0170	0.0859	0.8646	33.650

Tablo 7'de ortanca değerler, n arttıkça artmasına paralel olarak değişim aralığı da artmaktadır. Yani, 12 tane üstel dağılım gösteren rastgele değişkenin toplanmasına rağmen bazı p-değerleri hala sıfır değeri alabilmektedir. Bu ise, p-değerlerinin güvenirliliğini sorgulatmaktadır.

Yukarıda verilen simülasyon çalışmasında elde edilen veriler için önerilen yöntemle Π değerleri elde edilmiş ve p-değerleri ile arasındaki korelasyon hesaplanarak sonuçlar Tablo 8'de gösterilmiştir.

Tablo 8. Üstel dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri, Π ve p-değeri arasındaki korelasyon katsayısı

n	$Q_0\,(min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	r
2	0.6823	0.7594	0.7743	0.7894	0.8678	0.530
3	0.7123	0.8030	0.8181	0.8331	0.9207	0.493
4	0.7472	0.8289	0.8438	0.8587	0.9447	0.527
5	0.7542	0.8457	0.8608	0.8758	0.9596	0.570
6	0.7732	0.8582	0.8731	0.8881	0.9684	0.604
7	0.7838	0.8675	0.8825	0.8974	0.9794	0.634
8	0.7886	0.8754	0.8902	0.9049	0.9794	0.656
9	0.8013	0.8815	0.8964	0.9111	0.9881	0.677
10	0.8042	0.8867	0.9017	0.9165	0.9933	0.696
11	0.8105	0.8912	0.9061	0.9209	0.9933	0.709
12	0.8203	0.8950	0.9098	0.9246	0.9993	0.719

Tablo 8'deki olabilirlik indeks değerlerinin çeyreklikleri, n arttıkça artış göstermekte ve aynı zamanda değişim aralığı da sabit kalmaktadır. Bu sonuçlar olabilirlik indeks değerlerinin p-değerlerine göre daha stabil davrandığını göstermektedir. Sonuç olarak p-değerleri ile olabilirlik indeks değerleri arasındaki korelasyon katsayısının n arttıkça artış eğilimde olduğu tablodan gözlenmektedir.

3.1.2. Büyük Sayılar Kanununa Göre Olabilirlik İndeksinin Değerlendirilmesi

Büyük sayılar kanunu bir rastgele değişkenin uzun vadeli kararlılığını yansıtır. İstatistikte bir örneklemden elde edilen istatistikler, örneklemin kaynağı olan kitlenin parametrelerinin bir tahmini olarak verilmektedir. Dolayısıyla, istatistikler parametre etrafında bir saçılım göstermektedir. Büyük sayılar kanununa göre, örneklemin boyutu arttıkça istatistikler parametrelere yaklaşmaktadır. Başka bir ifadeyle, parametre etrafında saçılım gösteren istatistiklerin varyansları küçülmektedir (Heyde, 1975; Kobayashi vd., 2012). Simülasyon çalışmasının bu aşamasında, büyük sayılar kanunundan yararlanılarak önerilen yönsel olabilirlik indeksinin geçerliliği ve başarısı irdelenmiştir.

Bu aşamada kitle olarak normal dağılım seçilmiştir $\left(X_n \sim Normal(0,1)\right)$ ve örneklem boyutu $n = 10, 20, \ldots, 3000$ olarak alınmıştır. Büyük sayılar kanununa göre örneklem boyutu arttıkça istatistiklerin parametrelere yaklaşması gerekmektedir. Bu durumda da p ve Π değerlerinin örneklem boyutu ile ilişkili olarak artması beklenmektedir. $X_n \sim Normal(0,1)$ dağılımından çekilen örneklemlere KS uyum iyiliği testi uygulanmış ve elde edilen p-değerlerinin çeyreklikleri ve yokluk hipotezinin $\alpha = 0.05$ anlamlılık düzeyi ile kabul yüzdeleri Tablo 9'da verilmiştir.

Tablo 9. Normal dağılımdan elde edilen örneklemin KS testi sonucundaki p-değerlerinin çeyreklikleri ve yokluk hipotezinin kabul yüzdesi

n	$Q_0\,(min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	$\%H_0$
10	0.0024	0.2266	0.4839	0.7511	0.9997	94.7
20	0.0029	0.2641	0.5018	0.7607	0.9979	95.7
30	0.0001	0.2530	0.4947	0.7274	0.9992	95.1
40	0.0009	0.2643	0.5013	0.7419	0.9971	95.7
50	0.0014	0.2359	0.4911	0.7389	0.9996	95.8
100	0.0005	0.2648	0.4900	0.7439	0.9996	95.6
200	0.0037	0.2374	0.5118	0.7590	0.9999	95.4
300	0.0025	0.2707	0.5008	0.7718	0.9995	95.0
400	0.0007	0.2669	0.5072	0.7506	0.9984	94.9
500	0.0023	0.2275	0.5101	0.7482	0.9982	93.7
1000	0.0036	0.2775	0.5288	0.7409	0.9973	96.5
2000	0.0057	0.2505	0.4898	0.7509	0.9996	95.5
3000	0.0014	0.2446	0.4938	0.7508	0.9997	94.1

Büyük sayılar kanununa göre, örneklem boyutuna bağlı olarak p-değerlerinin ve yokluk hipotezinin kabul yüzdelerinin artması beklenmektedir. Tablo 9 incelendiğinde ise, p-değerlerinde küçük sapmalar olmasına rağmen sabit bir seyir izlediği görülmektedir. Bu durum, p-değerinin büyük sayılar kanununa cevap vermediğini göstermektedir.

Yapılan simülasyon çalışmasında elde edilen veriler için önerilen Π değerleri elde edilmiş ve p-değerleri ile arasındaki korelasyon katsayısı hesaplanarak sonuçlar Tablo 10'da verilmiştir.

Tablo 10. Normal dağılımdan elde edilen örneklemin Π değerlerinin çeyreklikleri ve Π ve p-değerleri arasındaki korelasyon katsayısı

n	$Q_0\,(min)$	Q_1	$Q_2(med)$	Q_3	$Q_4(max)$	r
10	0.1583	0.6244	0.7354	0.8309	0.9941	0.6997
20	0.3476	0.7394	0.8163	0.8797	0.9925	0.7085
30	0.4836	0.7794	0.8482	0.9020	0.9983	0.6682
40	0.5656	0.8132	0.8655	0.9155	0.9970	0.6821
50	0.6193	0.8335	0.8818	0.9212	0.9964	0.6883
100	0.7174	0.8781	0.9159	0.9439	0.9983	0.7082
200	0.8086	0.9172	0.9426	0.9635	0.9977	0.6974
300	0.8584	0.9324	0.9513	0.9689	0.9979	0.6909
400	0.8642	0.9404	0.9601	0.9741	0.9990	0.6997
500	0.8730	0.9477	0.9642	0.9763	0.9994	0.7197
1000	0.9198	0.9641	0.9745	0.9833	0.9986	0.6510
2000	0.9217	0.9740	0.9813	0.9881	0.9991	0.7011
3000	0.9515	0.9781	0.9844	0.9899	0.9995	0.7255

Tablo 10'daki Π değerlerinin değişimi incelendiğinde, örneklem boyutu arttıkça Π değerlerinin de arttığı görülmektedir. Bu durum da önerilen yönsel olabilirlik indeks değerlerinin büyük sayılar kanununa göre geçerliliğini ve başarısını ortaya koymaktadır. Ayrıca, p-değeri ve Π değerlerinin arasındaki korelasyonun yaklaşık 0.7 olması, p ve Π değerlerinin ilişkili olduğunu göstermektedir.

3.1.3. Farklı Dağılımlara Göre Olabilirlik İndeksinin Hesaplanması

Simülasyon çalışmasının bu bölümünde, simetrik dağılımlardan düzgün dağılım $\left(F_U(x)\right)$, üçgen dağılım $\left(F_T(x)\right)$, normal dağılım $\left(F_N(x)\right)$, student-t dağılımı $\left(F_{St}(x)\right)$ ve laplace dağılımı $\left(F_L(x)\right)$ kullanılarak, her birinden örneklem büyüklüğü $n = 1000$ olan rastgele değerler üretilmiştir. Bu işlem 1000 kez tekrarlanmıştır ve her bir dağılım birbiri ile test edilerek dağılımlar arasındaki Π değerleri, p-değerleri ve yokluk hipotezinin $\alpha = 0.05$ anlamlılık düzeyinde kabul yüzdeleri hesaplanmıştır.

Tablo 11'de beş dağılımdan elde edilen örneklemlerin hepsi birbiri ile ayrı ayrı ele alınarak, ortalama yönsel olabilirlik indeks değerleri hesaplanmıştır. Her satırda veya sütundaki en büyük ortalama yönsel olabilirlik indeks değerleri '*' ile gösterilmiştir.

Tablo 11. Dağılımların yönsel olabilirlik indeks değerleri (Π)

		Varsayılan Kümülatif Dağılımlar				
		$F_U(x)$	$F_T(x)$	$F_N(x)$	$F_{St}(x)$	$F_L(x)$
Gerçek Kümülatif Dağılımlar	$F_U(x)$	**0.9729***	0.8801	0.8356	0.9134	0.6649
	$F_T(x)$	0.8800	**0.9709***	0.9538	0.9601	0.7824
	$F_N(x)$	0.8364	0.9535	**0.9718***	0.9302	0.8107
	$F_{St}(x)$	0.9100	0.9633	0.9329	**0.9721***	0.7625
	$F_L(x)$	0.6754	0.7854	0.8179	0.7647	**0.9717***

Tablo 11 incelendiğinde, en büyük olabilirlik indeks değerlerinin her dağılımın kendisiyle testi sonucunda elde edildiği görülmüştür. Yani, önerilen olabilirlik indeks değerlerinin doğru ve stabil sonuçlar verdiği söylenebilir.

Tablo 12'de beş dağılımdan elde edilen örneklemlerin hepsi birbiri ile ayrı ayrı ele alınarak KS uyum iyiliği testi uygulanmıştır ve test sonucunda elde edilen ortalama p-değerleri hesaplanmıştır. Her satırda veya sütundaki en büyük ortalama p-değeri '*' ile gösterilmiştir.

Tablo 12. KS uyum iyiliği testi sonucunda elde edilen p-değerleri

		Varsayılan Kümülatif Dağılımlar				
		$F_U(x)$	$F_T(x)$	$F_N(x)$	$F_{St}(x)$	$F_L(x)$
Gerçek Kümülatif Dağılımlar	$F_U(x)$	**0.5115***	0.0070	0.0005	0.0001	0.0000
	$F_T(x)$	0.0050	**0.4813***	0.2849	0.0832	0.0000
	$F_N(x)$	0.0006	0.3008	**0.5010***	0.0853	0.0002
	$F_{St}(x)$	0.0001	0.0773	0.0801	**0.5046***	0.0000
	$F_L(x)$	0.0000	0.0000	0.0002	0.0000	**0.4960***

Tablo 12'ye göre, en büyük p-değerlerinin her dağılımın kendisiyle testi sonucunda elde edildiği görülmüştür. En büyük ortalama p-değerlerine bakıldığında, yaklaşık 0.5 civarında olduğu görülmüştür ve dolayısıyla doğru sonuçlar elde edildiği söylenebilir.

Tablo 13'te ise, Tablo 12'de verilen KS testi sonucunda elde edilen p-değerlerine göre yokluk hipotezinin kabul yüzdeleri verilmiştir. Her satır ve sütundaki en büyük kabul yüzdeleri '*' ile gösterilmiştir.

Tablo 13. KS testi sonucundaki p-değerlerine göre yokluk hipotezlerinin kabul yüzdeleri

		Varsayılan Kümülatif Dağılımlar				
		$\boldsymbol{F_U(x)}$	$\boldsymbol{F_T(x)}$	$\boldsymbol{F_N(x)}$	$\boldsymbol{F_{St}(x)}$	$\boldsymbol{F_L(x)}$
Gerçek Kümülatif Dağılımlar	$\boldsymbol{F_U(x)}$	**95.3***	0	0	0	0
	$\boldsymbol{F_T(x)}$	0.3	**94.5***	83.9	80.3	0
	$\boldsymbol{F_N(x)}$	0	88.5	**95.7***	58.4	0
	$\boldsymbol{F_{St}(x)}$	0	56.9	50.6	**94.6***	0
	$\boldsymbol{F_L(x)}$	0	0	0	0	**95.9***

Tablo 13'ten görüldüğü üzere, en büyük kabul yüzdeleri, her dağılımın kendisiyle testi sonucunda elde edilmiştir. Buna göre, testlerin doğru sonuçlar verdiği söylenebilir.

3.1.4. Deneysel Rastgele Sayıların Değiştirilmesi ile Olabilirlik İndeksinin Hesaplanması

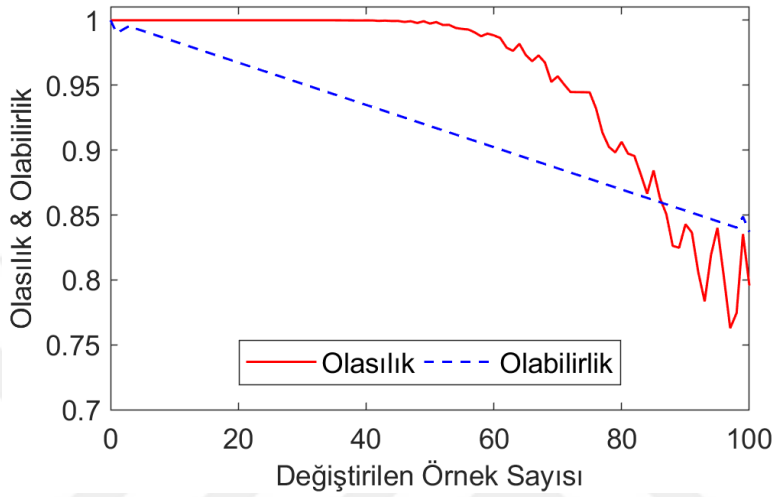
Deneysel rastgele sayılar,

$$\widetilde{X}_\iota = Shuffle\left(F^{-}\left(\widetilde{U}_\iota\right)\right), \qquad \widetilde{U}_\iota = \frac{[i-1/2]}{n}, \qquad (i = 1, 2, \dots, n) \tag{44}$$

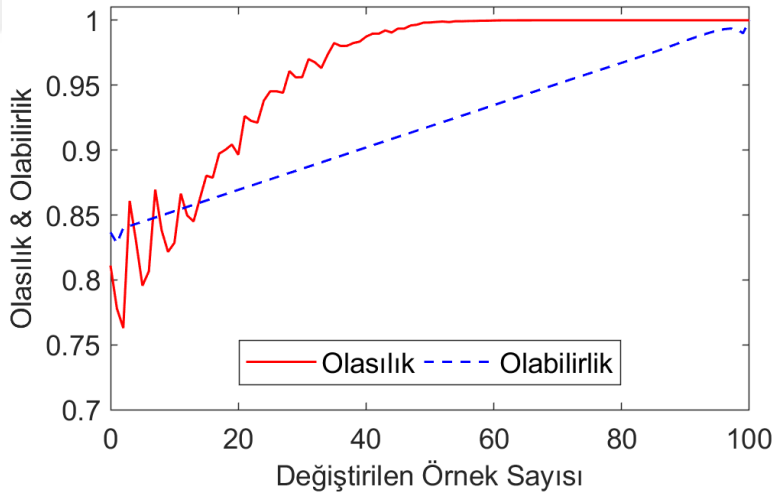
biçiminde hesaplanan ve istatistiksel açıdan hatasız olarak dağılıma uyan verilerdir.

Simülasyon çalışmasının bu bölümünde deneysel rastgele sayılardan faydalanılmıştır. İlk olarak normal $\left(X_n \sim Normal(0,1)\right)$ ve düzgün dağılımdan $\left(Y_n \sim Düzgün(0,1)\right)$, $n = 100$ büyüklüğünde iki örneklem için deneysel rastgele değerler üretilmiştir. Öncelikle, bu iki örnekleme, normal ve düzgün dağılıma göre iki ayrı KS uyum iyiliği testi uygulanarak p-değerleri ve aynı örneklemlerin Π değerleri hesaplanmıştır. Daha sonra, verilere bağlı

olarak sistemin değişimini gözlemlemek için rastgele üretilen örneklem verileri birer birer değiştirilmiştir ($swap(X_i, Y_i),\ i = 1, 2, \ldots, n$). Her bir değişimde oluşan yeni örneklemler için KS uyum iyiliği testi ve önerilen yöntem uygulanarak p ve Π değerleri hesaplanmıştır. Her aşamada p ve Π değerlerinde meydana gelen değişim Şekil 18'de gösterilmiştir.



(a)



(b)

Şekil 18. Verilere bağlı olarak p ve Π değerlerinin değişimleri (a) Normal dağılıma göre elde edilen p ve Π değerleri, (b) Düzgün dağılıma göre elde edilen p ve Π değerleri

KS testi sonucunda elde edile p ve önerilen olabilirlik indeks değerlerindeki değişimin, değiştirilen örnek sayısıyla orantılı ve uyum iyiliği testinin dağılımına bağlı olarak artması

veya azalması gerekmektedir. Ancak, Şekil 18'deki her iki grafik için de p-değerlerinin değişimine bakıldığında, p-değerlerinde iniş çıkışlar görülmektedir. Bu durum ise, p-değerinin tutarlı bir davranış sergilemediğini ortaya koymaktadır. Ayrıca, her iki grafikte de olabilirlik indeks değerlerinin lineer bir değişim gösterdiği görülmektedir. Böylece, olabilirlik indeksini temsil eden Π değerlerinin p-değerlerine göre daha tutarlı olduğunu ve daha iyi çalıştığını göstermektedir.

3.2. Bağımsız Ki-Kare Testinin I. Tip Hata ve Güç Bakımından Karşılaştırılması

Monte Carlo simülasyonu yardımıyla, önerilen bağımsız ki-kare testi (FCS) literatürde yaygın bir şekilde kullanılan ki-kare uyum iyiliği testi (BCS), Cramér-von Mises testi (CVM), Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), I. tip hata ve güç bakımından karşılaştırılmıştır. Tüm karşılaştırmalarda anlamlılık düzeyi $\alpha = 0.05$ olarak alınmıştır.

3.2.1. Simetrik Dağılımlarda I. Tip Hata ve Güç Değerlerinin Karşılaştırılması

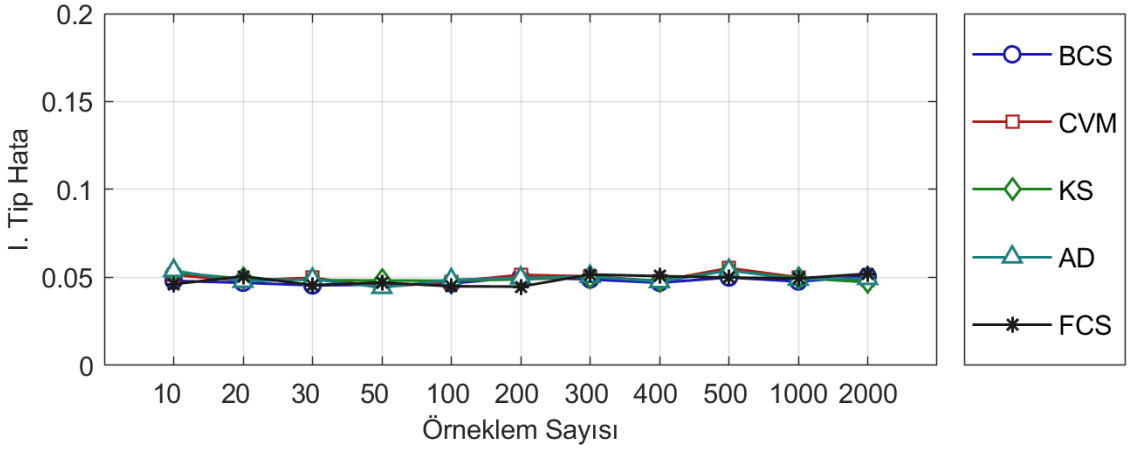
Simetrik dağılımlarda güç değerlerinin karşılaştırılmasında, her simülasyon $n = \{10, 20, 30, 50, 100, 200, 300, 400, 500, 1000, 2000\}$ için 10000 kez tekrarlanmıştır. Bu simülasyon için normal dağılım, düzgün dağılım, Laplace dağılımı, student-t dağılımı ve üstel dağılım seçilmiştir. İlk dört dağılım simetrik dağılımlardan seçilmesine rağmen, bu simülasyona simetrik olmayan üstel dağılım da dahil edilmiştir.

$(\mu = 0, \sigma^2 = 1)$ parametreli normal dağılımdan farklı büyüklüklerde çekilen örneklemlerin $(\mu = 0, \sigma^2 = 1)$ parametreli normal dağılıma uyup uymadığı seçilen beş ayrı uyum iyiliği yöntemi kullanılarak test edilmiştir. Normal dağılım için hesaplanan I. tip hatalar Tablo 14'te verilmiştir.

Tablo 14. Normal dağılım $(\mu = 0, \sigma^2 = 1)$ için uyum iyiliği testlerinin I. tip hataları

	Yöntem				
n	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0447	0.0468	0.0492	0.0493	0.0455
20	0.0453	0.0478	0.0506	0.0485	0.0456
30	0.0466	0.0459	0.0459	0.0467	0.0473
50	0.0457	0.0478	0.0498	0.0457	0.051
100	0.0524	0.0498	0.0497	0.049	0.0497
200	0.0477	0.0472	0.0464	0.0486	0.0477
300	0.0527	0.0481	0.0463	0.0487	0.0508
400	0.0481	0.0487	0.0477	0.0486	0.0476
500	0.0440	0.0519	0.0529	0.0504	0.054
1000	0.0537	0.0486	0.0489	0.0493	0.0524
2000	0.0514	0.0524	0.0524	0.0522	0.0521

Tablo 14 incelendiğinde, I. tip hataların 0.05 etrafında değiştiği görülmektedir. Bu sonuca göre, önerilen FCS yönteminin karşılaştırılan diğer uyum iyiliği testleri gibi tutarlı bir şekilde çalıştığı söylenebilir. Ayrıca, Tablo 14'teki I. tip hataların değişimi Şekil 19'da gösterilmiştir.



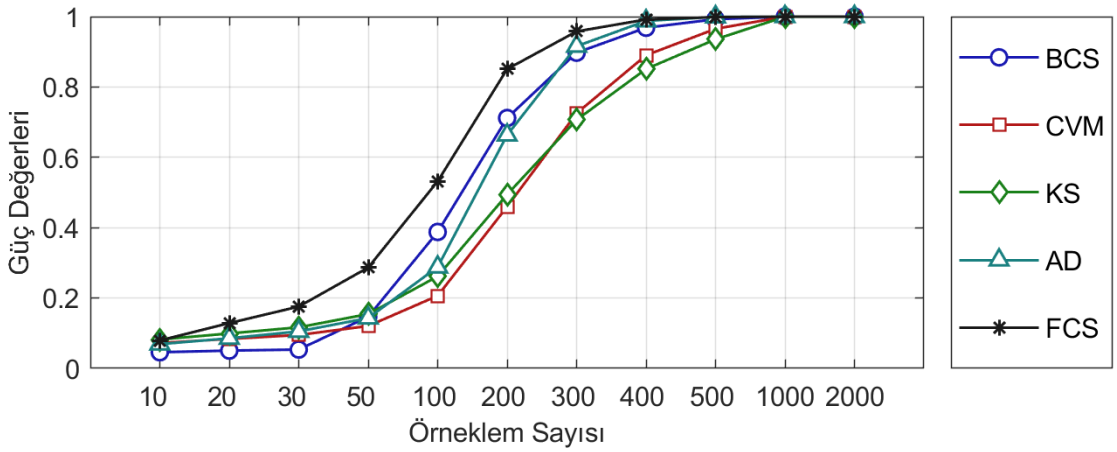
Şekil 19. Normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen I. tip hataları

$(a = -\sqrt{3}, b = \sqrt{3})$ parametreli düzgün dağılımından farklı büyüklüklerde çekilen örneklemlerin $Normal(\mu = 0, \sigma^2 = 1)$ dağılımına uyup uymadığı belirlenen beş ayrı uyum iyiliği yöntemi ile test edilmiştir. Düzgün dağılım için uyum iyiliği testlerinin güç karşılaştırmaları Tablo 15'te yer almaktadır.

Tablo 15. Düzgün dağılım için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
$\boldsymbol{n}$	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0464	0.0730	0.0835	0.0700	0.0809
20	0.0495	0.0836	0.0961	0.0868	0.1287
30	0.0486	0.0934	0.1178	0.1027	0.1775
50	0.1412	0.1196	0.1506	0.1464	0.2827
100	0.3890	0.2093	0.2617	0.2914	0.5288
200	0.7111	0.4626	0.4941	0.6638	0.8427
300	0.8982	0.7270	0.7065	0.9135	0.9613
400	0.9685	0.8907	0.8510	0.9860	0.9906
500	0.9925	0.9651	0.9344	0.9997	0.9987
1000	1	1	0.9999	1	1
2000	1	1	1	1	1

Tablo 15'teki güç değerleri incelendiğinde, örneklem sayısı arttıkça beş uyum iyiliği yönteminin de düzgün dağılım için güç değerleri artmıştır. BCS yöntemi, örneklem sayısı $n < 50$ olduğu durumlarda diğer yöntemlere göre daha düşük güce sahiptir. Örneklem sayısı arttıkça bütün yöntemler için güç değerlerinin bir olduğu gözlenmiştir. Ayrıca, Tablo 15'teki güç değerlerinin değişimi Şekil 20'de gösterilmiştir.



Şekil 20. Düzgün dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

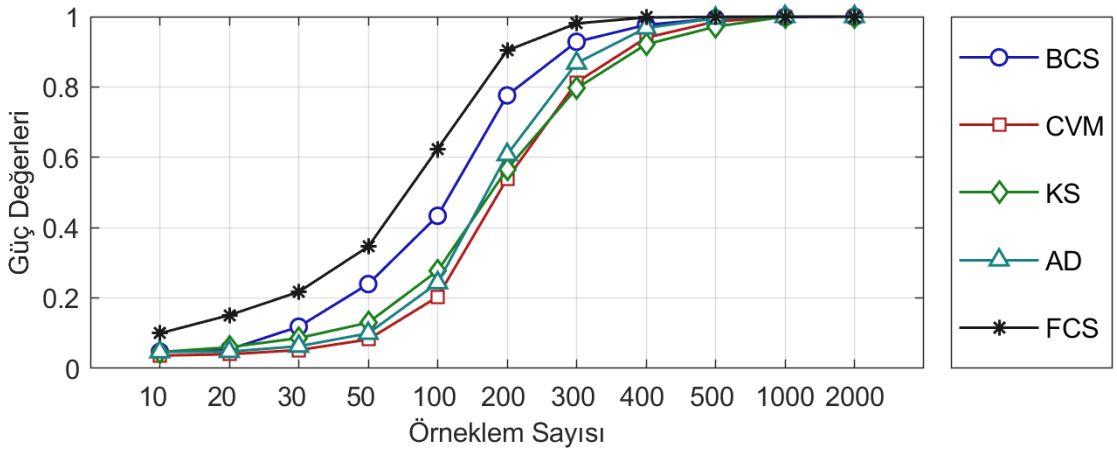
Şekil 20'deki güç değerlerine göre, CVM ve KS yöntemleri en düşük duyarlılığı gösterirken, FCS yöntemi en iyi duyarlılığa sahiptir. BCS ile AD yöntemleri benzer duyarlılıkları göstermesine rağmen, BCS yönteminin $n < 50$ durumunda düşük duyarlılığa sahip olduğu gözlenmektedir.

$Laplace\ (\mu = 0, \sigma^2 = 1)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Normal\ (\mu = 0, \sigma^2 = 1)$ dağılımına uyup uymadığı belirlenen beş ayrı uyum iyiliği yöntemi kullanılarak test edilmiştir. Laplace dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları Tablo 16'da verilmiştir.

Tablo 16. Laplace dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
$\boldsymbol{n}$	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0398	0.0320	0.0395	0.0423	0.0945
20	0.0540	0.0452	0.0629	0.0503	0.1575
30	0.1156	0.0516	0.0857	0.0605	0.2166
50	0.2434	0.0842	0.1332	0.0990	0.3471
100	0.4189	0.1967	0.2648	0.2339	0.6163
200	0.7701	0.5384	0.5715	0.6129	0.9056
300	0.9266	0.8004	0.7873	0.8597	0.9816
400	0.9785	0.9399	0.9207	0.9668	0.9967
500	0.9922	0.9838	0.9713	0.9952	0.9998
1000	0.9986	1	1	1	1
2000	0.9993	1	1	1	1

Tablo 16'daki güç değerlerine göre, örneklem sayısı arttıkça beş yöntemin de Laplace dağılımı için güç değerleri artmıştır. $n < 300$ durumunda CVM test yöntemi diğer yöntemlere göre daha düşük güce sahiptir. Önerilen FCS test yönteminin tüm örneklem boyutları için diğer yöntemlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir. Örneklem sayısı arttıkça BCS yöntemi hariç diğer yöntemler için güç değerlerinin bir olduğu gözlenmiştir. Ayrıca, Tablo 16'daki güç değerlerindeki değişimler Şekil 21'de verilmiştir.



Şekil 21. Laplace dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

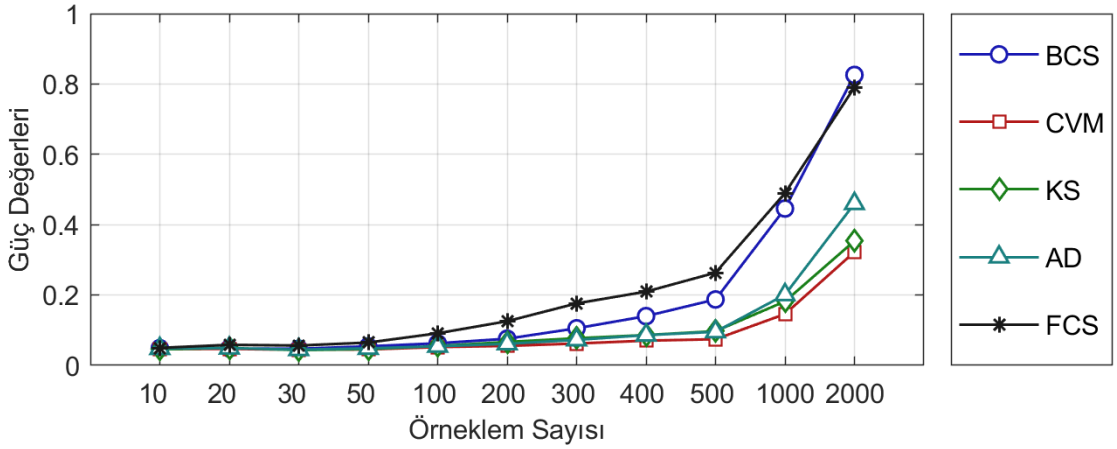
Şekil 21 incelendiğinde, uyum iyiliği testlerinin güç değerlerinin değişimine göre CVM ve KS test yöntemleri en düşük duyarlılığı gösterirken, önerilen FCS yöntemi en iyi duyarlılığı sağlamaktadır. BCS yönteminin ise AD yönteminden özellikle yüksek örneklem boyutlarında daha yüksek duyarlılığa sahip olduğu gözlenmiştir.

$Student - t\ (v = 10)$ dağılımından farklı örneklem boyutlarında çekilen örneklemlerin $Normal\ (\mu = 0, \sigma^2 = 10/8)$ dağılıma sahip olup olmadığı belirlenen beş ayrı uyum iyiliği yöntemi kullanılarak test edilmiştir. Tablo 17'de student-t dağılımı için karşılaştırılan uyum iyiliği testlerinin güç değerleri verilmiştir.

Tablo 17. Student-t dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
n	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0477	0.0428	0.0417	0.0470	0.0466
20	0.0488	0.0439	0.0443	0.0461	0.0536
30	0.0508	0.0433	0.0473	0.0450	0.0585
50	0.0555	0.0474	0.0495	0.0487	0.0668
100	0.0545	0.0507	0.0556	0.0520	0.0832
200	0.0795	0.0578	0.0664	0.0639	0.1353
300	0.1063	0.0596	0.0778	0.0721	0.1713
400	0.1390	0.0640	0.0855	0.0799	0.2135
500	0.1867	0.0734	0.0992	0.0963	0.2577
1000	0.4323	0.1364	0.1764	0.1866	0.4774
2000	0.8299	0.3354	0.3684	0.4738	0.7930

Tablo 17'deki güç değerleri incelendiğinde, örneklem sayısı arttıkça beş uyum iyiliği yönteminde de student-t dağılımı için güç değerleri artmıştır. $n > 10$ olduğu durumlarda CVM testi diğer test yöntemlerine göre daha düşük güç değerlerine sahiptir. Önerilen FCS yöntemi sonucunda genellikle tüm örneklem boyutları için en yüksek güç değerleri elde edilmiştir. Tablo 17'de sadece KS ve FCS yöntemleri için tüm örneklem boyutlarında güç değerleri artmıştır. Diğer test yöntemleri için güç değerleri monoton artan değildir. Ayrıca, Tablo 17'deki güç değerlerindeki değişimler Şekil 22'de verilmiştir.



Şekil 22. Student-t dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

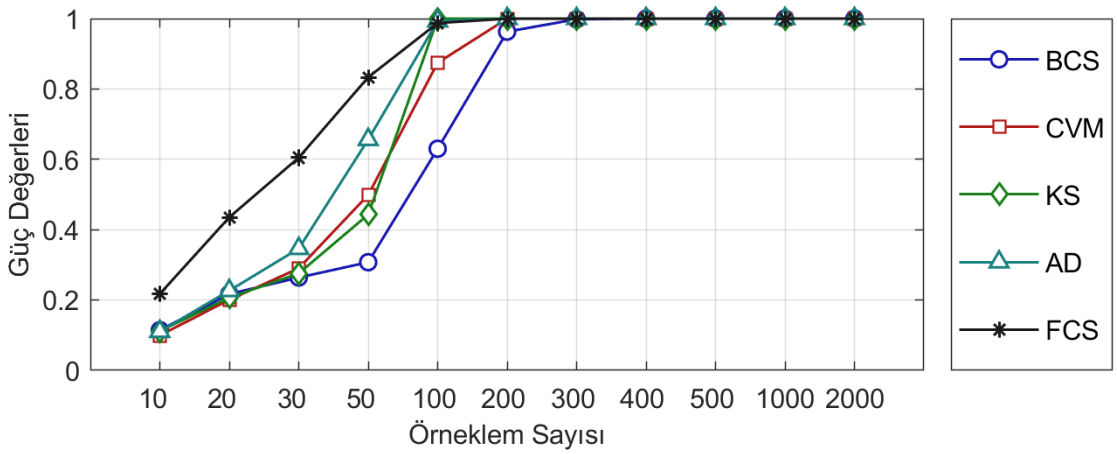
Şekil 22'deki güç değerlerinin değişimi incelendiğinde, CVM ve KS yöntemleri en düşük duyarlılığı gösterirken, FCS en iyi duyarlılığı sağlamıştır. BCS yöntemi ise AD yönteminden özellikle yüksek örneklem boyutlarında daha yüksek duyarlılığa sahiptir.

Üstel $(\lambda = 1)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Normal\ (\mu = 1, \sigma^2 = 1)$ dağılımına sahip olup olmadığı belirlenen beş ayrı uyum iyiliği test yöntemiyle incelenmiştir. Tablo 18'de üstel dağılım için uyum iyiliği testleri sonucunda elde edilen güç değerleri verilmiştir.

Tablo 18. Üstel dağılım için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
n	BCS	CVM	KS	AD	FCS
10	0.1196	0.1035	0.1156	0.1168	0.2259
20	0.2042	0.1879	0.1926	0.2115	0.4329
30	0.2684	0.2920	0.2793	0.3530	0.6062
50	0.3015	0.4829	0.4332	0.6457	0.8330
100	0.6281	0.8765	1	0.9937	0.9882
200	0.9637	1	1	1	1
300	0.9980	1	1	1	1
400	1	1	1	1	1
500	1	1	1	1	1
1000	1	1	1	1	1
2000	1	1	1	1	1

Tablo 18'e göre, örneklem sayısı arttıkça tüm yöntemlerin güç değerleri artmıştır. Önerilen FCS yöntemi $n = 100$ için KS ve AD yöntemlerine göre daha düşük güce sahiptir fakat FCS yöntemi diğer boyutlarda diğer yöntemlere göre testin gücü bakımından üstünlük göstermektedir. $n > 400$'den sonra tüm yöntemlerin güç değerleri 1 olmuştur. Ayrıca, Tablo 18'deki güç değerlerindeki değişimler Şekil 23'te verilmiştir



Şekil 23. Üstel dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

Şekil 23'teki güç değerlerinin değişimi incelendiğinde, BCS yöntemi en düşük duyarlılığı gösterirken, önerilen FCS yöntemi en iyi duyarlılığı sahiptir. AD yöntemi ise, KS ve CVM yöntemlerinden bazı örneklem boyutlarında daha yüksek duyarlılık göstermiştir.

3.2.2. Simetrik Olmayan Dağılımlarda I. Tip Hata ve Güç Değerlerinin Karşılaştırılması

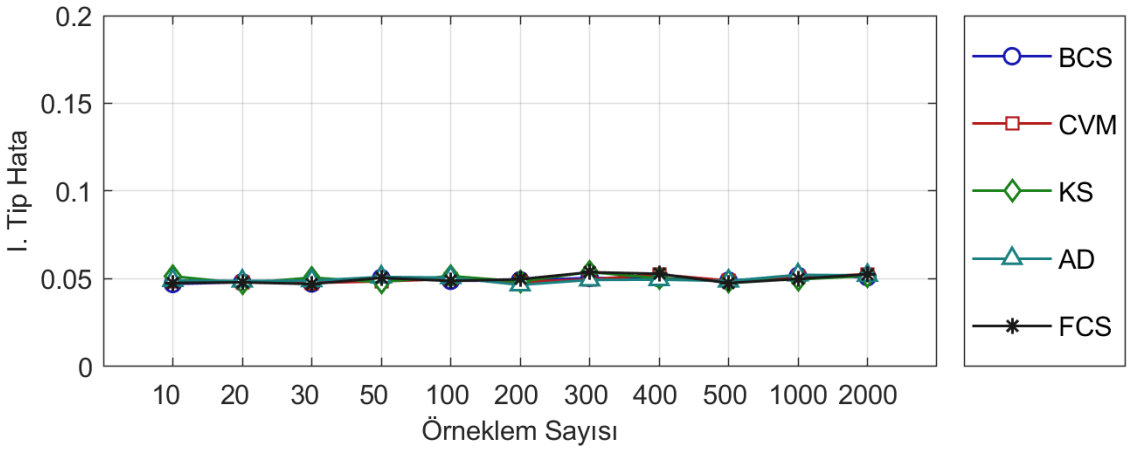
Simetrik olmayan dağılımlarda güç değerlerinin karşılaştırılmasında her simülasyon $n = \{10, 20, 30, 50, 100, 200, 300, 400, 500, 1000, 2000\}$ için 10000 kez tekrarlanmıştır. Bu simülasyon için, normal dağılım, log-normal dağılım, Weibull dağılımı, gamma dağılımı ve ki-kare dağılımı seçilmiştir. Normal dağılım simetrik dağılım olmasına rağmen diğer dört dağılım simetrik olmayan dağılımlardan seçilmiştir.

$Lognormal\ (\mu = 0.8432, \sigma^2 = 0.7147)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Lognormal\ (\mu = 0.8432, \sigma^2 = 0.7147)$ dağılımına uyup uymadığı beş ayrı uyum iyiliği yöntemi kullanılarak test edilmiştir. Log-normal dağılım için hesaplanan I. tip hatalar Tablo 19'da verilmiştir.

Tablo 19. Log-normal dağılım için uyum iyiliği testlerinin I. tip hataları

	Yöntem				
$\boldsymbol{n}$	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0494	0.0542	0.0543	0.0546	0.0489
20	0.0476	0.0467	0.0470	0.0476	0.0464
30	0.0470	0.0497	0.0500	0.0495	0.0501
50	0.0494	0.0468	0.0473	0.0461	0.0492
100	0.0512	0.0493	0.0507	0.0473	0.0519
200	0.0518	0.0494	0.0495	0.0485	0.0481
300	0.0508	0.0524	0.0549	0.0537	0.0540
400	0.0487	0.0480	0.0467	0.0492	0.0521
500	0.0506	0.0463	0.0499	0.0474	0.0514
1000	0.0514	0.0535	0.0512	0.0539	0.0519
2000	0.0462	0.0490	0.0496	0.0499	0.0453

Tablo 19 incelendiğinde, I. tip hataların 0.05 değeri etrafında dağıldığı görülmektedir. Bu sonuca göre, önerilen FCS yönteminin diğer yöntemler gibi simetrik olmayan dağılımlar için de tutarlı bir şekilde çalıştığı söylenebilir. Ayrıca, Tablo 19'daki I. tip hataların değişimi Şekil 24'te gösterilmiştir.



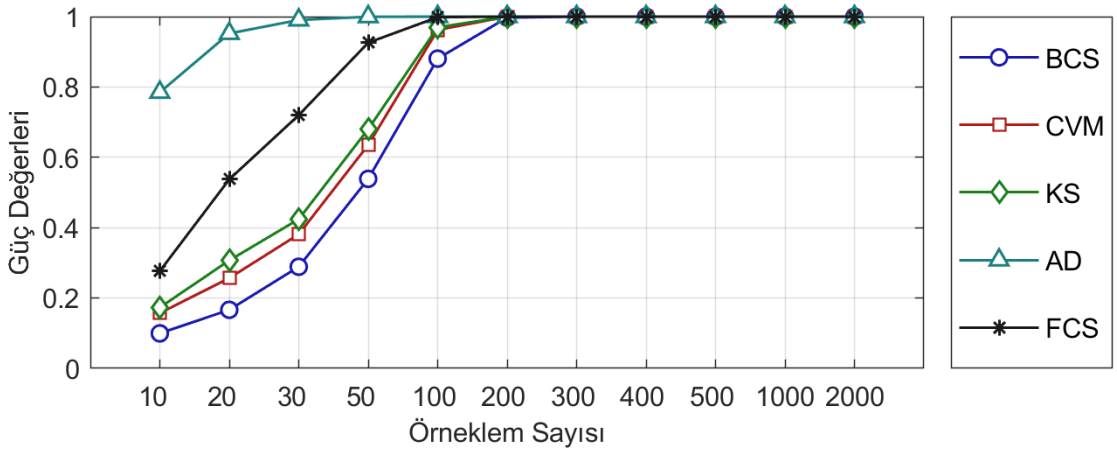
Şekil 24. Log-normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen I. tip hataları

$Normal\ (\mu = 3, \sigma^2 = 6)$ dağılımdan farklı büyüklüklerde çekilen örneklemlerin $Lognormal\ (\mu = 0.8432, \sigma^2 = 0.7147)$ dağılımına sahip olup olmadığı belirlenen beş ayrı uyum iyiliği yöntemi ile test edilmiştir. Normal dağılım için uyum iyiliği testlerinin güç karşılaştırmaları Tablo 20'de verilmiştir.

Tablo 20. Normal dağılım için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
n	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0994	0.1554	0.1660	0.7760	0.2640
20	0.1652	0.2496	0.2925	0.9520	0.5266
30	0.2919	0.3787	0.4218	0.9918	0.7259
50	0.5292	0.6360	0.6724	0.9998	0.9268
100	0.8849	0.9588	0.9711	1	0.9982
200	0.9966	1	1	1	1
300	0.9998	1	1	1	1
400	1	1	1	1	1
500	1	1	1	1	1
1000	1	1	1	1	1
2000	1	1	1	1	1

Tablo 20'deki beş uyum iyiliği yönteminin güç değerleri incelendiğinde, örneklem sayısı arttıkça beş yöntemin de normal dağılım için güç değerleri artmıştır. Tüm örneklem boyutlarında, BCS yöntemi en düşük güç değerlerine sahipken, AD yöntemi en yüksek güç değerlerini sağlamıştır. Önerilen FCS yöntemi, AD yöntemine göre daha düşük güç değerlerine sahip olmasına rağmen, karşılaştırılan diğer yöntemlere göre testin gücü bakımından üstünlük göstermektedir. Örneklem sayısı arttıkça, tüm yöntemler için güç değerleri $n > 400$ olduğu durumlarda 1 olmuştur. Ayrıca, Tablo 20'deki güç değerlerindeki değişimler Şekil 25'te verilmiştir.



Şekil 25. Normal dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

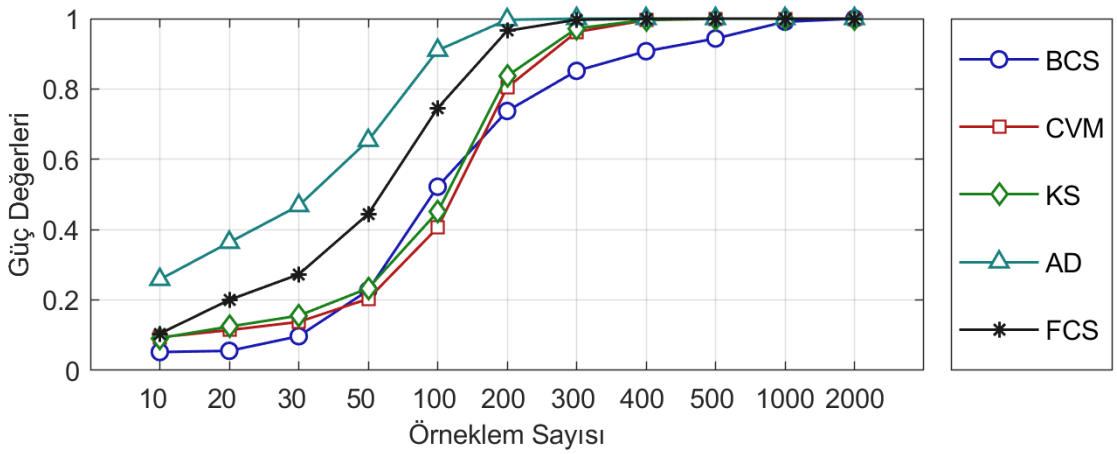
Şekil 25'teki güç değerlerindeki değişimler incelendiğinde, BCS uyum iyiliği testi en düşük duyarlılığa sahipken AD testi en yüksek duyarlılığı göstermiştir. Önerilen FCS testinin AD testine göre daha az duyarlı olmasına rağmen, karşılaştırılan diğer uyum iyiliği test yöntemlerine göre daha duyarlı olduğu gözlenmiştir.

$Weibull\ (\beta = 3.21,\ \lambda = 1.23)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Lognormal\ (\mu = 0.8432,\ \sigma^2 = 0.7147)$ dağılımına uyup uymadığı belirlenen beş ayrı uyum iyiliği yöntemi kullanılarak test edilmiştir. Weibull dağılımı için hesaplanan testin gücü değerleri Tablo 21'de verilmiştir.

Tablo 21. Weibull dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
n	BCS	CVM	KS	AD	FCS
10	0.0473	0.0968	0.0979	0.2620	0.1103
20	0.0588	0.1147	0.1216	0.3639	0.1927
30	0.0987	0.1380	0.1571	0.4690	0.2814
50	0.2201	0.1928	0.2216	0.6426	0.4364
100	0.5181	0.4002	0.4527	0.9112	0.7454
200	0.7455	0.8063	0.8410	0.9987	0.9656
300	0.8501	0.9680	0.9736	1	0.9964
400	0.9040	0.9960	0.9980	1	0.9997
500	0.9431	0.9997	1	1	1
1000	0.9908	1	1	1	1
2000	0.9998	1	1	1	1

Tablo 21'e göre, örneklem sayısı arttıkça beş yöntemin de Weibull dağılımı için güç değerleri artmıştır. Tüm örneklem boyutlarında, önerilen FCS yöntemi, AD yöntemine göre daha düşük güce sahip olmasına rağmen, diğer yöntemlere göre üstünlük göstermektedir. Örneklem sayıları arttıkça, BCS yöntemi hariç tüm yöntemler için güç değerleri 1 olmuştur. Ayrıca, Tablo 21'deki güç değerlerindeki değişimler Şekil 26'da verilmiştir.



Şekil 26. Weibull dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

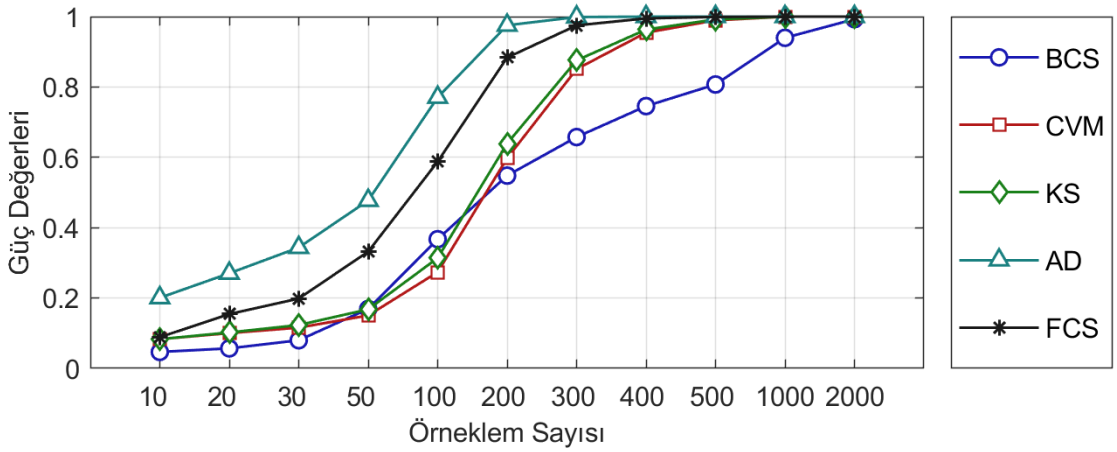
Şekil 26'daki güç değerlerinde meydana gelen değişimler incelendiğinde, BCS testi düşük ve yüksek örneklem büyüklüklerinde en düşük duyarlılığa sahipken AD testi en yüksek duyarlılığı göstermiştir. Önerilen FCS testinin, AD testine göre daha az duyarlı olmasına rağmen, diğer testlere göre daha duyarlı olduğu söylenebilir.

$Gamma$ $(k = 1.5, \theta = 2)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Lognormal$ $(\mu = 0.8432, \sigma^2 = 0.7147)$ dağılımına uyup uymadığı seçilen beş ayrı uyum iyiliği yöntemi ile test edilmiştir. Gamma dağılımı için hesaplanan güç değerleri Tablo 22'de verilmiştir.

Tablo 22. Gamma dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
$\boldsymbol{n}$	**BCS**	**CVM**	**KS**	**AD**	**FCS**
10	0.0489	0.0894	0.0878	0.2073	0.0941
20	0.0523	0.0987	0.1007	0.2711	0.1513
30	0.0851	0.1121	0.1208	0.3380	0.2109
50	0.1660	0.1483	0.1690	0.4785	0.3276
100	0.3680	0.2681	0.3114	0.7683	0.5918
200	0.5406	0.6002	0.6490	0.9721	0.8847
300	0.6537	0.8470	0.8762	0.9985	0.9757
400	0.7360	0.9541	0.9675	1	0.9953
500	0.8027	0.9909	0.9943	1	0.9991
1000	0.9362	1	1	1	1
2000	0.9918	1	1	1	1

Tablo 22'ye göre, örneklem sayısı arttıkça beş yöntemin de gamma dağılımı için güç değerleri artmaktadır. BCS yöntemi, $n = 50$ ve $n = 100$ olduğu durumlar haricinde tüm örneklem boyutlarında diğer yöntemlere göre daha düşük güç değerlerine sahiptir. Tüm örneklem boyutlarında, önerilen FCS yöntemi, AD yöntemine göre daha düşük güce sahip olmasına rağmen, karşılaştırılan diğer yöntemlere göre testin gücü bakımından üstünlük göstermektedir. Örneklem sayısı arttıkça, BCS yöntemi hariç tüm yöntemler için güç değerleri 1 olmuştur. Ayrıca, Tablo 22'deki güç değerlerinin değişimi Şekil 27'de verilmiştir.



Şekil 27. Gamma dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

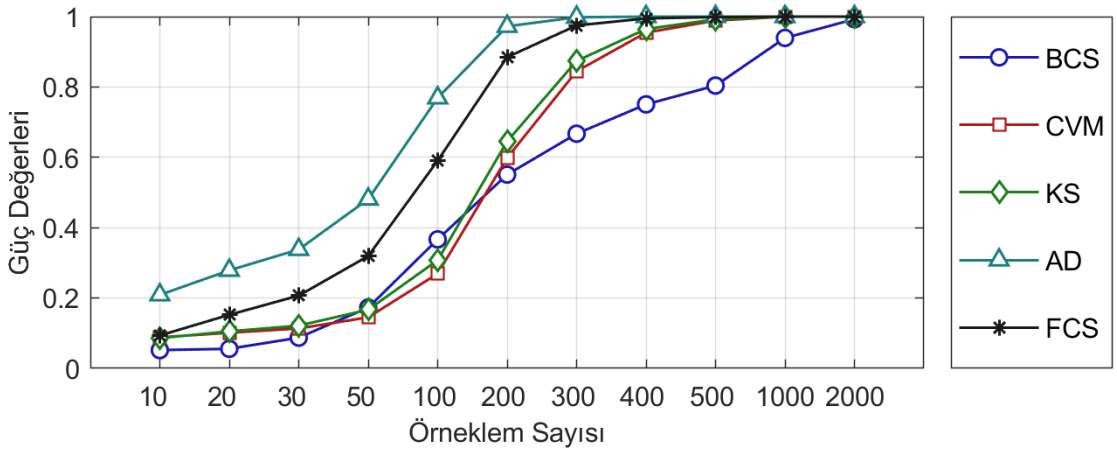
Şekil 27 incelendiğinde, BCS testi düşük ve yüksek örneklem büyüklüklerinde en düşük duyarlılığa sahipken, AD testinin en yüksek duyarlılığı sağladığı görülmektedir. Önerilen FCS testinin, AD testine göre daha düşük duyarlılığa sahip olmasına rağmen, diğer testlere göre daha duyarlı olduğu söylenebilir.

$Kikare\ (k = 3)$ dağılımından farklı büyüklüklerde çekilen örneklemlerin $Lognormal\ (\mu = 0.8432, \sigma^2 = 0.7147)$ dağılımına uyup uymadığı belirlenen beş ayrı uyum iyiliği yöntemiyle test edilmiştir. Ki-kare dağılımı için hesaplanan güç değerleri Tablo 23'te verilmiştir.

Tablo 23. Ki-kare dağılımı için uyum iyiliği testlerinin güç karşılaştırmaları

	Yöntem				
n	BCS	CVM	KS	AD	FCS
10	0.0514	0.0917	0.0901	0.2084	0.0925
20	0.0560	0.0996	0.1073	0.2742	0.1560
30	0.0832	0.1124	0.1230	0.3463	0.2089
50	0.1644	0.1447	0.1663	0.4795	0.3225
100	0.3668	0.2705	0.3076	0.7664	0.5917
200	0.5508	0.5990	0.6534	0.9769	0.8903
300	0.6650	0.8459	0.8731	0.9981	0.9766
400	0.7418	0.9550	0.9655	0.9999	0.9953
500	0.8077	0.9908	0.9938	1	0.9991
1000	0.9375	1	1	1	1
2000	0.9940	1	1	1	1

Tablo 23'teki uyum iyiliği testlerine ait güç değerleri incelendiğinde, örneklem sayısı arttıkça beş yöntemin de ki-kare dağılımı için güç değerlerinin arttığı görülmektedir. BCS yöntemi $n = 50$ ve $n = 100$ olduğu durumlar haricinde tüm örneklem boyutlarında diğer yöntemlere göre daha düşük güç değerlerine sahiptir. Tüm örneklem boyutlarında, önerilen FCS yöntemi, AD yöntemine göre daha düşük güce sahip olmasına rağmen, diğer karşılaştırılan yöntemlere göre testin gücü bakımından üstünlük göstermektedir. Örneklem sayısı arttıkça, BCS yöntemi hariç karşılaştırılan tüm yöntemler için güç değerleri 1 olmuştur. Ayrıca, Tablo 23'teki güç değerlerinin değişimi Şekil 28'de gösterilmiştir.



Şekil 28. Ki-kare dağılımdan üretilmiş örneklemlerin uyum iyiliği testleri sonucunda elde edilen güç değerleri

Şekil 28'deki güç değerlerindeki değişimler incelendiğinde, düşük ve yüksek örneklem büyüklüklerinde BCS yönteminin en düşük duyarlılığa, AD yönteminin ise en yüksek duyarlılığa sahip olduğu görülmektedir. Önerilen FCS yönteminin, AD yöntemine göre daha az duyarlı olmasına rağmen, diğer testlere göre daha duyarlı olduğu gözlenmektedir.

3.3. Bağımsız Ki-Kare Testinin Gerçek Veri Seti Uygulamaları

Çalışmanın bu bölümünde, gerçek hayatta karşılaşılan sorunlara ait veri setleri kullanılmış ve bu verilerin davranışını karakterize eden dağılımları sınamak için verilere önerilen bağımsız ki-kare uyum iyiliği testi (FCS) ile literatürde yaygın bir şekilde kullanılan ki-kare uyum iyiliği testi (BCS), Cramér-von Mises testi (CVM), Kolmogorov-Smirnov uyum iyiliği testi (KS), Anderson-Darling uyum iyiliği testi (AD) uygulanmış ve sonuçları karşılaştırılmıştır.

3.3.1. Küresel Isınma Veri Seti

Küresel ısınma, atmosferik sıcaklık ve atmosferdeki karbondioksit miktarı gibi birbiriyle etkileşime giren iki elemandan etkilenmektedir. Bu çalışmada, küresel ısınmaya ilişkin CO_2 verileri kullanılmıştır. Kullanılan veri seti http://scrippsco2.ucsd.edu/data/atmosperic-co2.html web sitesinden alınmıştır. 1974-2004

yılları arasındaki Alaska'nın Point Barrow noktasına ait CO_2 miktarları kaydedilmiştir (Ramachandran ve Tsokos, 2015). Rastgele çekilen 31 gözlemin yer aldığı CO_2 miktarlarını içeren veri seti Tablo 24'te verilmiştir.

Tablo 24. Küresel ısınma veri seti

1974-2004 yılları arasındaki Alaska'nın Point Barrow noktasına ait CO_2 miktarları						
332.435	333.273	334.148	335.300	336.930	338.418	340.429
340.129	342.483	344.064	345.644	346.980	348.588	350.051
352.929	355.392	356.362	357.396	357.873	358.280	359.785
362.095	364.064	365.086	367.356	369.746	370.822	372.376
374.533	376.9970	378.371				

1974-2004 yılları arasındaki 31 yıllık CO_2 verilerinin gamma dağılımından geldiği düşünülmektedir ve hipotezler,

$$H_0: CO_2 \; verileri \; gamma \; dağılımından \; gelmektedir.$$

$$H_1: CO_2 \; verileri \; gamma \; dağılımdan \; gelmemektedir$$

biçiminde yazılır. Bu yokluk hipotezini sınamak için veri setine Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), Cramér-von Mises testi (CVM), ki-kare testi (BCS) ve önerilen bağımsız ki-kare testi (FCS) $\alpha = 0.05$ anlamlılık düzeyinde uygulanmıştır. Uyum iyiliği testlerinden elde edilen sonuçlar, p-değerleri ve test istatistikleri Tablo 25'te verilmiştir.

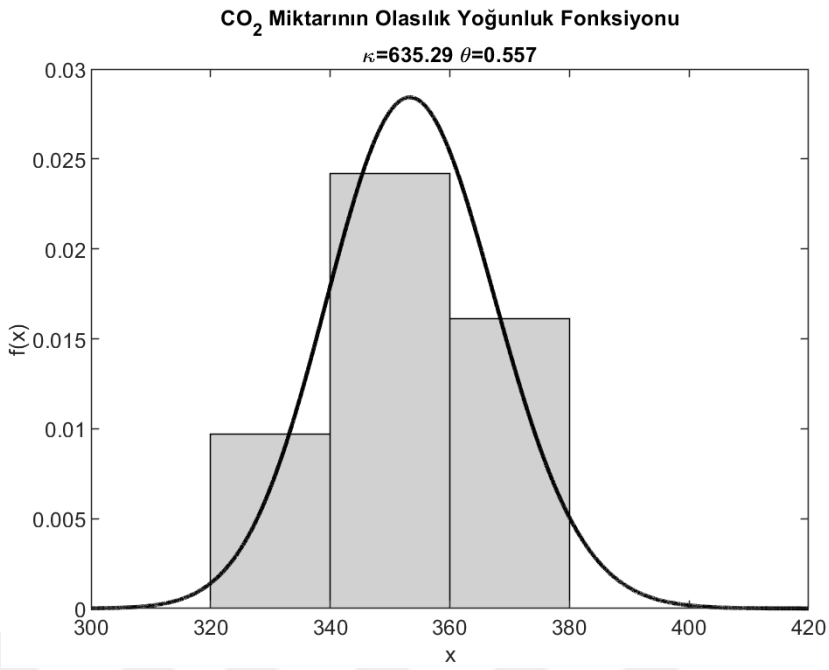
Tablo 25. Küresel ısınma veri setinin uyum iyiliği testi sonuçları

		Test istatistiği	p-değeri	Hipotez
Ki-kare (BCS)	:	0.9582	0.8110	Kabul
Cramér-von Mises (CVM)	:	0.0496	0.8812	Kabul
Kolmogorov-Smirnov (KS)	:	0.0884	0.9510	Kabul
Anderson-Darling (AD)	:	0.3632	0.8835	Kabul
Bağımsız ki-kare (FCS)	:	1.2701	0.7351	Kabul

Küresel ısınma veri setine uygulanan tüm uyum iyiliği testleri için $p > \alpha = 0.05$ olduğundan yokluk hipotezi reddedilemez. Yani, uyum iyiliği testleri sonucunda CO_2 miktarına ilişkin verilerin gamma dağılımından geldiği söylenebilir. Gamma dağılımının parametreleri için en çok olabilirlik tahmin edicileri kullanılmıştır ve böylece $\hat{k} = 635.29$ ve $\hat{\theta} = 0.557$ olarak bulunmuştur. Uyum iyiliği testi ve parametre tahmini sonucunda Alaska'nın Point Barrow noktasına ait CO_2 miktarı verilerinin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{x^{635.29-1}}{(0.557)^{635.29}\,\Gamma(635.29)} \exp\left(-\frac{x}{0.557}\right) \quad x > 0 \tag{45}$$

biçiminde yazılabilir. Belirlenen parametrelere bağlı olarak CO_2 miktarı verilerinin olasılık yoğunluk fonksiyonunun grafiği Şekil 29'da verilmiştir.



Şekil 29. CO_2 miktarı verilerinin olasılık yoğunluk fonksiyonu

3.3.2. Katrina Kasırgası Veri Seti

Katrina kasırgası son 100 yılda Amerika Birleşik Devletleri'nde gerçekleşen en yıkıcı kasırgalardan biridir. Atlantik merkezli bu kasırga, 23-31 Ağustos 2005 tarihleri arasında 9 gün sürmüştür. Katrina kasırgasının en önemli değişkenlerinden biri basınç rüzgâr hızıdır. Bu veri seti uygulamasında, Katrina kasırgasının basınç rüzgâr hızının karakteristiğini ortaya koyacak olan hangi dağılımdan geldiği belirlenmeye çalışılmıştır. Katrina kasırgasına ait basınç rüzgâr hızı verileri ttp://weather.unisys.com/hurricane/atlantic/2005H/KATRINA web sitesinden alınmıştır. Bu veri setinde basınç rüzgâr hızına ait 63 gözlem bulunmaktadır ve maksimum basınç rüzgâr hızının saatte 150 mph olarak ölçülmüştür. Katrina kasırgasına ait basınç rüzgâr hızı gözlemlerinin yer aldığı veri seti Tablo 26'da verilmiştir.

Tablo 26. Katrina kasırgasına ait veri seti

No	Rüzgâr Hızı	No	Rüzgâr Hızı	No	Rüzgâr Hızı	No	Rüzgâr Hızı	No	Rüzgâr Hızı
1	30	**14**	50	**27**	85	**40**	125	**53**	110
2	30	**15**	55	**28**	85	**41**	125	**54**	90
3	30	**16**	60	**29**	85	**42**	140	**55**	80
4	30	**17**	65	**30**	85	**43**	150	**56**	65
5	30	**18**	70	**31**	90	**44**	150	**57**	55
6	30	**19**	70	**32**	95	**45**	145	**58**	50
7	35	**20**	65	**33**	100	**46**	140	**59**	45
8	40	**21**	60	**34**	100	**47**	140	**60**	30
9	40	**22**	60	**35**	100	**48**	140	**61**	25
10	50	**23**	65	**36**	100	**49**	135	**62**	25
11	45	**24**	65	**37**	100	**50**	130	**63**	15
12	45	**25**	65	**38**	100	**51**	125		
13	45	**26**	70	**39**	100	**52**	115		

Rüzgâr hızına ilişkin verilerin Weibull dağılımı gösterdiği düşünülmektedir ve hipotezler,

$$H_0: Rüzgâr\ hızı\ verileri\ Weibull\ dağılımından\ gelmektedir.$$

$$H_1: Rüzgâr\ hızı\ verileri\ Weibull\ dağılımdan\ gelmemektedir.$$

biçiminde yazılır. Bu yokluk hipotezini sınamak için veri setine Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), Cramér-von Mises testi (CVM), ki-kare testi (BCS) ve önerilen bağımsız ki-kare testi (FCS) $\alpha = 0.05$ anlamlılık düzeyinde uygulanmıştır. Uyum iyiliği testlerinden elde edilen sonuçlar, p-değerleri ve test istatistikleri Tablo 27'de verilmiştir.

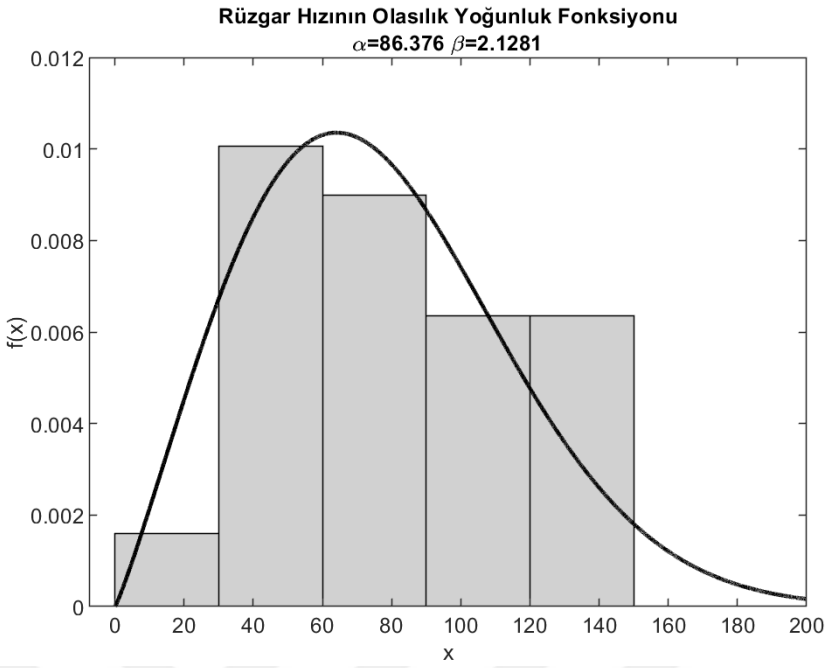
Tablo 27. Katrina kasırgası veri setinin uyum iyiliği testi sonuçları

		Test istatistiği	p-değeri	Hipotez
Ki-kare (BCS)	:	3.0141	0.8071	Kabul
Cramér-von Mises (CVM)	:	0.0734	0.7323	Kabul
Kolmogorov-Smirnov (KS)	:	0.0792	0.7948	Kabul
Anderson-Darling (AD)	:	0.5949	0.8628	Kabul
Bağımsız ki-kare (FCS)	:	1.1040	0.7121	Kabul

Katrina kasırgası veri setine uygulanan tüm uyum iyiliği testleri için $p > \alpha = 0.05$ olduğundan yokluk hipotezi reddedilemez. Yani, uyum iyiliği testleri sonucunda Katrina kasırgasına ait basınç rüzgâr hızı verileri Weibull dağılımından gelmektedir. Weibull dağılımının parametreleri için en çok olabilirlik tahmin edicileri kullanılmıştır ve bunun sonucunda $\hat{\beta} = 2.1281$ ve $\hat{\lambda} = 86.376$ olarak bulunmuştur. Böylece, Katrina kasırgasına ait basınç rüzgâr hızı verilerinin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{86.376}{2.1281}\left(\frac{x}{2.1281}\right)^{85.376} \exp\left(-\frac{x}{2.1281}\right)^{86.376} \tag{46}$$

biçiminde yazılabilir. Belirlenen parametrelere bağlı olarak basınç rüzgâr hızı verilerinin olasılık yoğunluk fonksiyonunun grafiği Şekil 30'da verilmiştir.



Şekil 30. Basınç rüzgâr hızı verilerinin olasılık yoğunluk fonksiyonu

3.3.3. Yağış Miktarı Veri Seti

Bu uygulamada, ABD'nin güney bölgesine ait 1975-2007 yılları arasındaki yıllık ortalama yağış verileri yer almaktadır. Veri setindeki 33 ölçüm kullanılarak, ABD'nin güney bölgesinin yıllık ortalama yağışının davranışını olasılıksal olarak karakterize eden olasılık yoğunluk fonksiyonu tanımlanmaya çalışılmıştır. Yağış miktarının olasılık yoğunluk fonksiyonunun elde edilmesiyle bölgede beklenen yağmur miktarı ve gerçek yıllık yağış miktarının güven sınırları belirlenebilmektedir (Ramachandran ve Tsokos, 2015). 1975-2007 yılları arasındaki yıllık ortalama yağış verilerini içeren veri seti Tablo 28'de verilmiştir.

Tablo 28. Yağış miktarı veri seti

Yıl	Yağış	Yıl	Yağış	Yıl	Yağış	Yıl	Yağış
1975	3.957	**1984**	3.563	**1993**	4.175	**2002**	5.037
1976	4.031	**1985**	3.592	**1994**	4.889	**2003**	4.633
1977	3.918	**1986**	4.307	**1995**	5.468	**2004**	5.219
1978	4.299	**1987**	4.465	**1996**	3.668	**2005**	5.220
1979	4.942	**1988**	4.487	**1997**	5.029	**2006**	3.526
1980	3.921	**1989**	3.612	**1998**	4.730	**2007**	3.211
1981	3.680	**1990**	3.322	**1999**	4.103		
1982	5.224	**1991**	4.463	**2000**	2.737		
1983	5.639	**1992**	4.514	**2001**	4.104		

Yıllık ortalama yağış miktarı verilerinin normal dağılıma sahip olduğu düşünülmektedir ve hipotezler,

$$H_0: Yağış\ miktarı\ verileri\ normal\ dağılımdan\ gelmektedir.$$

$$H_1: Yağış\ miktarı\ verileri\ normal\ dağılımdan\ gelmemektedir$$

biçiminde yazılabilir. Bu yokluk hipotezini sınamak için veri setine Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), Cramér-von Mises testi (CVM), ki-kare testi (BCS) ve önerilen bağımsız ki-kare testi (FCS) $\alpha = 0.05$ anlamlılık düzeyinde uygulanmıştır. Uyum iyiliği testlerinden elde edilen sonuçlar, p-değerleri ve test istatistikleri Tablo 29'da verilmiştir.

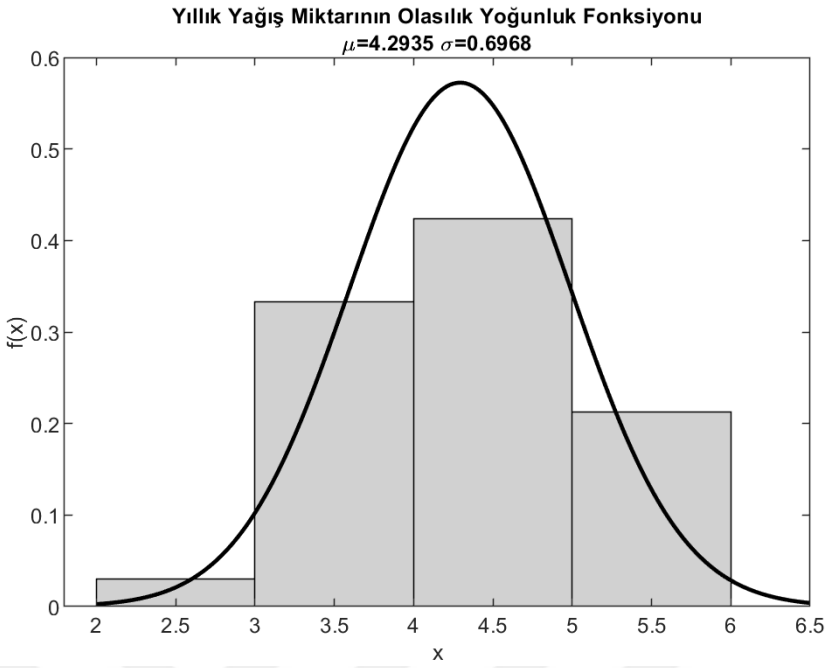
Tablo 29. Yağış miktarı veri setinin uyum iyiliği testi sonuçları

		Test istatistiği	p-değeri	Hipotez
Ki-kare (BCS)	:	1.0979	0.7776	Kabul
Cramer-von Mises (CVM)	:	0.0329	0.9667	Kabul
Kolmogorov-Smirnov (KS)	:	0.0834	0.9611	Kabul
Anderson-Darling (AD)	:	0.2259	0.9816	Kabul
Bağımsız ki-kare (FCS)	:	0.5614	0.6224	Kabul

ABD'nin güney bölgesine ait yağış miktarı veri setine uygulanan tüm uyum iyiliği testleri için $p > \alpha = 0.05$ olduğundan yokluk hipotezi reddedilemez. Yani, uyum iyiliği testleri sonucunda ABD'nin güney bölgesine ait yağış miktarının normal dağılımdan geldiği söylenebilir. Normal dağılımın parametreleri için en çok olabilirlik tahmin edicileri kullanılmıştır ve bunun sonucunda $\hat{\mu} = 4.2935$ ve $\hat{\sigma} = 0.6968$ olarak bulunmuştur. Böylece, yıllık ortalama yağış miktarı verilerinin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{0.6968\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - 4.2935}{0.6968}\right)^2\right) \tag{47}$$

biçiminde yazılabilir. Veri setine ait yıllık ortalama yağışın gerçek değerler etrafındaki güven sınırları da hesaplanmıştır ve bunun sonucunda %95 güvenle yıllık ortalama yağışın 4.0579 ile 4.5167 değerleri arasında olacağı söylenebilir. Belirlenen parametrelere bağlı olarak yağış miktarı verilerinin olasılık yoğunluk fonksiyonunun grafiği Şekil 31'de verilmiştir.



Şekil 31. Yağış miktarı verilerinin olasılık yoğunluk fonksiyonu

3.3.4. Akciğer Kanseri Veri Seti

Akciğer kanseri veri seti, erkek ve kadın olmak üzere 500 hastanın kötü huylu tümör boyutlarını (mm) içermektedir (Ramachandran ve Tsokos, 2015). Erkek hastalardan rastgele 60 tanesi seçilerek kanserli tümör boyutları incelenmiştir. Bu uygulamada, kanserli tümör boyutlarının karakteristiğini belirleyen olasılık yoğunluk fonksiyonu belirlenmeye çalışılmıştır. Kötü huylu tümör boyutlarına ait rastgele seçilen 60 gözlem Tablo 30'da verilmiştir.

Tablo 30. Akciğer kanseri veri seti

Kötü Huylu Tümör Boyutları (mm)									
0.94	0.27	0.31	0.33	0.56	0.93	0.66	0.60	0.95	0.32
0.97	0.37	0.65	0.32	0.45	0.15	0.50	0.83	1.10	0.11
0.11	0.29	0.94	0.48	0.50	0.16	0.05	0.96	0.20	0.25
0.17	0.08	0.31	0.15	0.40	0.13	0.55	0.55	0.03	0.16
0.24	0.25	0.88	0.30	0.35	0.13	0.23	0.08	0.51	0.42
0.14	0.12	0.56	0.85	0.30	0.15	0.15	0.52	0.24	0.82

Kötü huylu tümör boyutlarına ilişkin verilerin Weibull dağılımından geldiği düşünülmektedir ve hipotezler,

$$H_0: Kötü\ huylu\ tümör\ boyutları\ Weibull\ dağılımdan\ gelmektedir.$$

$$H_1: Kötü\ huylu\ tümör\ boyutları\ Weibull\ dağılımdan\ gelmemektedir.$$

biçiminde yazılır. Bu yokluk hipotezini sınamak için veri setine Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), Cramér-von Mises testi (CVM), ki-kare testi (BCS) ve önerilen bağımsız ki-kare testi (FCS) $\alpha = 0.05$ anlamlılık düzeyinde uygulanmıştır. Uyum iyiliği testlerinden elde edilen sonuçlar, p-değerleri ve test istatistikleri Tablo 31'de verilmiştir.

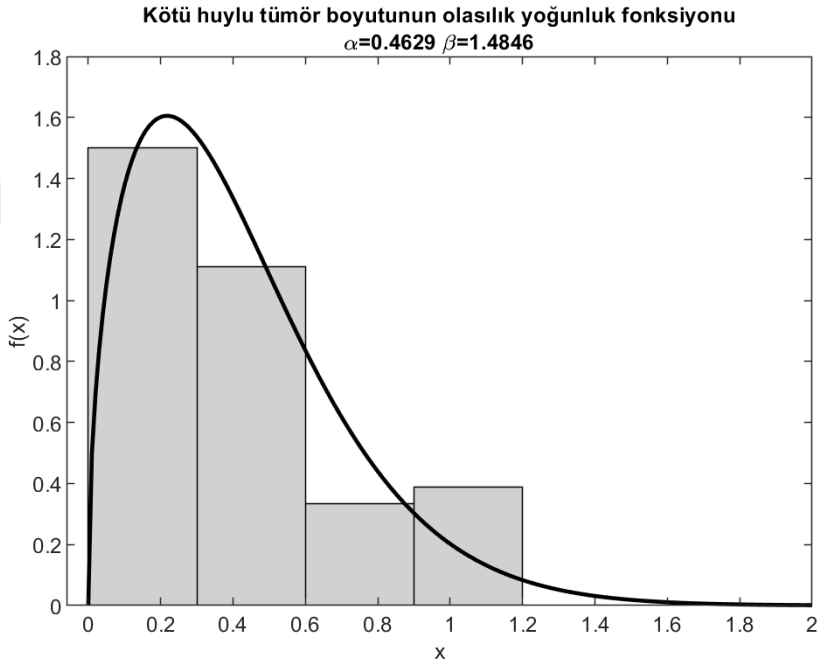
Tablo 31. Akciğer kanser veri setinin uyum iyiliği testi sonuçları

		Test istatistiği	***p*-değeri**	**Hipotez**
Ki-kare (BCS)	:	4.7954	0.5703	Kabul
Cramér-von Mises (CVM)	:	0.0784	0.7027	Kabul
Kolmogorov-Smirnov (KS)	:	0.0867	0.7245	Kabul
Anderson-Darling (AD)	:	0.5877	0.6589	Kabul
Bağımsız ki-kare (FCS)	:	1.5267	0.7669	Kabul

Akciğer kanseri veri setine uygulanan tüm uyum iyiliği testleri için $p > \alpha = 0.05$ olduğundan yokluk hipotezi reddedilemez. Yani, uyum iyiliği testleri sonucunda erkek hastalara ait kötü huylu tümör boyutları Weibull dağılımı göstermektedir. Weibull dağılımının parametreleri için en çok olabilirlik tahmin edicileri kullanılmıştır ve bunun sonucunda $\hat{\lambda} = 0.4629$ ve $\hat{\beta} = 1.4846$ olarak bulunmuştur. Belirlenen parametrelere bağlı olarak akciğer kanserine ait kötü huylu tümör boyutlarının olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1.4846}{0.4629}\left(\frac{x}{0.4629}\right)^{0.4846} \exp\left(-\frac{x}{0.4629}\right)^{1.4846} \tag{48}$$

biçiminde yazılır. Erkek hastalara ait kötü huylu tümör boyutlarını içeren verilerin olasılık yoğunluk fonksiyonunun grafiği Şekil 32'de verilmiştir.



Şekil 32. Kötü huylu tümör verilerinin olasılık yoğunluk fonksiyonu

3.3.5. Uyku Süresi Veri Seti

Öğrencilere iki haftalık süre boyunca uyku davranışlarının çeşitliği hakkında anket uygulanmıştır (Onyper vd., 2012). Bu anketle 100 öğrencinin günlük ortalama uyku süresi (h) gözlenmiştir. Ortalama uyku sürelerine ait veri seti Tablo 32'de verilmiştir.

Tablo 32. Uyku süresi veri seti

Günlük Ortalama Uyku Süresi (h)									
5.75	6.51	7.21	7.50	7.64	7.86	8.00	8.29	8.71	9.05
5.93	6.57	7.25	7.52	7.67	7.86	8.04	8.29	8.71	9.15
5.96	6.69	7.26	7.53	7.71	7.87	8.07	8.43	8.75	9.19
6.00	6.78	7.30	7.54	7.73	7.88	8.11	8.49	8.79	9.25
6.19	6.92	7.32	7.57	7.75	7.89	8.17	8.49	8.81	9.32
6.36	7.04	7.39	7.60	7.79	7.93	8.18	8.52	8.82	9.80
6.36	7.05	7.40	7.61	7.81	7.96	8.18	8.54	8.88	9.85
6.43	7.05	7.43	7.63	7.83	7.98	8.20	8.59	8.89	9.87
6.50	7.11	7.43	7.64	7.83	7.99	8.21	8.61	8.93	9.96
6.51	7.18	7.50	7.64	7.84	8.00	8.21	8.68	9.00	10.62

Ortalama uyku süresindeki verilerin normal dağılım gösterdiği düşünülmektedir ve hipotezler,

$$H_0: Ortalama\ uyku\ süreleri\ normal\ dağılımdan\ gelmektedir.$$

$$H_1: Ortalama\ uyku\ süreleri\ normal\ dağılımdan\ gelmemektedir.$$

biçiminde yazılır. Bu yokluk hipotezini sınamak için veri setine Kolmogorov-Smirnov testi (KS), Anderson-Darling testi (AD), Cramér-von Mises testi (CVM), ki-kare testi (BCS) ve önerilen bağımsız ki-kare testi (FCS) $\alpha = 0.05$ anlamlılık düzeyinde uygulanmıştır. Uyum

iyiliği testlerinden elde edilen sonuçlar, p-değerleri ve test istatistikleri Tablo 33'te verilmiştir.

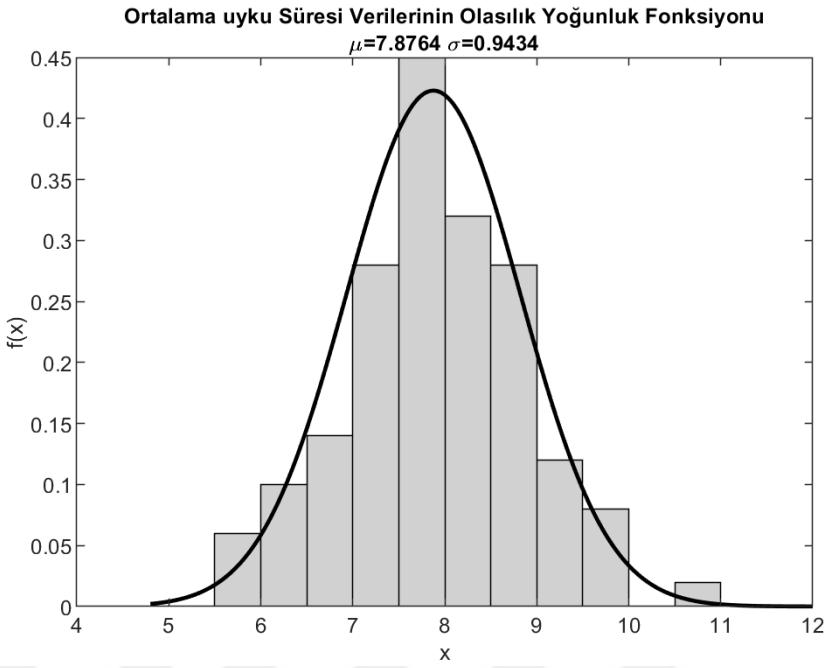
Tablo 33. Ortalama uyku süresi veri setinin uyum iyiliği testi sonuçları

		Test istatistiği	***p*-değeri**	**Hipotez**
Ki-kare (BCS)	:	5.5484	0.4756	Kabul
Cramér-von Mises (CVM)	:	0.0662	0.7758	Kabul
Kolmogorov-Smirnov (KS)	:	0.0618	0.8166	Kabul
Anderson-Darling (AD)	:	0.3969	0.8515	Kabul
Bağımsız ki-kare (FCS)	:	1.0827	0.7090	Kabul

Ortalama uyku süresi veri setine uygulanan tüm uyum iyiliği testleri için $p > \alpha = 0.05$ olduğundan yokluk hipotezi reddedilemez. Yani, uyum iyiliği testleri sonucunda ortalama uyku süreleri normal dağılımdan gelmektedir. Normal dağılımın parametreleri için en çok olabilirlik tahmin edicileri kullanılmıştır ve bunun sonucunda $\hat{\mu} = 7.8764$ ve $\hat{\sigma} = 0.9434$ olarak bulunmuştur. Tahmin edilen parametre değerlerine göre ortalama uyku süresi verilerinin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{0.9434\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - 7.8764}{0.9434}\right)^2\right) \tag{49}$$

biçiminde yazılabilir. Ortalama uyku süresi verilerinin olasılık yoğunluk fonksiyonunun grafiği Şekil 33'te verilmiştir.



Şekil 33. Ortalama uyku süresi verilerinin olasılık yoğunluk fonksiyonu

3.4. Bağımsız Ki-Kare Testinin Değerlendirilmesi

Bu bölümde, önerilen FCS uyum iyiliği testinin geçerliliğini ve kullanılabilirliğini desteklemek için farklı simülasyon çalışmaları yapılmıştır. Bu amaç doğrultusunda önerilen test yöntemi olan FCS literatürde en çok kullanılan KS test yöntemiyle karşılaştırılmıştır.

3.4.1. Farklı Dağılımlara Göre Bağımsız Ki-Kare Testi

Farklı dağılımlara göre FCS ve KS uyum iyiliği testlerinin davranışları gözlenmek istenmiştir. Bu nedenle, simetrik dağılımlardan olan düzgün dağılım $\left(F_U(x)\right)$, üçgen dağılım $\left(F_T(x)\right)$, normal dağılım $\left(F_N(x)\right)$, student-t dağılımı $\left(F_{St}(x)\right)$ ve Laplace dağılımı $\left(F_L(x)\right)$ kullanılarak her birinden $n = 1000$ örneklem büyüklüğünde rastgele değerler üretilmiştir. Her dağılımdan çekilen bu örneklemler 1000 kez tekrarlanmıştır. Her bir dağılım birbiri ile KS ve FCS yöntemleri kullanılarak test edilmiştir ve yokluk hipotezinin kabul yüzdeleri hesaplanmıştır. KS testi sonucundaki yokluk hipotezinin kabul yüzdeleri Tablo 34'te verilmiştir. Her satırda ve sütundaki en büyük kabul yüzdeleri '*' ile gösterilmiştir.

Tablo 34. KS testi sonucundaki yokluk hipotezinin kabul yüzdeleri

		Varsayılan Kümülatif Dağılımlar				
		$F_U(x)$	$F_T(x)$	$F_N(x)$	$F_{St}(x)$	$F_L(x)$
Gerçek Kümülatif Dağılımlar	$F_U(x)$	**95.9***	0	0	0	0
	$F_T(x)$	0.4	**95.1***	85.3	79.2	0
	$F_N(x)$	0.1	88.2	**96.1***	56.7	0
	$F_{St}(x)$	0	56.7	49.5	**94.6***	0
	$F_L(x)$	0	0	0	0	**94.7***

Tablo 34'e göre, KS uyum iyiliği testinin en büyük kabul yüzdelerini her dağılımın kendisiyle testi sonucunda elde ettiği görülmüştür.

FCS testi sonucundaki yokluk hipotezinin kabul yüzdeleri Tablo 35'te verilmiştir. Her satırda ve sütundaki en büyük kabul yüzdeleri '*' ile gösterilmiştir.

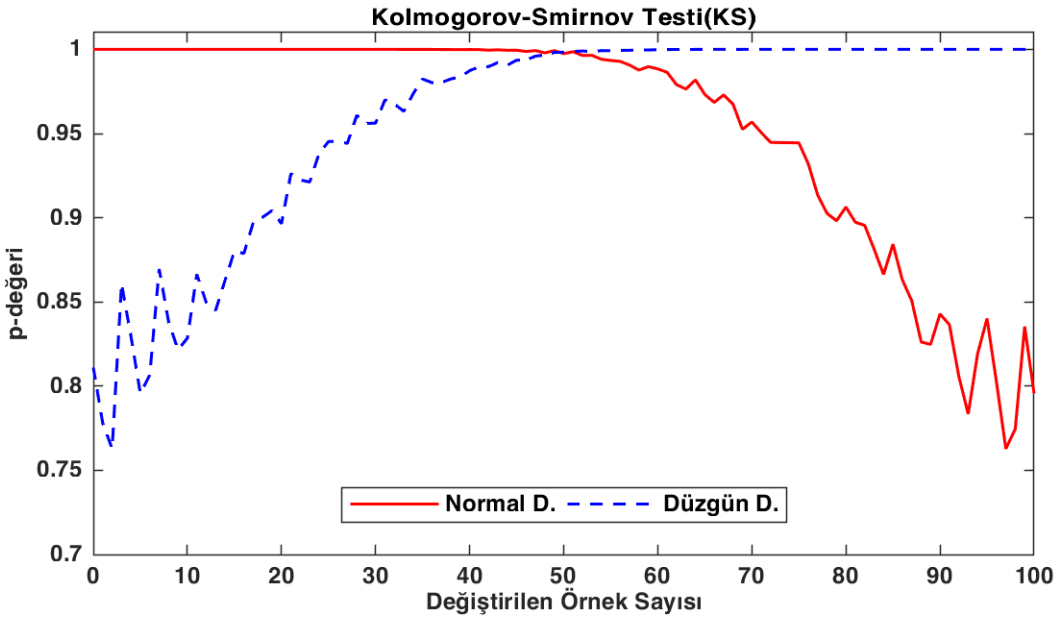
Tablo 35. FCS testi sonucundaki yokluk hipotezinin kabul yüzdeleri

		Varsayılan Kümülatif Dağılımlar				
		$F_U(x)$	$F_T(x)$	$F_N(x)$	$F_{St}(x)$	$F_L(x)$
Gerçek Kümülatif Dağılımlar	$F_U(x)$	**95.7***	0	0	5.6	0
	$F_T(x)$	0.3	**95.3***	68.2	81.5	0
	$F_N(x)$	0	64.2	**94.8***	21	0
	$F_{St}(x)$	7.4	82.9	29	**95***	0
	$F_L(x)$	0	0	0	0	**95.7***

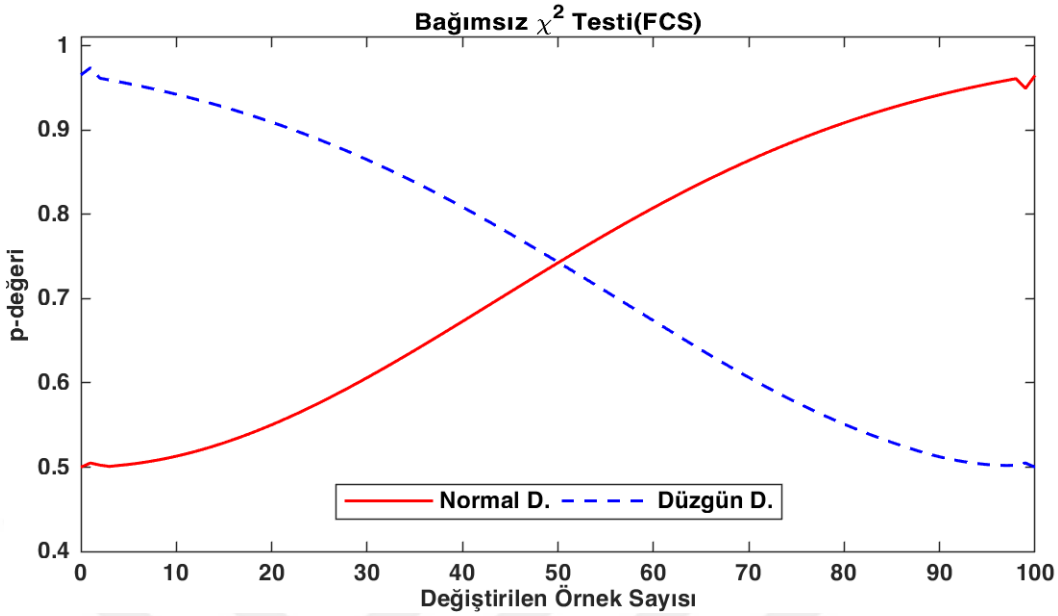
Tablo 35'teki yokluk hipotezindeki kabul yüzdeleri incelendiğinde, FCS uyum iyiliği testinin en büyük değerleri her dağılımın kendisiyle testi sonucunda elde ettiği görülmüştür. Yani, önerilen FCS testinin literatürde en çok kullanılan KS testi gibi doğru ve stabil sonuçlar sağladığı söylenebilir.

3.4.2. Deneysel Rastgele Sayıların Değiştirilmesi ile Bağımsız Ki-Kare Testinin Değerlendirilmesi

Normal $(X_n \sim Normal(0,1))$ ve düzgün dağılımdan $(Y_n \sim Düzgün(0,1))$, $n = 100$ büyüklüğünde iki örneklem için deneysel rastgele sayılar üretilmiştir. İlk olarak, bu iki örnekleme normal ve düzgün dağılıma göre KS ve önerilen FCS uyum iyiliği testleri uygulanarak p-değerleri hesaplanmıştır. Daha sonra, verilere bağlı olarak sistemin değişimini gözlemlemek için rastgele üretilen örneklem verileri birer birer değiştirilmiştir. Her bir değişimde oluşan yeni örneklemler için KS ve FCS uygulanarak p-değerleri hesaplanmıştır. Her aşamadaki p-değerlerinin değişimleri Şekil 34 ve Şekil 35'te verilmiştir.



Şekil 34. Normal ve düzgün dağılıma göre KS uygulanarak elde edilen p-değerlerinin değişimi



Şekil 35. Normal ve düzgün dağılıma göre FCS uygulanarak elde edilen p-değerlerinin değişimi

KS ve önerilen FCS uyum iyiliği testleri uygulanarak elde edilen p-değerlerindeki değişimin, değiştirilen örnek sayısıyla orantılı ve uyum iyiliği testinin dağılımına bağlı olarak artması veya azalması gerekmektedir. Ancak, Şekil 34'teki p-değerlerinin değişimleri incelendiğinde, p-değerlerinde iniş çıkışlar görülmektedir. Şekil 35'teki p-değerlerinin değişimlerine bakıldığında p-değerlerinin düzgün bir şekilde azalıp arttığı görülmektedir. Bu durumda, önerilen FCS uyum iyiliği testinin KS uyum iyiliği testine göre p-değerleri açısından daha tutarlı ve istikrarlı çalıştığı söylenebilir.

4. SONUÇLAR

Bir araştırma sonucunda elde edilen verileri karakterize eden dağılımın bilinmesi istatistikte verilerin analizi aşamasında uygulanacak yöntemin doğru bir şekilde belirlenmesi amacıyla oldukça önemlidir. Hem rastgele üretilen hem de gerçek hayat problemlerinden elde edilen verilerin dağılımının sınanması amacıyla uyum iyiliği testleri kullanılmaktadır. Uyum iyiliği testleriyle ilgili literatürde çok sayıda farklı test yöntemi geliştirilmiştir. Uyum iyiliği testlerinde karar kuralı olarak test istatistiğinin yanı sıra p-değeri kullanılmaktadır. Dolayısıyla, p-değeri istatistikte önemli bir anlam taşımaktadır. Ancak p-değerinin küçük bir değişiminden veya anlamlılık düzeyinin iyi seçilememesinden dolayı testlerin karar aşamasında hatalar yapılmaktadır.

Bu tez çalışmasında, p-değerinden kaynaklı hataları ve stabilite sorunlarını ortadan kaldıracak, literatürdeki uyum iyiliği testlerine alternatif olarak yeni bir test yöntemi ve uyum iyiliği testlerinin güvenilirlik derecesi için ölçüt olarak kullanılabilecek bir olabilirlik indeksi geliştirilmiştir. Çalışmada, öncelikle yönsel yaklaşım temel alınarak örneklemin test edilen dağılıma ne kadar iyi uyup uymadığının derecesini veren bir olabilirlik indeks değeri (Π) hesaplanmış ve ardından bu indeks değerinin dağılımı incelenerek FCS uyum iyiliği testi önerilmiştir.

Önerilen olabilirlik indeks değerine göre örneklemin iddia edilen dağılıma ne kadarlık bir uyum gösterdiği yorumlanabilecektir. Böylece, $p < \alpha$ olması durumunda, kurulan yokluk hipotezini doğrudan reddetmek yerine, olabilirlik indeksi kullanılarak örneklemin dağılımı ne kadar yansıttığı ortaya koyulabilecek ve ne kadarlık bir uyumun çalışma için kabul edilebileceği araştırmacıya bırakılacaktır. Olabilirlik indeks değerinin doğruluğunun ve geçerliliğinin belirlenmesi amacıyla merkezi limit teoremi, büyük sayılar kanunu ve deneysel rastgele sayılardan yararlanılarak yapılan simülasyon çalışmalarında, olabilirlik indeksini temsil eden Π değerinin p-değerine göre daha tutarlı ve daha iyi çalıştığı gösterilmiştir.

Önerilen FCS uyum iyiliği testi, literatürdeki diğer test yöntemlerine göre dağılımdan, sınıf sayısından, serbestlik derecesinden bağımsızlık, duyarlılık, hesaplama ve kullanım açısından kolaylık ve hızlılık olmak üzere birçok avantaja sahiptir. BCS yöntemi, sınıf seçiminden dolayı dezavantajları olan bir uyum iyiliği testidir ve χ^2 yaklaşımının geçerli olması için örneklem boyutunun yeterli olması gerekmektedir. CVM ve KS testleri sınıf

sayısından bağımsız olduğundan tercih edilebilir uyum iyiliği testlerindendir. Ancak KS testi, dağılımın merkezine yakın sapmalara, kuyruklara göre daha duyarlı olması ve testin karar aşamasında KS kritik değer tablosu kullanılması bakımından dezavantajlara sahiptir. Daha duyarlı bir test olan AD testi ise bazı dağılımlar için iyi sonuçlar vermesine rağmen, her dağılıma uygulanamamaktadır. Ayrıca, her dağılım için ayrı bir kritik değer tablosunun olması yöntemin dağılıma bağımlı bir test yöntemi olduğunu göstermektedir.

FCS uyum iyiliği testinin performansını kapsamlı bir şekilde ölçmek için farklı simülasyon çalışmaları yapılmıştır. Monte Carlo yaklaşımı yardımıyla FCS yöntemi, literatürde en çok bilinen ve yaygın bir şekilde kullanılan BCS, CVM, KS ve AD test yöntemleriyle I. tip hata ve testin gücü bakımında karşılaştırılmıştır. Bu karşılaştırmalar hem simetrik hem de simetrik olmayan dağılımlar kullanılarak gerçekleştirilmiştir. FCS yönteminin, simetrik dağılımlar için karşılaştırılan diğer yöntemlere göre daha duyarlı ve daha yüksek güç değerlerine sahip olduğu görülürken, I. tip hatalara göre yöntemin doğru bir şekilde çalıştığı gözlenmiştir. Simetrik olmayan dağılımlar için testlerin güç değerleri incelendiğinde, AD testi daha yüksek sonuçlar vermesine rağmen, önerilen FCS yöntemi kullanılarak karşılaştırılan diğer yöntemlerden daha iyi sonuçlar elde edilmiştir. Fakat en yüksek testin gücü değerlerine sahip olan AD testi her dağılıma uygulanamamaktadır yani, bu test yöntemi dağılıma bağımsız bir yöntem değildir.

Önerilen FCS uyum iyiliği testi, beş farklı gerçek veri setine de uygulanmış ve başarımı gösterilmiştir. Simülasyon çalışmasının son kısmında da FCS yönteminin geçerliliğini ve kullanılabilirliğini desteklemek için KS test yöntemiyle karşılaştırmalar yapılmıştır. Bu simülasyon sonucunda da FCS testinin KS testi gibi doğru ve stabil sonuçlar verdiği gözlenmiştir. Fakat p-değerleri açısından karşılaştırıldıklarında, FCS uyum iyiliği testinin KS testine göre daha tutarlı ve istikrarlı çalıştığı söylenebilir.

Sonuç olarak, önerilen FCS test yöntemi duyarlı bir test olmasının yanı sıra aynı zamanda sınıflandırmadan ve dağılımdan bağımsız bir testtir. Bu avantajlara ek olarak, önerilen test yöntemi verileri gruplandırma veya sıralama işlemine ihtiyaç duymadığı gibi deneysel dağılım fonksiyonuna da ihtiyaç duymamaktadır. Sadece iki serbestlik dereceli bir ki-kare tablosu kullandığı için serbestlik derecesinden de bağımsız bir yöntemdir. Ayrıca, bu çalışmada, önerilen yöntemin basitliği ve hesaplama açısından kolaylığı gösterilmiştir. Bu durumun, önerilen test yöntemini birçok araştırmacının tercih edeceği bir yöntem haline getireceği düşünülmektedir.

5. ÖNERİLER

Doktora tezi olarak sunulan bu çalışmada, uyum iyiliği testlerinde örneklemin iddia edilen dağılımı ne derece yansıttığını ortaya koyabilecek bir olabilirlik indeksi ve bu indeksin dağılımdan yararlanarak yeni bir uyum iyiliği testi geliştirilmiştir. Bu tez konusu ile çalışma yapmak isteyecek araştırmacılara ve uygulayıcılara iletilmek istenen öneriler aşağıdaki gibi belirtilmiştir:

- Olabilirlik indeksi, örneklemin dağılımı ne kadar yansıttığını ortaya koymaktadır ve çalışma için ne kadarlık bir uyumun kabul edilebileceğini araştırmacıya bırakmaktadır. Bu nedenle, önerilen indeks değeri çok örneklem ve çok değişkenli durumlar için geliştirilip bulanık çıkarsama sistemlerinde kullanılabilir.
- Bu tezde literatürde yaygın bir şekilde kullanılan simetrik ve simetrik olmayan dağılımlar kullanılarak simülasyon çalışmaları yapılmıştır. Çalışmayı geliştirmek için daha karmaşık yapıdaki dağılımlar da kullanılabilir.
- FCS testi tek örneklem için geliştirilmiş ve uygulanmıştır. Önerilen yöntemin iki örneklem için geliştirilip, mevcut yöntemlerle karşılaştırılması yapılabilir.
- Önerilen FCS uyum iyiliği testi tek değişkenli durum için geliştirilmiştir. Çok değişkenli durum için uyarlanabilir.

6. KAYNAKLAR

Abelson, R. P. 1997. "A retrospective on the significance test ban of 1999 (If there were no significance tests, they would be invented)." 117-141 in *What if there were no significance tests?*, edited by Harlow, L. L., Mulaik, S. A. ve Steiger, J. H. New Jersey, Mahwah: Erlbaum.

Altın Yavuz, A., 2021. İlerleyen Sansürlü Örneklemlere Dayalı Olarak Weibull Dağılımının Şekil Parametresinin Sağlam Tahmin Edicileri, Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 25, 2, 410-418.

Anderson, T. W., 1962. On the distribution of the two-sample Cramer-von Mises criterion, The Annals of Mathematical Statistics, 33, 3, 1148-1159.

Anderson, T. W. ve Darling, D. A., 1952. Asymptotic theory of certain goodness of fit criteria based on stochastic processes, The Annals of Mathematical Statistics, 23, 2, 193-212.

Anderson, T. W. ve Darling, D. A., 1954. A test of goodness of fit, Journal of the American Statistical Association, 49, 268, 765-769.

Angus, J. E., 1994. The probability integral transform and related results, SIAM review, 36, 4, 652-654.

Arizono, I. ve Ohta, H., 1989. A test for normality based on Kullback—Leibler information, The American Statistician, 43, 1, 20-22.

Arshad, M., Rasool, M. ve Ahmad, M., 2003. Anderson Darling and modified Anderson Darling tests for generalized Pareto distribution, Pakistan Journal of Applied Sciences, 3, 2, 85-88.

Bain, L. J. ve Engelhardt, M., 1987. Introduction to probability and mathematical statistics, Brooks/Cole.

Balakrishnan, N. ve Nevzorov, V. B., 2004. A primer on statistical distributions, John Wiley & Sons Inc., Hoboken. New Jersey.

Balakrishnan, N., Voinov, V. ve Nikulin, M. S., 2013. Chi-squared goodness of fit tests with applications, Elsevier, United States of America.

Baringhaus, L. ve Henze, N., 1988. A consistent test for multivariate normality based on the empirical characteristic function, Metrika, 35, 1, 339-348.

Batsidis, A., Economou, P. ve Bar-Lev, S., 2022. A Comparative Study of Goodness-of-Fit Tests for the Laplace Distribution, Austrian Journal of Statistics, 51, 2, 91-123.

Bayoud, H. A., 2021. Tests of normality: new test and comparative study, Communications in Statistics-Simulation and Computation, 78, 1, 1-22.

Bera, A. K. ve Jarque, C. M., 1981. Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence, Economics Letters, 7, 4, 313-318.

Bonett, D. G. ve Seier, E., 2002. A test of normality with high uniform power, Computational Statistics & Data Analysis, 40, 3, 435-445.

Bontemps, C. ve Meddahi, N., 2005. Testing normality: a GMM approach, Journal of Econometrics, 124, 1, 149-186.

Brys, G., Hubert, M. ve Struyf, A., 2008. Goodness-of-fit tests based on a robust measure of skewness, Computational statistics, 23, 3, 429-442.

Brys, G., Hubert, M. ve Struyf, A., 2008. Goodness-of-fit tests based on a robust measure of skewness, Computational Statistics, 23, 3, 429-442.

Cabaña, A. ve Cabaña, E. M., 2003. Tests of normality based on transformed empirical processes, Methodology and Computing in Applied Probability, 5, 3, 309-335.

Chaichatschwal, R. ve Budsaba, K., 2007. A power comparison of goodness-of-fit tests for normality based on the likelihood ratio and the non-likelihood ratio, Thailand Statistician, 5, 2007, 57-68.

Chen, L. ve Shapiro, S. S., 1995. An alernative test for normality based on normalized spacings, Journal of Statistical Computation and Simulation, 53, 3-4, 269-287.

Cochran, W. G., 1952. The $\chi 2$ test of goodness of fit, The Annals of Mathematical Statistics, 23, 3, 315-345.

Coin, D., 2008. A goodness-of-fit test for normality based on polynomial regression, Computational Statistics & Data Analysis, 52, 4, 2185-2198.

Cox, D. R., 1961. Tests of separate families of hypotheses, Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, 1, 105-123.

Cox, D. R., 1962. Further results on tests of separate families of hypotheses, Journal of the Royal Statistical Society. Series B (Methodological), 24, 2, 406-424.

Cramér, H., 1928. On the composition of elementary errors, Scandinavian Actuarial Journal, 1928, 1, 13-74.

Crzcgorzewski, P. ve Wirczorkowski, R., 1999. Entropy-based goodness-of-fit test for exponentiality, Communications in Statistics-Theory and Methods, 28, 5, 1183-1202.

D'Agostino, R. B. ve Stephens, M. A., 1986. Goodness-of-Fit Techniques, Marcel Dekker Inc., New York.

D'Agostino, R. B., 1971. An omnibus test of normality for moderate and large size samples, Biometrika, 58, 2, 341-348.

D'agostino, R. ve Pearson, E. S., 1973. Tests for departure from normality. Empirical results for the distributions of b 2 and√ b, Biometrika, 60, 3, 613-622.

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P. ve Meester, L. E., 2005. A Modern Introduction to Probability and Statistics: Understanding why and how, Springer, Delft. Netherlands.

del Barrio, E., Cuesta-Albertos, J. A., Matrán, C. ve Rodríguez-Rodríguez, J. M., 1999. Tests of goodness of fit based on the L2-Wasserstein distance, Annals of Statistics, 27, 4, 1230-1239.

Dodge, Y., 2006. The Oxford dictionary of statistical terms, Oxford University Press , New York.

Domanski, C. ve Szczepocki, P., 2020. Comparison of selected tests for univariate normality based on measures of moments., Statistics in Transition New Series, 21, 5, 151-179.

Dong, L. B. ve Giles, D. E., 2007. An empirical likelihood ratio test for normality, Communications in Statistics—Simulation and Computation, 36, 1, 197-215.

Doornik, J. A. ve Hansen, H., 2008. An omnibus test for univariate and multivariate normality, Oxford Bulletin of Economics and Statistics, 70, 1, 927-939.

Dudewicz, E. J., van der Meulen, E. C., SriRam, M. G. ve Teoh, N. K. W., 1995. Entropy-based random number evaluation, American Journal of Mathematical and Management Sciences, 15, 1-2, 115-153.

Epps, T., 2005. Tests for location-scale families based on the empirical characteristic function, Metrika, 62, 1, 99-114.

Epps, T. W. ve Pulley, L. B., 1983. A test for normality based on the empirical characteristic function, Biometrika, 70, 3, 723-726.

Epps, T., Singleton, K. ve Pulley, L., 1982. A test of separate families of distributions based on the empirical moment generating function, Biometrika, 69, 2, 391-399.

Esteban, M., Castellanos, M., Morales, D. ve Vajda, I., 2001. Monte Carlo comparison of four normality tests using different entropy estimates, Communications in Statistics-Simulation and Computation, 30, 4, 761-785.

Esteban, M. D., Marhuenda, Y., Morales, D. ve Sanchez, A., 2007. New goodness-of-fit tests based on sample quantiles, Communications in Statistics—Simulation and Computation, 36, 3, 631-642.

Evren, A. ve Tuna, E., 2012. On some properties of goodness of fit measures based on statistical entropy, International Journal of Research and Reviews in Applied Sciences, 13, 1, 192-205.

Fan, J., 1996. Test of significance based on wavelet thresholding and Neyman's truncation, Journal of the American Statistical Association, 91, 434, 674-688.

Fan, Y., 1997. Goodness-of-fit tests for a multivariate distribution by the empirical characteristic function, Journal of Multivariate Analysis, 62, 1, 36-63.

Fan, Y., 1998. Goodness-of-fit tests based on kernel density estimators with fixed smoothing parameters, Econometric Theory, 14, 5, 604-621.

Filliben, J. J., 1975. The probability plot correlation coefficient test for normality, Technometrics, 17, 1, 111-117.

Forbes, C., Evans, M., Hastings, N. ve Peacock, B., 2011. Statistical distributions, Fourth Edition ed., John Wiley & Sons Inc., Hoboken. New Jersey.

Fortiana, J. ve Grané, A., 2003. Goodness-of-fit tests based on maximum correlations and their orthogonal decompositions, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65, 1, 115-126.

Freeman, M. F. ve Tukey, J. W., 1950. Transformations related to the angular and the square root, The Annals of Mathematical Statistics, 21, 4, 607-611.

Frosini, B. V. 1987. "On the distribution and power of a goodness-of-fit statistic with parametric and nonparametric applications." 133-154 in *Goodness-of-fit,* edited by Revesz, P., Sarkadi, K. ve Sen, P. K. Amsterdam: North Holland Publishing Company.

Gel, Y. R. ve Gastwirth, J. L., 2008. A robust modification of the Jarque-Bera test of normality, Economics Letters, 99, 1, 30-32.

Glen, A. G., Leemis, L. M. ve Barr, D. R., 2001. Order statistics in goodness-of-fit testing, IEEE Transactions on Reliability, 50, 2, 209-213.

Gokhale, D., 1983. On entropy-based goodness-of-fit tests, Computational Statistics & Data Analysis, 1, 157-165.

Grané, A., 2012. Exact goodness-of-fit tests for censored data, Annals of the Institute of Statistical Mathematics, 64, 6, 1187-1203.

Hegazy, Y. ve Green, J., 1975. Some New Goodness-Of-Fit Tests Using Order Statistics, Journal of the Royal Statistical Society: Series C (Applied Statistics), 24, 3, 299-308.

Heyde, C., 1975. A supplement to the strong law of large numbers, Journal of Applied Probability, 12, 1, 173-175.

Hosking, J. R., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics, Journal of the Royal Statistical Society: Series B (Methodological), 52, 1, 105-124.

Hubbard, R. ve Lindsay, R. M., 2008. Why P values are not a useful measure of evidence in statistical significance testing, Theory & Psychology, 18, 1, 69-88.

Jarque, C. M. ve Bera, A. K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals, Economics letters, 6, 3, 255-259.

Johnson, N. L., Kotz, S. ve Balakrishnan, N., 1995. Continuous univariate distributions, volume 2, Second Edition ed., John Wiley & Sons Inc., Hoboken. New Jersey.

Johnson, N. L., Kotz, S. ve Balakrishnan, N. 2019. "Genesis." 1-4 in *Exponential Distribution: Theory, Methods and Applications*, edited by Balakrishnan, N. ve Basu, A. P. New York: Taylor & Francis Group.

Khinchin, A. I., 1957. Mathematical foundations of information theory, Dover Publication Inc., New York.

Kobayashi, H., Mark, B. L. ve Turin, W., 2012. Probability, random processes, and statistical analysis, Cambridge University Press, New York. USA.

Kolmogorov, A. N., 1933. Grundbegriffe der Wahrscheinlichkeitsrechnung, Julius Springer, Berlin.

Kuiper, N. H., 1960. Tests concerning random points on a circle, Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen: Series A, 63, 38-47.

Kullback, S. ve Leibler, R. A., 1951. On information and sufficiency, The Annals of Mathematical Statistics, 22, 1, 79-86.

Laio, F., 2004. Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, Water Resources Research, 40, 9.

Ledwina, T., 1994. Data-driven version of Neyman's smooth test of fit, Journal of the American Statistical Association, 89, 427, 1000-1005.

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

Mardia, K. V., 1972. Statistics of Directional Data, Academic Press, London. England.

Marques de Sá, J., 2007. Applied statistics using SPSS, statistica, Matlab and R, Springer Company, USA.

Massey Jr, F. J., 1951. The Kolmogorov-Smirnov test for goodness of fit, Journal of the American Statistical Association, 46, 253, 68-78.

Matsui, M. ve Takemura, A., 2005. Empirical characteristic function approach to goodness-of-fit tests for the Cauchy distribution with parameters estimated by MLE or EISE, Annals of the Institute of Statistical Mathematics, 57, 1, 183-199.

Mendes, M. ve Pala, A., 2003. Type I error rate and power of three normality tests, Pakistan Journal of Information and Technology, 2, 2, 135-139.

Nakas, C. T., 2007. Performance of the One-Sample Goodness-of-Fit P-P-Plot Length Test, Communications in Statistics—Simulation and Computation, 36, 5, 1053-1059.

Neyman, J., 1949. Contribution to the theory of $\chi 2$ test, Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability, 1, 239-273.

Noughabi, H. A. ve Arghami, N. R., 2011. Monte Carlo comparison of seven normality tests, Journal of Statistical Computation and Simulation, 81, 8, 965-972.

Onyper, S. V., Thacher, P. V., Gilbert, J. W. ve Gradess, S. G., 2012. Class start times, sleep, and academic performance in college: A path analysis, Chronobiology International, 29, 3, 318-335.

Owen, A. B., 2001. Empirical likelihood, 1st ed., Chapman and Hall/CRC, New York.

Park, S., 1999. A goodness-of-fit test for normality based on the sample entropy of order statistics, Statistics & probability letters, 44, 4, 359-363.

Pearson, K., 1900. X. 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, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 50, 302, 157-175.

Pettitt, A., 1977. Testing the normality of several independent samples using the Anderson-Darling statistic, Journal of the Royal Statistical Society: Series C (Applied Statistics), 26, 2, 156-161.

Rahman, M. M. ve Govindarajulu, Z., 1997. A modification of the test of Shapiro and Wilk for normality, Journal of Applied Statistics, 24, 2, 219-236.

Ramachandran, K. ve Tsokos, C. P., 2015. Mathematical Statistics with Applications in R, Second Edition ed., Academic Press.

Ramachandran, K. M. ve Tsokos, C. P., 2021. Mathematical statistics with applications in R, Third Edition ed., Academic Press.

Razali, N. M. ve Yap, B. W., 2011. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests, Journal of Statistical Modeling and Analytics, 2, 1, 21-33.

Reschenhofer, E. ve Bomze, I., 1991. Length tests for goodness of fit, Biometrika, 78, 1, 207-216.

Romao, X., Delgado, R. ve Costa, A., 2010. An empirical power comparison of univariate goodness-of-fit tests for normality, Journal of Statistical Computation and Simulation, 80, 5, 545-591.

Rosenblatt, M., 1956. A central limit theorem and a strong mixing condition, Proceedings of the National Academy of Sciences, 42, 1, 43-47.

Rosin, P. ve Rammler, E., 1933. Laws governing the fineness of a powdered coal, Journal of Institute of Fuel, 7, 29-36.

Saksena, R. S., 1980. A Handbook of Statistics, Motilal Banarsidass Publishe.

Seier, E., 2002. Comparison of tests for univariate normality, InterStat Statistical Journal, 1, 2002, 1-17.

Shannon, C., 1948. A mathematical theory of communication, The Bell System Technical Journal, 27, 3, 379-423.

Shan, G., Vexler, A., Wilding, G. E. ve Hutson, A. D., 2010. Simple and exact empirical likelihood ratio tests for normality based on moment relations, Communications in Statistics—Simulation and Computation, 40, 1, 129-146.

Shapiro, S. S. ve Francia, R. S., 1972. An approximate analysis of variance test for normality, Journal of the American Statistical Association, 67, 337, 215-216.

Shapiro, S. S. ve Wilk, M. B., 1965. An analysis of variance test for normality (complete samples), Biometrika, 52, 3/4 , 591-611.

Shapiro, S. S., Wilk, M. B. ve Chen, H. J., 1968. A comparative study of various tests for normality, Journal of the American Statistical Association, 63, 324, 1343-1372.

Siraj-Ud-Doulah, M., 2019. A Comparison among Twenty-Seven Normality Tests, Research & Reviews: Journal of Statistics, 8, 3, 41-59.

Smirnov, N. V., 1939. Estimate of deviation between empirical distribution functions in two independent samples, Bulletin Moscow University, 2, 2, 3-16.

Song, K.-S., 2002. Goodness-of-fit tests based on Kullback-Leibler discrimination information, IEEE Transactions on Information Theory, 48, 5, 1103-1117.

Stephens, M. A., 1974. EDF statistics for goodness of fit and some comparisons, Journal of the American Statistical Association, 69, 347, 730-737.

Student, 1908. The probable error of a mean, Biometrika, 1, 6, 1-25.

Sullivan, G. M. ve Feinn, R., 2012. Using effect size-or why the P value is not enough, Journal of Graduate Medical Education, 4, 3, 279-282.

Tolikas, K. ve Heravi, S., 2008. The Anderson-Darling goodness-of-fit test statistic for the three-parameter lognormal distribution, Communications in Statistics—Theory and Methods, 37, 19, 3135-3143.

Torabi, H., Montazeri, N. H. ve Grané, A., 2016. A test for normality based on the empirical distribution function, Statistics and Operations Research Transactions, 40, 1, 55-88.

Towhidi, M. ve Salmanpour, M., 2007. A new goodness-of-fit test based on the empirical characteristic function, Communications in Statistics—Theory and Methods, 36, 15, 2777-2785.

Uhm, T. ve Yi, S., 2021. A comparison of normality testing methods by empirical power and distribution of P-values, Communications in Statistics-Simulation and Computation, 1-14.

Uyanto, S. S., 2022. An Extensive Comparisons of 50 Univariate Goodness-of-fit Tests for Normality, Austrian Journal of Statistics, 51, 3, 45-97.

van Zyl, J. M., 2018. The performance of univariate goodness-of-fit tests for normality based on the empirical characteristic function in large samples, Communications in Statistics-Simulation and Computation, 47, 4, 1146-1156.

Vasicek, O., 1976. A test for normality based on sample entropy, Journal of the Royal Statistical Society. Series B (Methodological), 38, 1, 54-59.

von Plato, J. 2005. "AN Kolmogorov, Grundbegriffe der wahrscheinlichkeitsrechnung (1933)." 960-969 in *Landmark Writings in Western Mathematics 1640-1940*, edited by Guinness, I. G. Elsevier.

Wang, J. C., 2012. Sample distribution function based goodness-of-fit test for complex surveys, Computational Statistics & Data Analysis, 56, 3, 664-679.

Watson, G. S., 1961. Goodness-of-fit tests on a circle, Biometrika, 48, 1/2, 109-114.

Watson, G. S., 1962. Goodness-of-fit tests on a circle. II, Biometrika, 49, 1/2, 57-63.

Weibull, W., 1939. A statistical theory of the strength of material, Ingeniors Vetenskapa Acadamiens Handligar, 151, 5-45.

Weibull, W., 1951. A statistical distribution function of wide applicability, Journal of Applied Mechanics, 18, 192-297.

Wijekularathna, D. K., Manage, A. B. ve Scariano, S. M., 2019. Power analysis of several normality tests: A Monte Carlo simulation study, Communications in Statistics-Simulation and Computation, 51, 3, 757-773.

Wilks, S. S., 1935. The likelihood test of independence in contingency tables, The Annals of Mathematical Statistics, 6, 4, 190-196.

Yap, B. W. ve Sim, C. H., 2011. Comparisons of various types of normality tests, Journal of Statistical Computation and Simulation, 81, 12, 2141-2155.

Yazici, B. ve Yolacan, S., 2007. A comparison of various tests of normality, Journal of Statistical Computation and Simulation, 77, 2, 175-183.

Yıldırım, N. ve Gökpınar, F., 2012. Bazı Normallik Testlerinin 1. Tip Hataları ve Güçleri Bakımından Kıyaslanması, Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 16, 1, 109-115.

Zghoul, A. A., 2010. A goodness of fit test for normality based on the empirical moment generating function, Communications in Statistics-Simulation and Computation, 39, 6, 1292-1304.

Zhang, P., 1999. Omnibus test of normality using the Q statistic, Journal of Applied Statistics, 26, 4, 519-528.

Zhang, J., 2002. Powerful goodness-of-fit tests based on the likelihood ratio, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64, 2, 281-294.

Zhang, J. ve Wu, Y., 2005. Likelihood-ratio tests for normality, Computational Statistics & Data Analysis, 49, 3, 709-721.

ÖZGEÇMİŞ

Özge TEZEL, İlköğrenimini Yavuz Selim İlköğretim Okulu'nda bitirdikten sonra lise eğitimini 2008 yılında Trabzon (YDA) Lisesi'nde tamamladı. 2008-2009 öğretim yılında Karadeniz Teknik Üniversitesi Fen Fakültesi İstatistik ve Bilgisayar Bilimleri bölümünde lisans eğitimine başladı. 2012 yılında bu bölümden mezun oldu. Aynı yıl Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik ve Bilgisayar Bilimleri Anabilim dalında tezli yüksek lisans programına başladı. 2015 yılında "Yönsel Verilerin Kümelenmesinde Bulanık C-Ortalamalar Algoritması" başlıklı yüksek lisans tezini sunmasının ardından doktora eğitimine başladı. 2014 yılından itibaren Karadeniz Teknik Üniversitesi İstatistik ve Bilgisayar Bilimleri Bölümü'nde araştırma görevlisi olarak görev yapmaktadır. İyi derece İngilizce bilmektedir.

Tezel, Ö., Tiryaki, B. K., Özkul, E., ve Kesemen, O., (2021). A New Goodness-of-Fit Test: Free Chi-Square (FCS). Gazi University Journal of Science, 34(3), 879-897.