

ÇUKUROVA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

Karma Dağılımların Hizmet Sektöründe Uygulanması

Selin SARAÇ GÜLERYÜZ

Endüstri Mühendisliği Anabilim Dalı

Haziran, 2023

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZ ONAYI

Karma Dağılımların Hizmet Sektöründe Uygulanması

Selin SARAÇ GÜLERYÜZ

Endüstri Mühendisliği Anabilim Dalı

Bu Doktora Tezi 02/06/2023 Tarihinde Aşağıdaki Jüri Üyeleri Tarafından Değerlendirilmiş ve Oy Birliği ile Kabul Edilmiştir.

Jüri : Doç. Dr. Melik KOYUNCU (Danışman)
: Prof. Dr. Ali İhsan GENÇ
: Doç. Dr. Z. Figen ANTMEN
: Dr. Öğr. Üyesi Fikri EGE
: Dr. Öğr. Üyesi M. Zühal E. BARAK

Bu Tez Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalında Hazırlanmıştır.

Tez No:

Prof. Dr. Sadık DİNÇER
Enstitü Müdürü

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

İÇİNDEKİLER

ÖZ	I
ABSTRACT	II
TEŞEKKÜR	III
ÇİZELGELER DİZİNİ	IV
ŞEKİLLER DİZİNİ	V
SİMGELER VE KISALTMALAR	VII
1. GİRİŞ	1
1.1. Kapasite Planlaması	1
1.2. Yoğun Bakım Üniteleri	3
1.3. Simülasyon Modelleme	7
1.4. Karma Dağılımların Kullanımı	8
1.5. Problemin Tanımı	10
1.6. Çalışmanın Amacı ve Kapsamı	10
1.7. Çalışmanın Orijinal Katkısı	12
1.8. Tezin Yapısı	12
2. ÖNCEKİ ÇALIŞMALAR	13
3. MATERYAL VE METOT	17
3.1. Temel İstatistik Dağılımlar	19
3.1.1. Normal Dağılım	19
3.1.2. Lognormal Dağılım	21
3.1.3. Üstel Dağılım	23
3.1.4. Gamma Dağılımı	25
3.1.5. Weibull Dağılımı	28
3.1.6. Ampirik Dağılım	31
3.2. Veri Toplama	31
3.3. Olasılık Dağılım Ailesinin Seçimi	34
3.3.1. Yer Gösteren Ölçüler	36
3.3.2. Yayınlık Ölçüleri	40
3.3.3. Grafikselsel Yöntemler	41
3.4. Parametre Tahmini	47
3.5. Uygunluk Testi	48
3.6. Karma Dağılımlar	49
3.6.1. Karma Dağılımlarda Bileşen Sayısı Belirleme	55
3.6.2. Karma Dağılımlarda Parametre Tahmini	64
3.6.3. Performans Kriteri	67

3.7. Yoğun Bakım Ünitesi Modelinin Oluşturulması	69
3.7.1. Kuyruk Teorisi	70
3.7.2. Simülasyon Modeli	75
4. BULGULAR VE TARTIŞMA	79
4.1. Veri Analizi.....	79
4.2. Karma Dağılım Analizi.....	108
4.3. Ampirik Dağılım Oluşturma	116
4.4. Performans Kriterlerini Değerlendirme	117
4.5. Simülasyon Modeli Sonuçları.....	121
5. SONUÇLAR VE ÖNERİLER	127
5.1. Çalışmanın Özeti ve Sonuçlar.....	127
5.2. Sonraki Çalışmalar için Öneriler.....	128
KAYNAKLAR	131
ÖZGEÇMİŞ	139

Karma Dağılımların Hizmet Sektöründe Uygulanması

Selin SARAÇ GÜLERYÜZ

Danışman: Doç Dr. Melik KOYUNCU

Endüstri Mühendisliği Anabilim Dalı

ÖZ

Yoğun bakım üniteleri (YBÜ) hastanelerin en önemli unsurlarından biridir. Bu üniteler sağlık sisteminde ve kritik hastalara bakım sağlamada merkezi bir rol oynar, bu nedenle bu birimlerde kapasite planlamasında yüksek öneme sahiptir. YBÜ yataklarının ve personelin yetersiz olması, hasta ölümü de dahil olmak üzere onarılmaz sonuçlara yol açabilir. Sonuç olarak hastane yöneticileri bu ünitelerde uygun yatak sayısını belirlemek için bir takım metotlar geliştirmelidir. Bununla birlikte, hastaların YBÜ'lere gelişleri arası süresi ve bu ünitelerde aldığı servisin süresi doğası gereği stokastiktir. Sonuç olarak, kapasite planlaması dinamik bir operasyon yönetimi problemidir. Gelişler arası süre ve servis süresi genellikle çok çarpık dağılımlara sahiptir ve bilinen saf bir istatistiksel dağılıma uydurmak mümkün olmayabilir. Bu durumda gelişler arası süre ve servis süresini tanımlayabilmek için ampirik dağılımlar oluşturulması gerekir. Ancak ampirik dağılımların parametresi az sayıdadır ve verinin yapısını iyi bir şekilde temsil edemez. Bu tezin temelini, YBÜ'lerdeki hastaların gelişler arası sürelerini ve servis sürelerini tahmin etmek için karma dağılım yaklaşımı kullanılması oluşturmaktadır. Elde edilen karma dağılım parametrelerinin ampirik istatistiksel dağılımlardan daha iyi bir tahmin sağladığını görülmüştür. Tahmin edilen karma dağılım parametreleri, bekleme süresi olmaksızın, gelen tüm hastalara hizmet verebilmek için gereken yatak sayısını belirlemeye yardımcı olan simülasyon modeline dahil edilerek YBÜ analiz edilmiştir.

Anahtar Kelimeler: Servis Süresi, Gelişler arası Süre, Karma Dağılım, Yatak Kapasitesi, Yoğun Bakım Ünitesi.

Application of Mixture Distribution in the Service Sector

Selin SARAÇ GÜLERYÜZ

Advisor: Assoc. Prof. Dr. Melik KOYUNCU

Department of Industrial Engineering

ABSTRACT

Intensive care units (ICUs) are one of the most crucial elements of hospitals. These units play a central role in the healthcare system and providing care to critically ill patients, so capacity planning in these units is of critical importance. Shortages of ICU beds and personnel can lead to irreversible consequences, including patient mortality. As a result, hospital administrators make efforts to determine the appropriate number of beds in these units. However, the inter-arrival time (IAT) of patients to the ICUs and the duration of their service time (ST) in these units are inherently stochastic. Consequently, capacity planning becomes a dynamic operational management problem. The inter-arrival time and service time often exhibit highly skewed distributions, making it difficult to fit a known probability distribution. In such cases, empirical distributions need to be constructed to define the inter-arrival time and service time. Nonetheless, empirical distributions have a small number of parameters and may not represent the structure of the data adequately. This thesis is based on the use of a mixture distribution approach to estimate the inter-arrival times and service time of patients in ICUs. It has been observed that the estimated parameters of the mixture distribution provide better predictions than empirical statistical distributions. The estimated parameters of the mixture distribution were included in a simulation model that helps determine the required number of beds to serve all incoming patients without any waiting time, and the ICU was analyzed accordingly.

Keywords: Service Time, Inter-arrival Time, Mixture Distribution, Bed Capacity, Intensive Care Unit.

TEŞEKKÜR

İlk teşekkürü Çukurova Üniversitesi Endüstri Mühendisliği Anabilim Dalının tüm idari ve akademik üyelerine etmek istiyorum. Çukurova Üniversitesi'nde lisansüstü eğitimim boyunca kendimi büyük bir ailenin bir parçası gibi hissettim. Bu üniversitenin mezunu olmak hayatım boyunca bir gurur kaynağı olacak.

Ardından en büyük teşekkürüm tüm ilgisi, sabrı ve rehberliği için çok değerli danışmanım Sayın Doç Dr. Melik KOYUNCU'ya olacak. Verdiğiniz motivasyon, teşvik ve bilgiyle bana yeni yollar açtınız, size her zaman minnettar olacağım.

Tez İzleme Komitemin değerli üyeleri, Prof. Dr. Ali İhsan GENÇ, Doç. Dr. Z. Figen ANTMEN; tüm doktora eğitimim boyunca her ihtiyacım olduğunda ayırdığınız vakit ve emekleriniz için sizlere çok teşekkür ederim. Sizlerin gösterdiği yeni ufuklar sayesinde bu çalışma var olabildi.

Bir diğer teşekkürüm ise, 2014 yılından beri Araştırma Görevlisi olarak bir parçası olduğum Toros Üniversitesi Mühendislik Fakültesi aileme olacak. Tüm lisansüstü eğitimi sürecinde verdikleri destek, sağladıkları kolaylık ve gösterdikleri anlayış için tüm akademik ve idari personele teşekkürlerimi sunuyorum.

Ve tabi ki sevgili ailem; hep beraber çok zor bir yıl geçirdik. Bu uzun yolda her zaman beni koşulsuz desteklediğiniz, varlığınıza güç verdiğiniz, hayatın tüm iniş çıkışlarında beraber olduğumuz için; annem Gamze SARAÇ, babam Şükrü SARAÇ, kardeşim Cem SARAÇ, eşim Cem GÜLERYÜZ, kedilerim Marla, Tyler, Münevver, Faça ve Patron, sizlere teşekkür ederim.

6 Şubat 2023'te gerçekleşen deprem, ülkemizde büyük bir yıkım ve kayıp yaşattı. Depremden dolayı tüm yaşamı etkilenen ve hayatını kaybeden dostlarımın acısı ve hüznünü hala yüreğimde derin bir şekilde hissediyorum. Bu tezimi onlara ve hala çok güzel olan kadim şehir Antakya'ya ithaf ediyorum.

ÇİZELGELER DİZİNİ

Çizelge 3.1.	Çarpıklık Katsayısı Yorumu	38
Çizelge 3.2.	Dağılım Özellikleri	41
Çizelge 3.3.	İstatistiksel Dağılımların En Çok Olabilirlik Tahmin Edicileri	48
Çizelge 3.4.	$M/M/1/N$ Kuyruk Sistemi için Formüller	72
Çizelge 3.5.	$M/G/1$ Kuyruk Sistemi için Formüller	73
Çizelge 4.1.	Yoğun Bakım Ünitelerinin Tanımlayıcı İstatistikleri	79
Çizelge 4.2.	Ki-kare (χ^2) Uyum İyiliği Testi p -değerleri	107
Çizelge 4.3.	Kolmogorov-Smirnov (K-S) Uyum İyiliği Testi p -değerleri	108
Çizelge 4.4.	Anderson-Darling (A-D) Uyum İyiliği Testi p -değerleri	108
Çizelge 4.5.	Pediatric YBÜ GAS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri	109
Çizelge 4.6.	Pediatric YBÜ SS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri	110
Çizelge 4.7.	Reanimasyon YBÜ GAS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri	110
Çizelge 4.8.	Reanimasyon YBÜ SS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri	111
Çizelge 4.9.	EM Algoritması Kullanılarak Tahmin Edilen Karma Dağılım Parametreleri	112
Çizelge 4.10.	Tahmin Edilen Karma Dağılım Modellerinin Genel Gösterimi	113
Çizelge 4.11.	Tahmin Edilen Ampirik Dağılım Modellerinin Genel Gösterimi	117
Çizelge 4.12.	Ampirik ve Karma Dağılımlarının Performans Kriterlerinin Karşılaştırılması	120

ŞEKİLLER DİZİNİ

Şekil 3.1.	Çukurova Üniversitesi Balcalı EAH akademik personelin istidam bilgileri	18
Şekil 3.2.	Çalışmanın Aşamaları	19
Şekil 3.3.	Standart Normal Dağılım Fonksiyon Grafiği	21
Şekil 3.4.	Farklı Parametre Değerlerinde Lognormal Dağılımları	23
Şekil 3.5.	Farklı Parametre Değerlerinde Üstel Dağılımlar	24
Şekil 3.6.	$k = 1, \theta = 1$ olan Gamma Dağılımı	26
Şekil 3.7.	Farklı Parametre Değerlerinde Gamma Dağılımları	27
Şekil 3.8.	Farklı Parametre Değerlerinde Weibull Dağılımları	29
Şekil 3.9.	$k = 1$ Olduğu Durumda Weibull Dağılımı	30
Şekil 3.10.	Olasılık Dağılımı Seçilmesi	32
Şekil 3.11.	RStudio Ara yüzü	33
Şekil 3.12.	Tanımlayıcı İstatistik Sınıflandırılması	35
Şekil 3.13.	Sola, Sağa Çarpık ve Simetrik Dağılım Grafikleri	39
Şekil 3.14.	$N(50, 5)$ Dağılımının Normal Dağılıma Uyum Grafikleri	44
Şekil 3.15.	$N(50, 5)$ Dağılımının Weibull Dağılıma Uyum Grafikleri	45
Şekil 3.16.	$N(50, 5)$ Dağılımının Üstel Dağılıma Uyum Grafikleri	46
Şekil 3.17.	Pearson'un Yengeç Verisi Histogramı ve Yoğunluk Grafiği	52
Şekil 3.18.	Pearson'un Yengeç Verisi Normallik Kontrolü	53
Şekil 3.19.	Pearson'un Yengeç Verisi Kutu Grafiği	54
Şekil 3.20.	$N(20, 2)$ OYF ve Q-Q Grafiği	58
Şekil 3.21.	$N(20, 2)$ %50 + $N(40, 2)$ %50 OYF ve Q-Q Grafiği	59
Şekil 3.22.	$N(20, 2)$ %30 + $N(40, 2)$ %30 + $N(60, 2)$ %40 OYF ve Q-Q Grafiği	59
Şekil 3.23.	$N(15, 2)$ %50 + $N(20, 2)$ %50 OYF ve Q-Q Grafiği	60
Şekil 3.24.	$N(3, 2)$ %50 + $N(5, 2)$ %50 OYF ve Q-Q Grafiği	60
Şekil 3.25.	Pearson Yengeç Verisi Q-Q Grafiği	61
Şekil 3.26.	Pearson Yengeç Verisi BIC Grafiği	64
Şekil 3.27.	Pearson Yengeç Verisi ve Karma Dağılım OYF'leri	66
Şekil 4.1.	Pediatric YBÜ GAS Histogramı	81
Şekil 4.2.	Reanimasyon YBÜ GAS Histogramı	82
Şekil 4.3.	Pediatric YBÜ SS Histogramı	83
Şekil 4.4.	Reanimasyon YBÜ SS Histogramı	84
Şekil 4.5.	Pediatric ve Reanimasyon YBÜ'lerin GAS ve SS Kutu Grafikleri	85
Şekil 4.6.	Pediatric YBÜ'ye GAS Verisinin Normal Dağılım Uydurma Grafikleri	87
Şekil 4.7.	Pediatric YBÜ'ye GAS Verisinin Weibull Dağılım Uydurma Grafikleri	88
Şekil 4.8.	Pediatric YBÜ'ye GAS Verisinin Gamma Dağılım Uydurma Grafikleri	89

Şekil 4.9. Pediatri YBÜ'ye GAS Verisinin Lognormal Dağılım Uydurma Grafikleri.....	90
Şekil 4.10. Pediatri YBÜ'ye GAS Verisinin Üstel Dağılım Uydurma Grafikleri.....	91
Şekil 4.11. Pediatri YBÜ'deki SS Verisinin Normal Dağılım Uydurma Grafikleri.....	92
Şekil 4.12. Pediatri YBÜ'deki SS Verisinin Weibull Dağılım Uydurma Grafikleri.....	93
Şekil 4.13. Pediatri YBÜ'deki SS Verisinin Gamma Dağılım Uydurma Grafikleri	94
Şekil 4.14. Pediatri YBÜ'deki SS Verisinin Lognormal Dağılım Uydurma Grafikleri	95
Şekil 4.15. Pediatri YBÜ'deki SS Verisinin Üstel Dağılım Uydurma Grafikleri	96
Şekil 4.16. Reanimasyon YBÜ'ye GAS Verisinin Normal Dağılım Uydurma Grafikleri.....	97
Şekil 4.17. Reanimasyon YBÜ'ye GAS Verisinin Weibull Dağılım Uydurma Grafikleri	98
Şekil 4.18. Reanimasyon YBÜ'ye GAS Verisinin Gamma Dağılım Uydurma Grafikleri	99
Şekil 4.19. Reanimasyon YBÜ'ye GAS Verisinin Lognormal Dağılım Uydurma Grafikleri	100
Şekil 4.20. Reanimasyon YBÜ'ye GAS Verisinin Üstel Dağılım Uydurma Grafikleri.....	101
Şekil 4.21. Reanimasyon YBÜ'deki SS Verisinin Normal Dağılım Uydurma Grafikleri	102
Şekil 4.22. Reanimasyon YBÜ'deki SS Verisinin Weibull Dağılım Uydurma Grafikleri.....	103
Şekil 4.23. Reanimasyon YBÜ'deki SS Verisinin Gamma Dağılım Uydurma Grafikleri.....	104
Şekil 4.24. Reanimasyon YBÜ'deki SS Verisinin Lognormal Dağılım Uydurma Grafikleri.....	105
Şekil 4.25. Reanimasyon YBÜ'deki SS Verisinin Üstel Dağılım Uydurma Grafikleri	106
Şekil 4.26. Pediatri YBÜ'ye GAS Verisinin Karma Dağılım Yoğunluk Grafiği.....	114
Şekil 4.27. Pediatri YBÜ'deki SS Verisinin Karma Dağılım Yoğunluk Grafiği	114
Şekil 4.28. Reanimasyon YBÜ'ye GAS Verisinin Karma Dağılım Yoğunluk Grafiği	115
Şekil 4.29. Reanimasyon YBÜ'deki SS Verisinin Karma Dağılım Yoğunluk Grafiği	115
Şekil 4.30. Gerçek, Ampirik ve Karma Dağılımların Yoğunluk Grafiği	119
Şekil 4.31. Senaryo I, Kuyrukta Beklemeye İzin Verilen Simülasyon Modeli	122
Şekil 4.32. Senaryo II, Kuyrukta Beklemeye İzin Verilmeyen Simülasyon Modeli.....	123
Şekil 4.33. Simülasyon Modeli Ara Yüzü.....	124
Şekil 4.34. Senaryo III, Pediatri YBÜ Yatak Sayısına Simülasyon Sonuçları.....	125
Şekil 4.35. Senaryo III, Reanimasyon YBÜ Yatak Sayısına Simülasyon Sonuçları.....	126

SİMGELER VE KISALTMALAR

YBÜ	: Yoğun Bakım Ünitesi
GAS	: Gelişler arası Süre
SS	: Servis/Kalış Süresi
EAH	: Eğitim Araştırma Hastanesi
OYF	: Olasılık Yoğunluk Fonksiyonu
KDF	: Kümülatif Dağılım Fonksiyonu
LRT	: Olabilirlik Oran İstatistiği
AIC	: Akaike Bilgi Kriteri
BIC	: Bayesci Bilgi Kriteri
ML	: En Çok Olabilirlik
EM	: Beklenti Maksimizasyonu
MAPE	: Ortalama Mutlak Yüzde Hatası
MWU	: Mann-Whitney U

1. GİRİŞ

Karma dağılımların tarihi, 19. yüzyılın sonlarına dayanır. Pearson, doğal fenomenlerin istatistiksel modellenmesinde basit tek bir dağılımın yetersiz olduğunu fark etmiş ve farklı bileşenlerin birleşimiyle oluşan karma dağılımların kullanılmasını önermiştir. Karma dağılımların gelişimi, 20. yüzyıl boyunca istatistiksel teori ve yöntemlerdeki ilerlemelerle hız kazanmıştır. İstatistiksel hesaplama gücünün artması, istatistiksel modelleme tekniklerinin daha karmaşık veri setlerine uygulanabilmesini sağlamıştır. Bu dönemde, daha etkili parametre tahmin yöntemleri ve model seçimi teknikleri geliştirilmiştir. Son yıllarda, karma dağılımların gelişimi, makine öğrenmesi ve derin öğrenme alanlarındaki ilerlemelerle daha da hızlanmıştır. Yeni algoritmalar ve yaklaşımlar, karma dağılımların daha büyük ve karmaşık veri setlerine uyarlanmasını sağlamıştır. Bu sayede, gerçek dünya problemlerinin daha iyi modellenmesi ve çözülmesi mümkün hale gelmiştir. Bu problemlerden bir tanesi de sağlık sistemindeki yoğun bakım üniteleridir.

Hastanelerin yoğun bakım üniteleri, kritik hastalara hizmet veren bölümlerdir. Bu ünitelerin yanlış yönetilmesi, hastaların ciddi sağlık problemleri yaşamasına neden olabilir. Genel olarak yoğun bakım ünitelerinde gelen talepler ve bu ünitelerde kalış süreleri stokastik ve istatistiki olarak çarpık olarak nitelendirilebilir. Bu nedenle, bilinen saf istatistiksel dağılımlarla tanımlanamayabilir. Bu aşamada, yoğun bakım ünitelerinde, özellikle simülasyon modellemede gerekli olan girdi parametrelerinin karma dağılımla temsil edilmesi fikri ortaya çıkmıştır.

Bu tezin odak noktası, yoğun bakım ünitelerinin yatak kapasitesi planlamasının karma dağılımlar kullanılarak simülasyon yöntemiyle belirlenmesidir. Bu amaçla, bu bölümde talep ve kapasite planlaması, yoğun bakım ünitelerinin önemi, simülasyon modelleme ve karma dağılımlar hakkında genel bilgiler verilir. Bu bölüm ayrıca tezin kapsamını net bir şekilde açıklamaktadır. Bu şekilde, tezin amacı ve hedefleri daha iyi anlaşılabilir hale gelecektir.

1.1. Kapasite Planlaması

Üretim sistemlerinde planlama, bu sistemlerin verimli bir şekilde yönetilmesi için önemli bileşenlerden bir tanesidir. Planlama ve üretim yönetimi, üretim sistemlerinin başarısı için vazgeçilmez bir unsurdur. Ancak, gelecekteki belirsizlikler ve değişkenler, planlama sürecinde ciddi zorluklar yaratabilir. Bu nedenle, şirketler, geleceğe yönelik kararlarında sistematik bir yöntem kullanarak, doğru ve doğru olmayan planlamalardan kaçınmalıdır (Bulut, 2006). Ürün üretimi, hizmet sunumu ve personel yönetimi gibi faktörler, bir şirketin cevap aradığı en önemli soruların başında gelmektedir. Bu sorulara yanıt vermek, doğru kaynakların kullanımı için gereklidir. Özellikle talebin sürekli arttığı sağlık gibi sektörlerde, bu soruların yanıtı daha da önem kazanmaktadır.

Sağlık sektörü, sürekli artan talebin yanı sıra kaynakların sınırlı olması problemiyle karşı karşıya kalmaktadır. Bu kaynaklar, ekipman, personel, para, zaman ve teknoloji gibi unsurları içermektedir (Uz, 1998). Bu nedenle, sağlık sektöründe, talep tahmini, kapasite ve ekipman planlaması; hastalara zamanında müdahale edebilmek için son derece önemlidir. Ancak, sağlık sektörünün diğer sektörlerden önemli bir farkı, her hastanın farklı ihtiyaç ve beklentilerinin olması ve bu nedenle standart bir hizmet tanımının olmamasıdır (Yıldırım ve ark., 2014).

Sağlık hizmetlerinde talep tahmininin yanı sıra, kapasite planlaması da önemlidir. Doğru kapasite planlaması, gereksiz masrafların önlenmesine yardımcı olabilir ve sağlık hizmetlerinin daha etkili bir şekilde sunulmasına olanak tanır. Bununla birlikte, doğru kapasite planlaması yapmak, karmaşık bir süreçtir ve birden fazla faktör dikkate alınmalıdır. Sağlık sektöründe talep stokastik bir süreç içindedir. Hizmeti sağlayan personel, para, zaman, teknoloji gibi kaynaklar ise kısıtlıdır (Uz, 1998). Özellikle kritik durumdaki hastalar için, ihtiyaç anında doktor, ekipman ve yatak sağlanması hayati önem taşımaktadır. İhtiyaç anında elde olmayan bir donanım çok ciddi sonuçlara neden olabilir. Fazla miktarda bulunan ekipman ya da iş gücü maliyetleri olumsuz yönde etkileyebilir. Kişilerin sağlık hizmet talebi her zaman tedavi olarak değil, bilgi ya da destek amaçlı da olabilir (Mills ve Gilson, 1988). Ancak sağlık hizmeti talebi; gelir düzeyi, yaş, cinsiyet, sağlık kurumunun sağladığı fiziksel imkânlar, hastane çalışanlarının davranışlarına göre pek çok faktöre bağlıdır (Kurtulmuş, 1998). Sağlık sisteminde planlama iki aşamalı kullanılmaktadır. İlk aşama, mevcut durum analiz edilerek, gelecekte hiçbir değişiklik yapılmadığı durumunda olacak senaryodur. İkinci aşama ise, mevcut durum analiziyle gelecekte durumu etkileyecek farklı bir faktör eklendiğinde yaşanacak durum senaryosudur. Örneğin; birinci aşamada, bir hastaneden alınmış belirli bir periyottaki hasta sayısı verinin önümüzdeki periyotta ne olacağı tahmin edilmeye çalışılırken; ikinci aşamada yeni bir teknoloji eklenmesi ya da personel istihdamının bu sayıya nasıl etki edebileceği modellenir. Bu iki aşama ile ilk olarak mevcut sistemin kullanımı planlanabilir ve ikinci aşamada ihtiyaç duyulan insan gücü, donanım gibi hizmet ekipmanları ve türleri tanımlanabilir (Yıldırım ve ark., 2014). Talep tahmininin geçmişte olan olayların gelecekte de aynı şekilde devam etmesi varsayımı nedeniyle, geniş kapsamlı bir model denemekten ziyade hastane faaliyetinin bir bölümünü, örneğin bir acil servis departmanı veya ayakta tedavi kliniğini inceleyerek çözülebileceği ortaya atılmıştır (Günel ve Pidd, 2006). Bu sayede her bir bölümün tahmindeki hata oranı düştükçe, hastanenin genelinde talep tahmininde daha düşük hata oranı elde edilecektir.

Sağlık hizmetlerinde, teknolojik ilerlemeler de dikkate alınmalıdır. Yeni teknolojiler, sağlık hizmetlerinin daha da iyileştirilmesine olanak tanırken, bu teknolojilerin maliyetleri de hesaba katılmalıdır. Bu nedenle, sağlık hizmetleri yöneticileri, gelecekteki teknolojik gelişmeleri dikkate alarak doğru yatırım kararları vermeye çalışmalıdır. Son olarak, sağlık harcamaları da planlama sürecinde dikkate alınmalıdır. Kişi başına sağlık harcamaları sürekli artmaktadır ve bu harcamaların büyük bir kısmı devlet bütçesinden karşılanmaktadır. Bu nedenle, iyi bir planlama,

hem bireysel hem de ulusal katkı sağlar. 2021 yılında toplam sağlık harcamaları, bir önceki yıla göre %41,6 artarak 353 milyar 941 milyon TL'ye yükselmiştir. Bu artışın %41,5'i genel devlet sağlık harcamalarına ait olup, tutarları 280 milyar 220 milyon TL'ye ulaşmıştır. Özel sektör sağlık harcamaları ise %42,1 oranında bir artışla 73 milyar 721 milyon TL olarak tahmin edilmektedir (TUİK, 2021).

Tıp teknolojilerindeki ilerlemelerle birlikte, sağlık sistemleri hızlı bir şekilde değişmektedir. Yaşlı nüfusun artması, doğurganlık oranının düşmesi ve sağlık kurumlarına olan ihtiyacın artması gibi faktörler, kaynakların verimli bir şekilde yönetilmesi gerekliliğini ortaya çıkarmaktadır. Bu zorlukların üstesinden gelmek ve verimlilik ve etkinliği artırmak için, sağlık sistemi yöneticileri sürekli olarak araştırmalar yapmaktadır. Sağlık sistemlerinde verimliliğin artması, kaynakların daha iyi kullanılmasını ve sağlık hizmetlerinin daha geniş bir kesime erişimini sağlayacaktır (Harper ve Shahani, 2002).

Hastanelerdeki görev yapan doktorlar, ekipmanlar ve yataklar gibi kaynaklar, sağlık hizmetleri sağlamada kritik bir rol oynamaktadır. Bu kaynakların herhangi birinin eksikliği, hastaya uygun tedavinin sağlanamamasına neden olacaktır. Yataklar, hastane yöneticileri için yüksek maliyetli bir kaynak olarak görülmesi de, hastaların acil durumlarda hızlı ve etkili bir şekilde tedavi edilebilmeleri için yönetilmesi önemlidir. Yeterli yatak sayısının sağlanması, hastaların gereksinimlerini karşılayacak bir ortam sağlayarak hastane işleyişini optimize etmede kritik bir faktördür (Green, 2004). Özellikle hastanelerin yoğun bakım üniteleri (YBÜ) yatak kapasite planlamasının en kritik olduğu bölgelerden bir tanesidir. Poliklinik ünitelerinde yeterli doktor ya da yatak bulunmaması hastalar açısından her zaman ciddi sonuçlar doğurmaz ancak YBÜ'ler hayati açıdan kritik durumdaki hastalara hizmet vermesinden dolayı, her zaman erişilebilir ve hastaya hazır bir durumda olmalıdır.

1.2. Yoğun Bakım Üniteleri

Hastanelerde kritik durumdaki hastalara hizmet veren yoğun bakım ünitesi (YBÜ) sağlık sisteminde büyük öneme sahiptir ve hastaların acil bakım ihtiyaçlarını karşılamak için kullanılır (Zhu ve ark., 2012). Bu ünitelerde, kritik hastalıklar, yaralanmalar veya ameliyat sonrası komplikasyonlar gibi ciddi sağlık sorunları olan hastaların takibi ve tedavisi yapılmaktadır. YBÜ'ler, bu nedenle, sağlık sisteminde önemli bir yer tutmaktadır. Bu birimlerde yatak sayısı en önemli kaynakların başında gelmektedir ve doğrudan hastane performansını etkilemektedir. YBÜ'ler sağlık sistemlerinde önemli bir rol oynar ve bu birimlerde kapasite planlaması son derece kritiktir. Ancak, etkisiz yatak yönetimi birçok soruna neden olabilir. Örneğin, yetersiz yatak sayısı hastaların beklemesi, hizmetlerin tekrar planlanması veya hatta hastalarda ölümcül sonuçlara neden olabilir. Yatakların etkisiz kullanımı da hastane maliyetlerini artırabilir ve doktorlar ile hemşireler arasında dengesiz iş yüküne neden olabilir (Wright ve Bretthauer, 2010). YBÜ yataklarının fazla olması hastane bütçesi ve alanı gibi değerli kaynakların israfına yol açabilir. Yoğun bakım

hizmetleri, hastanenin toplam bütçesinin yaklaşık %25'ini oluşturur ve diğer hastane yatışlarına göre 2,5 kat daha maliyetlidir (Antmen ve Oğulata, 2013). Bu nedenlerden dolayı hastane yöneticileri, verimli ve tatmin edici bir şekilde sağlık hizmeti sunabilmek için yeterli sayıda yoğun bakım yatağı belirlemek için çalışırlar. YBÜ yataklarının yönetimi, hastane kaynaklarının etkili kullanımı için kritik öneme sahiptir ve doğru yatak yönetimi sayesinde hastaların gereksinimleri karşılanabilir, maliyetler düşürülebilir ve hemşireler ile doktorlar arasında iş yükü dengesi sağlanabilir. Sonuç olarak, hastane yöneticilerinin etkili bir şekilde yönetilen bir YBÜ'ye sahip olmaları gerekir. Yatak sayısı, hastane performansını doğrudan etkileyen en önemli kaynaklardan biridir.

YBÜ'lerde, hizmet verimliliğini sürdürmek için %10-15 yatak boşluğu önermektedir (Horrocks, 1986). YBÜ doluluğu için optimal hedef literatürde kesin olarak tanımlanmamıştır; ancak %70-%75 aralığında belirlenen bir çok çalışmaya rastlanmıştır (Andreas ve Patrick, 2011). YBÜ yatak talebinin doğru planlanması, düşük maliyetli yüksek verimli sistemin oluşmasını sağlar. Bu nedenle yatak kapasitesinin belirlenmesi gerek halk sağlığı gerekse yönetim açısından önemli bir parametredir.

Bu çalışmada Pediatri ve Reanimasyon yoğun bakım üniteleri analiz edilmiştir. Pediatri YBÜ, çocuklara yönelik sağlık hizmetleri sunan ve onların özel gereksinimlerini karşılamak için tasarlanmıştır. Bu üniteler, çocuklarda görülen ciddi sağlık sorunları, yaralanmalar, enfeksiyonlar, hastalıklar veya diğer acil durumlarda gerekli tıbbi bakımı sağlamaktadır. Sağlık Bakanlığına göre 28 günden büyük ve 18 yaşını doldurmamış tüm bireyler çocuk olarak tanımlanmaktadır (T.C. Sağlık Bakanlığı, 2008).

Reanimasyon YBÜ, kritik durumda olan hastaların hayatlarını kurtarmak için gereken tıbbi müdahaleleri sağlayan birimlerdir. Bu ünitelerde, kalp krizi, solunum problemleri, travmalar veya diğer acil durumlar gibi kritik sağlık sorunları olan hastaların takibi ve tedavisi yapılmaktadır. Reanimasyon YBÜ, hayat kurtarıcı müdahalelerin yanı sıra, hastaların stabilizasyonu için de gerekli tıbbi bakımı sağlamaktadır.

Sağlık Bakanlığı'nın yayınladığı Yoğun Bakım Ünitelerinin Standartları Yönergesi'ne göre YBÜ'lerin bilinmesi gereken özellikleri şu şekilde sıralanabilir (T.C. Sağlık Bakanlığı, 2008).

- YBÜ'ler, yatak kapasitesi, kabul edilecek hastaların özellikleri ve klinik durumları, bulunduğu fiziksel koşullar, uzmanlık dalları ve uzman hekim sayısı, diğer sağlık personelinin sayısı ve niteliği, tıbbi ekipman ve cihaz standartları ve hizmet verdiği sağlık tesisinin statüsü gibi faktörler göz önünde bulundurularak seviyelendirilmektedir.
- YBÜ'ler, erişkinler için birinci, ikinci ve üçüncü seviyelerde; çocuklar için ikinci ve üçüncü seviyelerde; yenidoğanlar için ise birinci, ikinci, üçüncü ve dördüncü seviyelerde seviyelendirilmektedir. Dördüncü seviye yenidoğan yoğun bakım servisleri, seviye IV A ve seviye IV B olarak ayrı tescil edilmelidir.

- Birinci basamak YBÜ, ilk basamak olarak kabul edilen üniteler, temel yoğun bakım hizmetlerini sunmaktadır. Bu hizmetler arasında hastanın vital fonksiyonlarını takip etmek, oksijen tedavisi uygulamak, ilaç tedavisi sağlamak ve basit mekanik ventilasyon uygulamak gibi hizmetler yer almaktadır. Bu ünitelerde hastaların durumu istikrarlıdır ve organ yetmezliği gibi durumlar söz konusu değildir. Bu basamakta en az iki yatak olmalıdır.
- İkinci basamak YBÜ, birinci basamak ünitelerin sunmuş olduğu hizmetlere ek olarak, daha ileri düzeyde tıbbi bakım sağlamaktadır. Bu ünitelerde hastaların durumu ciddidir ve hayat fonksiyonları yoğun bir şekilde izlenmektedir. Bu ünitelerde yoğun bakım ventilatörleri, kan gazı analiz cihazları, kalp monitörleri ve invaziv işlemler için ekipmanlar mevcuttur. Bu basamakta tavsiye edilen en düşük yatak sayısı dört olarak belirlenmiştir ve her vardiyada, dört yatak için en az bir hemşire bulundurulmalıdır.
- Üçüncü basamak YBÜ, en ileri düzeyde yoğun bakım hizmetlerini sunarlar ve hastaların durumu kritiktir. Bu ünitelerde genellikle ağır travma, yanık, kalp ve akciğer transplantasyonları gibi ciddi ameliyatlardan sonrası hastalar kabul edilmektedir. Bu ünitelerde çok sayıda medikal cihaz ve ekipman, uzman tıbbi personel ve özel bakım teknikleri kullanılmaktadır. Bu basamakta en az altı yatak ve her vardiyada, üç yatak için en az bir hemşire bulundurulmalıdır.
- YBÜ'ler, hastaların güvenliğini sağlamak için sağlık tesisindeki diğer birimlerden, genel kullanım alanlarından, hasta ve ziyaretçi trafiğinden ayrı bir şekilde tasarlanmalıdır. Bu servisler, tercihen asansör, ameliyathane, acil servis, laboratuvar ve görüntüleme birimlerine yakın konumlandırılmalıdır. Böylece, yoğun bakım servisine acil bir şekilde ulaşılabilirken hastaların ve personelin güvenliği de korunmuş olur.
- Hastane içerisinde, YBÜ'ye yakın bir konumda, hasta yakınlarının bilgilendirilmesi, görüşmeleri ve beklmeleri için yeterli büyüklükte uygun bir alan düzenlenmelidir. Bu görüşmeler, ilgili hekim veya sağlık personelinin talebi üzerine, bir güvenlik görevlisi refakatinde gerçekleştirilmelidir. Bu sayede, hasta yakınlarına konforlu bir bekleme alanı sağlanırken, hastane personeli ve hasta güvenliği de korunmuş olur.
- YBÜ'lerde, hasta alanları dışında, sık kullanılan malzemeler, gerekli ilaçlar ve serumlar için özel bir depo alanı ayrılmalıdır. Bu depo alanı, doğrudan güneş ışığından korunacak şekilde düzenlenmeli ve malzemelerin kolayca muhafaza edilmesi ve erişilmesi sağlanmalıdır. Böylece, yoğun bakım servisindeki tüm malzemelerin, ilaçların ve serumların doğru koşullarda muhafaza edilmesi sağlanırken, hasta bakımı ve tedavisi de en iyi şekilde gerçekleştirilebilir.

- YBÜ'lerde, eczaneden temin edilemeyen ve servis içinde hazırlanması gereken ilaçlar ve infüzyonlar için özel bir hazırlama alanı ayrılmalıdır. Bu alan, servisin destek alanı içinde ayrı bir bölüm oluşturacak şekilde düzenlenmeli ve ilaçların hazırlanması için gerekli tüm ekipmanlar ve malzemeler burada bulundurulmalıdır.
- Lavabo alanında en az bir adet, hasta alanında her altı yatağa kadar en az bir adet bulunacak şekilde genişlik ve derinlikleri uygun olan, suyun çevreye sıçramasını ve göllenmesini önleyecek özellikte sabun ve kâğıt havlu bulunduran lavabo konumlandırılmalıdır. Ayrıca sağlık çalışanlarının kolayca ulaşabileceği el antiseptiği de bulundurulmalıdır.
- Duvar ve tavan yüzeyleri için, kolay temizlenebilir ve parçacık dökülmeyen özellikte malzeme veya açık renkli boya tercih edilmelidir. Duvar yüzeyleri, hareketli cihazların ve eşyaların çarpma noktalarında korumalı olacak şekilde tasarlanmalıdır.
- Zemin yüzeyleri için, kolay kırılmayan ve mat yüzeyli malzemeler tercih edilmelidir. Ayrıca kaymayı önleyici, sık temizlemeye uygun, mikroorganizma üremesini en aza indiren, kimyasallara ihtiyaç duyulmadan ve kolayca temizlenebilen, antistatik özellikte malzemeler kullanılmalıdır.
- Merkezi havalandırma sistemi bulunan veya bulunması gereken her yoğun bakım servisi seviyesinde, pencerelerin açılabilir özelliğinin bulunmaması kaydıyla, hasta alanlarının gün ışığı alması sağlanmalıdır.
- YBÜ'lerde, hastaların sürekli gözetim ve izlenmesine uygun olarak, merkezi bir alanda bir bölüm ayrılmalıdır.
- YBÜ'lerde, izolasyon odaları dahil her yatak için ayrı olmak üzere zeminde bulunmayan ve hasta başında en az oniki çıkışlı elektrik paneli bulunmalıdır. Ancak birinci seviye yoğun bakım servislerinde, en az dört çıkışlı elektrik paneli bulundurulmalıdır. Tüm YBÜ'lerde birinci seviye dışında en az iki basınçlı hava çıkışı, iki oksijen çıkışı ve iki vakum sistemi bulundurulmalıdır.
- YBÜ'lerdeki musluklar, dirsek veya ayakla kontrol edilebilir veya otomatik açılır-kapanır özellikte olması tercih edilmelidir.
- Koroner, dahili ve cerrahi YBÜ yatakları, aynı seviyede bulunmak ve seviyesine uygun asgari şartları taşımak kaydıyla, aynı fiziksel alanda birleştirilebilir. Ancak, kardiyovasküler cerrahi YBÜ yatakları bu durumdan hariç tutulmaktadır. Bu şekilde yapılandırılan genel YBÜ, toplam yatak sayısı olarak kabul edilmekte ve kaydedilmektedir.
- Erişkin YBÜ'lerde destek alanları hariç, yataklar arasındaki mesafe en az 1,5 m pediatri YBÜ'lerinde destek alanları hariç, yataklar arasındaki mesafe en az 2 m olacak şekilde ayarlanmalıdır. Hem erişkin hem de pediatri YBÜ'lerde her yatak için en az 12 m² alan ayrılması gerekmektedir.

YBÜ yatış kriterleri, hastanın klinik durumuna, tedavi gereksinimlerine ve izlem gereksinimine bağlı olarak belirlenmektedir. YBÜ’de yatan hastalar, hayati fonksiyonlarının sürekli izlenmesi, gerekli tedavi ve müdahalelerin hızlı bir şekilde yapılması gereken kritik durumlardadır. Bu nedenle, YBÜ yatış kriterleri genellikle şu durumlarda uygulanmaktadır:

- Solunum yetersizliği veya solunum sıkıntısı yaşayan hastalar
- Ciddi kardiyovasküler hastalıklar (örneğin kalp krizi, kalp yetmezliği, aritmi vb.)
- Yoğun bakım gerektiren ciddi enfeksiyonlar (örneğin sepsis, pnömoni vb.)
- Ciddi travma veya yanık sonrası takip ve tedavi gerektiren durumlar
- Yoğun bakım gerektiren ciddi nörolojik bozukluklar (örneğin inme, nöbet vb.)
- Ciddi böbrek veya karaciğer fonksiyon bozuklukları
- Cerrahi müdahale sonrası veya ağır bir ameliyat geçirme riski olan hastalar (Köroğlu ve ark. 2006).

YBÜ yatış kriterleri, hastanın durumunun kritik olup olmadığına ve tedavi gereksinimlerinin YBÜ’de karşılanması gerekip gerekmediğine bağlı olarak değerlendirilmelidir. Bu karar, hastanın klinik durumunu değerlendiren bir uzman hekim tarafından verilmektedir.

Sağlık sistemleri, özellikle hastaneler, karar verilmesi gereken pek çok faktör olması nedeniyle oldukça sık çalışılan bir alandır. Endüstri Mühendisliği alanında çalışmaların önemli bir kısmını simülasyon modelleme oluşturur.

1.3. Simülasyon Modelleme

Simülasyon modelleme, gerçek dünya olaylarının matematiksel bir modelini oluşturarak, farklı senaryoların sanal ortamda test edilmesine izin veren bir tekniktir. Kapasite planlaması için kullanıldığında, bir organizasyonun iş yükünü, kaynaklarını ve operasyonel süreçlerini modellemek ve optimize etmek için kullanılabilir. Simülasyon modelleme, kapasite planlaması gibi karmaşık problemlerin çözümünde son derece yararlı bir araçtır, çünkü gerçek dünya verisini kullanarak daha iyi bir karar verme süreci sağlamaktadır.

Hastane yönetimleri, YBÜ’lerdeki yatak kapasitesinin artırılması veya azaltılması gerektiği konusunda bir karar alması gerektiğinde simülasyon modelleme kullanılarak, mevcut yoğun bakım ünitelerinin iş yükü, bekleme süreleri, doluluk oranları, personel sayısı ve diğer faktörler gibi bir dizi değişken modellenenir. Daha sonra, farklı senaryolar test eden bir sanal ortamda değişiklikler yapılarak, kapasite planlamasında en iyi sonucu sağlamak için hangi değişikliklerin yapılması gerektiğini belirlenebilir.

Yeterli sayıda YBÜ yatağını belirlemek için genellikle matematiksel modelleme, simülasyon modelleme ve kuyruk teorisi gibi nicel analiz teknikler kullanılır. Özellikle hastane birimlerinde tıbbi personel düzeylerinin ve sağlık malzemelerinin belirlenmesi için simülasyon modellemesi birçok farklı durum ve senaryoda kullanılmaktadır (Lee ve ark., 2017; Buschiazzi ve ark., 2020; Kokangül ve ark., 2017). Yararlı bir simülasyon modeli geliştirmek için sistemin bazı parametrelerinin doğru bir şekilde tahmin edilmesi gerekmektedir. Örneğin, bir YBÜ’de optimal yatak sayısını belirlemek için hasta geliş oranları ve hastaların sağlık sisteminde kalış süreleri tahmin edilmelidir. Ancak bu veri genellikle çarpık bir yapıya sahiptir ve bu nedenle saf bir istatistiksel dağılımla modellenmesi güç olabilmektedir. Çalışmanın bu aşamasında, istatistiksel olarak çarpık olan verinin simülasyon modeli için gereken girdi parametrelerinin karma dağılımla modellenmesi fikri ortaya atılmıştır.

1.4. Karma Dağılımların Kullanımı

Hastaların YBÜ’ye gelişler arası süreleri (GAS) ve hastaların kalış süreleri (SS) rasgele gerçekleşen stokastik olaylardır ve bu veri çarpık bir yayılım gösterebilmektedir. Bu nedenle, bazı durumlarda saf bir istatistik dağılıma uymayan veri setleri ortaya çıkmaktadır. SS, hastanın hastaneye ilk kabulünden taburculuk tarihine kadar geçen süre olarak tanımlanmaktadır ve bakım hizmetlerinin kalitesini, hastane performansını ve hastane kaynak tüketimini değerlendirmek için temel göstergelerden biri olarak kabul edilmektedir (Benzeval ve Judge, 1994; Shorten, 1995; Burchardi ve Moerer, 2001). YBÜ’de hastanın SS’i hastanın durumuna bağlı olarak uzun veya kısa olabilmektedir. Hastanın durumu kritik ise daha uzun süre yoğun bakımda kalmaktadır. Hastanenin operasyonel performansının kalitesini ve yatak kapasitesi, ekipman, personel ve zaman gibi kısıtlamaları yönetmek için hastaların kalış ve gelişler arası süresinin doğru bir şekilde belirlenebilmesi ve tahmin edilmesi gerekmektedir. Hastane yönetimi ve sigorta şirketleri bütçelerini yönetmek için bu tahmini strateji geliştirmek için kullanabilmektedir.

Kuyruk modelleri, bekleme hatları, mağazacılık, trafik planlaması ve iletişim sistemleri gibi kaynakları tahsis etmek ve yönetmeye yardımcı olmak için birçok alanda kullanılmaktadır. Bir YBÜ üç unsurdan oluşan bir kuyruk modelidir: gelişler arası süre veya geliş oranı (GAS), kalış veya hizmet süresi (SS) ve sunucu sayısıdır (Zhu ve ark., 2012). Gelişler arası zaman, kalış süresi istatistiksel olarak modellendikten ve yatak kapasitesi olan sunucu sayısı belirlendikten sonra, kuyruk teorisinde çıktı sonuçları hesaplanmaktadır (Gorunescu ve ark., 2002). Kuyruk modellerinin doğruluğu, iyi tahmin edilmiş girdi verisine bağlıdır. Bununla birlikte, simülasyon modelinde girdi parametreleri olarak kullanılan gelişler arası süre ve kalış süresinin saf istatistiksel dağılıma uydurmak zordur, çünkü veri genellikle aşırı derecede çarpıktır (McClean ve Millard, 1993).

Çarpık veri modellenmesi istatistiksel analizlerde sıkça karşılaşılan bir durumdur. Bu veri normal dağılıma uygun olmadıklarından ve simetrik olmayan bir yapıya sahip olduklarından, modelleme işlemi zorlu hale gelmektedir. Bu noktada karma dağılımların kullanımı önem

kazanmaktadır. Karma dağılımlar, iki ya da daha fazla dağılımın kombinasyonu ile oluşan ve çarpık veri modellenmesinde kullanılan bir istatistiksel dağılım türü olarak tanımlanabilir. Bu dağılımlar, simetrik olmayan dağılımların modellenmesinde etkili bir araçtır. Karma dağılımların kullanımı, hem modelin doğruluğunu arttırmakta hem de veri analizinde daha doğru sonuçlar elde edilmesini sağlamaktadır. Bu nedenle, çarpık veri modellenmesinde karma dağılımların kullanılması istatistiksel analizlerin doğruluğunu arttırmak için sıklıkla tercih edilmektedir.

Son zamanlarda karma dağılım yaklaşımı veri temsil etmek ve talep tahmini için kullanılmaktadır. Karma dağılımlar zaman serisi modellerinin temsilde iyi bir araç olarak tanımlanmakta ve kısa ila orta vadeli tahmin ve eksik değer tahmini için doğrudan uygulanabilmektedir (Eirola ve Lendasse, 2013). Karma dağılım modellemesi, temelde çalışılan popülasyonun grup yapısını tanımlamaktır ve kümelenme analizi gibi belli başlı istatistik alanlarında esnek ve güçlü bir olasılık yaklaşımı sağlamaktadır (Ng ve ark., 2019). Küme analizi bağlamındaki uygulamalarda, varlıkların birkaç kümeye gruplandırılması, karma dağılım aracılığıyla elde edilmektedir. Böylece kümeler arası benzerliğin en az, küme için benzerliğin en çok olduğu veri grupları elde edilmeye çalışılır. Karma dağılım temelli kümelenme yaklaşımları, gözlenen özellik verisi, her bir varlık ile ilişkili değişkenler temelinde, temel popülasyonun grup yapısı hakkında önceden bir bilgi olmadan ölçülmek suretiyle gerçekleştirilir. Bu durum, denetimsiz sınıflandırma olarak adlandırılır. Karma modelleri sadece kümelenme için sağlam bir matematiksel çerçeve sağlamakla kalmamakta, aynı zamanda bileşen sayısının belirlenmesi ve sınıflandırma belirsizliğinin değerlendirilmesine de istatistiksel çıkarımlar sağlamaktadır (McLachlan ve Peel, 2000). Uygulamalı istatistik, gözlemlenen verinin tek bir popülasyonda toplandığı varsayımına dayanmaktadır. Bu homojenlik varsayımı, çalışılan verisi doğru şekilde temsil etmek için çoğu zaman gerçekçi ve yeterli değildir. Bu noktada, karma dağılım modelleri, farklı parametrelere sahip heterojen alt popülasyonların varlığını dikkate alan bir yaklaşım sunduğundan veri temsilde önemli rol oynamaktadır. Maliyet etkinliği analizleri daha küçük veri kümeleriyle çalışma eğilimindedir ve parametrik modelleme kapsamı daha sınırlıdır. Son zamanlarda, lineer regresyon yerine doğrusal olmayan yöntemler kullanılmaya başlanmış ve bunların arasında sonu karma dağılımlar ve semiparametrik yaklaşımlar bulunmaktadır (Jone, 2010).

Esneklikleri nedeniyle, çarpık dağılımlar ve rasgele olaylar, karma dağılımları ile modellenebilir. Karma dağılım bileşenleri, yönetimin yatarak tedavi üzerindeki baskıyı takip edebilmesi için kısa, orta veya uzun vadeli kalış süresini tanımlayabilir. Böylece tahmin modeli klinik karar vermeyi geliştirmeye yardımcı olabilir. Tahmin edilen karma dağılım model parametreleri, veri bilinen herhangi bir istatistiksel saf dağılıma uymadığında ampirik dağılımlar yerine girdi parametreleri olarak simülasyon modellerinde kullanılabilir.

Simülasyon modellerinde karma dağılım uygulanması literatürde yaygın olmayan bir yöntemdir. Esneklikleri sayesinde, çarpık dağılımlar ve rasgele olaylar, karma dağılımlar ile

modellenebilir. Çarpık dağılımlardan oluşan verinin tek bir parametrik model ile modellenmesi zor olabilir ancak karma dağılım, bu tür verinin başa çıkmada oldukça etkili olmaktadır (Papi ve Pontecorvi, 2015). Literatürde karma dağılımlar kimya, mühendislik, genetik, tıp, fizik, kriminoloji, psikoloji, eğitim, ekonomi gibi sağlık alanı dışında da sıkça kullanılmaktadır. Sağlık alanında; anatomi, biyoinformatik, hücre biyolojisi, kronik hastalıklar, genetik, geriatri, bulaşıcı hastalıklar, görüntü işleme, ortopedi, farmakoloji, beslenme, jinekoloji, psikiyatri gibi alanlarda karma dağılımlar kullanılmıştır (Ng ve ark., 2019).

1.5. Problemin Tanımı

Heterojen veri yapıları hayatımızın her aşamasında olmasına rağmen, karma dağılımlara özellikle planlama konusunda tercih edilmemektedir. Aynı zamanda kurulan simülasyon modelinde de, girdi parametresi karma dağılımlar pek kullanılmamaktadır. (Veronica ve Bauer, 2016). Bunun sebebi karma modellerin parametre tahmininde kişi odaklı olduğu yanılgısıdır (Bergman ve Magnusson, 1997). Fakat çalışmalar talep tahmininde karma dağılımların oldukça iyi bir uyum gösterdiğini ortaya koymuştur. Prosedürler öğrenildiği zaman uygulaması oldukça kolay olan bu yaklaşımla aynı grup içerisindeki farklı kümeler başarılı bir şekilde temsil edilebilir ve bu sayede sistem daha rahat bir şekilde yönetilebilir. Örneğin, bir hastanede ortalama kalış süresi tespit edilmek istendiğinde, ayakta tedavi olan ile yatan hastaların kalış süresinin farklı dağılımlara sahip olması sebebiyle karma dağılım kullanılması talep tahmini performansını olumlu yönde etkileyecektir. Bu sayede ayakta tedavi gören hasta için bu küme için belirlenmiş parametreler kullanılarak talep tahmin yapılacaktır ve poliklinik planlamasında bu parametreler dikkate alınmayacaktır ya da yatak kapasite planlaması yapılırken yatan hasta için belirlenen parametreler göz önünde bulundurulacaktır. Böylece hem ekonomik hem de sosyal açıdan avantaj sağlanacaktır.

Bu çalışmada ele alınan ve çözülmesi hedeflenen Endüstri Mühendisliği problemi, analiz edilen YBÜ'lere gelen hastaların uzun bekleme süreleridir. Bu kapsamda; bu tezde hedef, YBÜ'ye gelen tüm hastalara bekleme süresi olmadan hizmet verebilmek için ilgili birimdeki yatak sayısını belirlemektir. Ancak hem etik açıdan hem de YBÜ'de verilen hizmetlerin karmaşıklığı nedeniyle bu alanda çalışmak oldukça karmaşıktır (Akalin, 2002). Ele alınan kapasite problemi simülasyon modeliyle çözmek istenmektedir. Simülasyon modelinde kullanılacak olan hastaların gelişler arası süresi ve servis süresi oldukça çarpık dağılımlar olduğu için karma dağılım yaklaşımıyla bu girdi parametreleri belirlenmektedir.

1.6. Çalışmanın Amacı ve Kapsamı

Sağlık sistemlerinde gerçek veri ile çalışırken, etik kurallar nedeniyle dikkatli olunması gerekmektedir. Hasta mahremiyeti sebebiyle sağlık sistemindeki gerçek veriye ulaşmak özel izne bağlıdır. Bu nedenle, sağlık sistemlerinde gerçek veriyle yapılmış çalışmalar sınırlıdır.

Karar vericiler, dinamik sorunları çözmek için genellikle nicel analiz tekniklerini kullanmaktadır. Simülasyon modelleme, kısıtlamalarla kapasite planlamasına kaynak sağlamak için en iyi yöntemlerden biri olarak tanımlanır (Ojstersek ve Buchmeister, 2021). YBÜ, varışlar arası süre veya varış oranı; kalış veya servis süresi ve sunucu sayısı olmak üzere üç girdi unsurdan oluşan bir kuyruk modelidir (Zhu ve ark., 2012). Bu girdi unsurları belirlendikten sonra kuyruk teorisinde sistemdeki ve kuyruktaki hasta sayısı, bir hastanın sistemdeki ve kuyruktaki toplam bekleme süresi, sistemin kullanım oranı gibi çıktı sonuçları hesaplanabilir (Gorunescu ve ark., 2002). Kuyruk modellerinin doğruluğu, iyi tahmin edilmiş girdi verisine bağlıdır. Bununla birlikte, simülasyon modelinde girdi parametreleri olarak kullanılan varışlar arası sürenin ve servis zamanının tek fazlı istatistiksel dağılımını uydurmak zordur, çünkü veri genellikle rasgele, oldukça sağa çarpık ve uzun kuyruklu dağılımlardır (McClean ve Millard, 1993).

Gelişler arası süre (GAS) ve servis süresini (SS) modellemek için yaygın olarak kullanılan Weibull, gamma ve lognormal dağılımları bu çarpıklığı ve kuyruk davranışını temsil edebilir (Marazzi ve ark., 1998). Bununla birlikte, bu saf istatistiksel dağılımlar, önemli sayıda aykırı değere sahip hasta popülasyonlarının GAS ve SS verisine uymamaktadır (Lee ve ark., 2001). Bu gibi durumlarda, genellikle ampirik dağılımı tercih edilmiştir. Ampirik dağılımlar, saf istatistiksel dağılımlardan daha az bilgi içermektedir. Bu sınırlamanın üstesinden gelmek için, yaklaşık ampirik dağılım yerine karma dağılımı önerilmiştir (Atienza ve ark., 2008; Papi ve ark., 2016). Sağlık hizmetlerinde yaklaşık ampirik dağılımlara bir alternatif olarak karma dağılım modellerinin önerilen kullanımına rağmen, yalnızca az sayıda araştırmacı, herhangi bir saf istatistiksel dağılıma uymayan veri ile sağlık sistemlerini analiz etmek için bu yöntemi kullanmıştır (McClean ve Millard, 1993).

Hastaların YBÜ’de GAS ve SS’nin çok değişken olması; yatak yönetimi, ameliyathane, tıbbi-cerrahi birimleri, personel ve zaman kaynaklarının yetersiz planlamasına yol açmaktadır. Hastaların gelişler arası süresinin ve kalış sürelerinin genellikle çeşitli faktörlere bağlı olan oldukça çarpık veri olması nedeniyle istatistiksel olarak kolaylıkla modellenememesi temel problemidir.

Bu çalışmada, oldukça çarpık ve uzun kuyruklu GAS ve SS dağılımlarının analizi için karma dağılım modellerinin kullanımı önerilmiştir. Bu karma dağılım parametreleri oldukça esnektir ve aykırı değerlerin silinmesini veya veri dönüştürülmesini gerektirmez. Karma dağılım modelindeki dağılımların parametrelerinin hesaplanması, Newton-Raphson veya beklenti maksimizasyon (EM) algoritması gibi karmaşık maksimum olabilirlik tahmin algoritmaları gerektirmesine rağmen, bu girdi parametrelerinin modellenmesi için oldukça yararlı bir yöntem olduğu ortaya konulmuştur.

YBÜ’ye gelen hastaların acil ve gelişmiş sağlık desteğine ihtiyacı olduğundan, bu tez; sisteme gelen tüm hastalara bekleme süresi olmadan hizmet verebilecek YBÜ yatak sayısını bulmayı amaçlamaktadır. Bu tezdeki amaç, genellikle oldukça çarpık olan ve önemli sayıda aykırı değer içeren bu girdiler için karma dağılım yaklaşımını kullanarak kapsamlı parametreler

geliştirmek ve bu tahmin edilen karma dağılım parametrelerini simülasyon modeline entegre ederek istenen hedefteki yatak kapasitesinin bulunmasıdır.

Özetle, bu tezin araştırma hedefleri aşağıdaki gibidir:

- Analiz edilen iki yoğun bakım ünitesindeki (Pediatri ve Reanimasyon) sürecin simüle edilmesi;
- Karma dağılım yaklaşımını kullanarak girdi parametrelerinin tahmin edilmesi ve bu parametrelerin simülasyon modeline entegre edilmesi;
- Hastalar için bekleme süresi olmaksızın hizmet alacağı yatak sayısının bulunması.

1.7. Çalışmanın Orijinal Katkısı

Gelişler arası süre (GAS) ve servis/kalış süresindeki (SS) çarpıklığın sınırlamaları göz önünde bulundurularak, pediatri ve reanimasyon YBÜ'lerin hasta akış simülasyonunda, girdi parametreleri olarak karma dağılım modeli kullanılmıştır. Karma dağılımlar, esneklikleri sayesinde, çarpık dağılımlar ve rasgele olayları istatistiksel olarak yüksek doğrulukla tahmin edebilmesi sebebiyle simülasyon modelinde temel parametreler olan gelişler arası süre ve servis süresinin tanımlanmasında kullanılması, bu çalışmanın temelini ve orijinal katkısını oluşturmaktadır. Bu tezin ana katkısı, gerçek hayat verisini geleneksel istatistiksel dağılımlardan daha iyi temsil etmek için girdi parametreleri olarak bir karma dağılım modeli kullanan bir simülasyon yaklaşımı kullanarak, gelen tüm hastalara bekleme süresi olmadan hizmet verebilmek için gereken yatak kapasitesini tahmin eden kapsamlı bir modelin geliştirilmesidir. Böylece tek fazlı dağılımlara göre, daha doğru parametreler belirlenerek simülasyon modelinin geçerliliği arttırılacaktır. Pediatri ve Reanimasyon YBÜ'lerde GAS ve SS'nin en uygun parametrelerini tahmin etmek için dört farklı sürekli karma dağılım modeli (normal, Weibull, gamma, lognormal) uygulanmış ve karşılaştırılmıştır. Ayrıca bu tezde kullanılan tüm verinin orijinal olması, kurulacak olan modelin gerçek durumu ortaya koymasında önemli bir rol oynamaktadır.

1.8. Tezin Yapısı

Bu tezin ikinci bölümünde Önceki Çalışmalar başlığı altında literatür incelenecektir. Üçüncü bölümde tezin hedefi, kapsamı ve metodoloji dahil olmak üzere araştırma tasarımı hakkında ayrıntılı bilgi verilecektir. Ardından dördüncü bölümde elde edilen çıktı bulgular sunulacak ve tartışılacak, son bölüm olan beşinci bölümde tezin nihai sonuçları ve çıktıları, kısıtlamaları paylaşılabilecek ve gelecek çalışmalar için önerileri sunulacaktır.

2. ÖNCEKİ ÇALIŞMALAR

Talep tahmini ile ilgili çalışmalar, 1960'lı yıllardan günümüze sürekli gelişmektedir (Ergün ve Şahin, 2017). Hizmet sistemlerinde talep tahmin, personel ve kapasite planlaması gibi alanlarda çalışmalar da literatürde sıkça çalışılmaktadır. En çok kullanılan tahmin metotları; zaman serisi analizi, ekonometrik modeller, regresyon analizi ve yapay sinir ağları modelleri olarak sıralanabilir (Akar ve Uğurlu, 2005). Mevcut sağlık sistemlerinin sürdürülebilir olmaması endişesiyle karar alma sistemleri geliştirilmeye çalışılmaktadır. Bu sayede altyapı ve insan gücü planlama kararları alınırken bu sistemlerden yararlanılabilir. Model karmaşıklıkla geleneksel tahmin yöntemleri ve varsayımlar yetersiz kalmaktadır. Planlama yapılırken en önemli hususlardan bir tanesi planlama yapılacak periyodun ve mevcut verinin uzunluğudur. Genelleme yapılacak olursak, mevsimsel modeller tahminlemede daha iyi performans göstermektedir. Yapılan bir çalışmada tek günlük sayım tipi modeller üretilmiştir. Bu modellerin performansı gerçek günlük verisine yakın olduğu fakat yıllık üretilen verisinden düşük olduğu ortaya konulmuştur (Mackay ve Lee, 2005).

Hastanede kalış süresi, hastane aktivitesinin ve sağlık sistemlerinin önemli bir göstergesi olarak kabul edilmektedir. Bu değişkenin gösterdiği çarpıklık, istatistiksel modellemede problemler doğurmaktadır. Literatürde, sağlık sistemlerinin talep tahmini konusunda da pek çok çalışma bulunmaktadır.

Hastanede kalış süresi ve gereken yatak sayısının tahmininin yapıldığı bir çalışmada mevsimsel modellerin daha iyi sonuç verdiği görülmüştür. Aynı çalışmada hastane için gereken yatak sayısına ortalama hastanede kalış süresiyle karar vermenin doğru bir yaklaşım olmadığı belirtilmiştir. Yatak sayısı modelleri geliştirilirken ortalama hastanede kalış süresi kullanmaktan kaçınılmasını gerektiren üç neden sıralanmıştır. Bunlardan ilki, kalış süresinin yüksek çarpıklıklara sahip olması ve normal dağılıma uymaması problemidir. İkincisi, kalış sürelerinin karmaşıklığıdır. Hastanede yatan kişiler cinsiyet, yaş, hastalık ve tedavi çeşidi karma bir yapıya sahiptir ve bu da karmaşık bir veri seti oluşturmaktadır. Üçüncüsü ise, kalış süresinin hastanenin en yoğun olduğu günün saatıyla ilgili olmayan bir ölçü olmasıdır (Markay ve Lee, 2005).

Yıllar boyunca hastane yatak planlamasıyla ilgili pek çok çalışma yapılmıştır. Ancak artan talep ve kısıtlı kaynaklar sebebiyle hastane yatak kapasitesi planlaması zorlu bir problem haline gelmiştir (Akçalı ve ark., 2006). Yatak kapasitesi planlama modeli kurulurken, karar verilmesi gereken hangi kriterin performansının amaçlandığıdır. Literatürde çoğunlukla kullanılan kriter yatakların toplam yatağa bölünmesiyle hesaplanan yatak doluluk oranıdır (Green, 2004). Ancak çalışılan zaman periyodu içerisinde ortalama bir değer olan yatak doluluk oranı, sistemi değerlendirmek için yeterli olmayabilir. Bu sebeple hizmet verilemeyen hasta sayısı, hedef hizmet seviyesi gibi farklı performans kriterleri modele eklenebilir (Kokangül, 2008). Birden fazla performans kriterinin birleştirilmesiyle toplam performans hakkında daha iyi bir genel bakış sağlanabilir.

Kuyruk teorisi müşteri hizmetlerini iyileştirmek için sağlık sektöründe sıklıkla kullanılan bir yöntemdir (Belciug ve Gorunescu, 2015). Çalışmalar incelendiği zaman pek çok araştırmacı simülasyon modelinde girdi olarak χ^2 ve Kolmogorov-Smirnov testlerinden elde edilen, bilinen saf bir dağılım kullanıldığı istatistiksel dağılıma uydurma yöntemi kullanıldığı görülmüştür (Ickowicz ve Sparks, 2016). Ancak çoğu zaman verinin çarpıklığı, istatistiksel olarak bilinen saf bir dağılıma uydurmayı zorlaştırır. Literatürde karma dağılım yaklaşımı kimya, mühendislik, genetik, tıp, fizik, kriminoloji, psikoloji, eğitim ve ekonomi gibi sağlık sistemleri dışında da sıklıkla kullanılmaktadır. Sağlık sistemlerinde; karma dağılımlar anatomi, biyoinformatik, hücre biyolojisi, kronik hastalıklar, genetik, geriatri, bulaşıcı hastalıklar, görüntü işleme, ortopedi, farmakoloji, beslenme, jinekoloji ve psikiyatri gibi alanlarda uygulanmaktadır (Ng ve ark., 2019). Bu nedenle, talep tahmini için bir uygulama olacak girdi analizinde çarpık veri setlerini daha uygun bir şekilde tanımlamak için karma dağılımları yaklaşımının uygulanması fikri bu çalışmanın temelini oluşturmaktadır.

Karma dağılımlar uzun yıllardır birçok alanda kullanılmasına rağmen talep tahmininde girdi analizinde kullanımı nadirdir. Son on yılda gelişler arası süre ve kalış süresi tahminleri için çeşitli hesaplama yöntemleri kullanılmıştır. Bazı makalelerde, araştırmacılar makine öğrenimi yöntemlerine dayalı olarak servis süresi ve gelişler arası süre tahmin edilmiştir (Verburg ve ark., 2014; Maharlou ve ark., 2018; Meyer ve ark., 2018; Singh, 2019; Wu ve ark., 2021). Özellikle sağlık sistemlerinde, ampirik modellemeye alternatif olarak karma dağılım modeli önerilmiştir (Deb ve Trivedi, 1998). Girdi parametrelerinin belirlenmesinde karma dağılımlarının kullanımı yaygın olarak kullanılmamaktadır (Veronica ve Bauer, 2016). Bir karma dağılım modelinin kullanılması, kuyruk modelinin gerçek zamanlı durumu doğru bir şekilde temsil edebildiği için simülasyon girdi parametreleri iyileştirilebilir. Karma dağılım modellerinin parametre tahmininde sık kullanılmamalarının nedenlerinden biri de insan odaklı olduklarına dair yanlış kanı olduğu ifade edilebilir (Bergman ve Magnusson, 1997).

Karma dağılım ile sağlık sistemlerinde talep tahmini yapılması konusunda da literatürde çalışmalara bulunmaktadır. Son yıllarda ortaya konan literatürde, sağlık sistemi yönetiminde karar vermeyi desteklemek için bir karma dağılım modeli kullanan modellerin umut verici sonuçları bildirmiştir. Sağlık sisteminde hizmet alan kişiler yaş, cinsiyet, sosyo-ekonomik durumu gibi heterojen bir yapıya sahiptir ve karma dağılım ile bu tip veri iyi bir şekilde temsil edilebilir. Bu modeller, kısa ve uzun süreli çalışanların olduğu işlerde kalış süresini tanımlamak için personel planlamada başarılı bir şekilde kullanılmıştır (Millard, 1988). Hastanede kalış süresi tahmininde veriyi, yatan hasta ile ayakta tedavi olan hasta olacak şekilde alt gruba ayırmak, yapılacak tahmin doğruluğunu arttırmaktadır (Xiao ve ark., 1999).

Literatürde, gelişler arası süre ve kalış sürelerinin normal, negatif binom, üstel, Weibull, Poisson ve lognormal karma dağılımlarının kullanıldığı çalışmalar taranmıştır. Amerika’da yapılan bir çalışmada hastanede yatan hastaların ilk kayıttan taburcu olana kadar olan süre, iki bileşenli

üstel karma dağılım ile temsil edilmiştir (Cleary, 1991). Başka bir çalışmada, hastanede gerçekleşen ölümü gerçekleştiren ve taburcu olan hastalar iki bileşenli üstel karma dağılım kullanılarak ifade edilmiştir (McClean ve Millard, 1993). Hastanede kalış süresini tanımlamak için log-normal, weibull ve gamma olmak üzere üç geleneksel parametrik model kullanılarak hasta gruplarının ortalama ve toplam kaynak tüketimi ölçümlerini belirlenmiş ve Weibull ve lognormal dağılımları ile sistemin başarılı bir şekilde temsil edildiği ortaya konulmuştur (Marazzi ve ark., 1998). Bir hastanede yapılan çalışmada, lösemi ve lenfoma hastaları için yüksek maliyete ve uzun hastanede kalış süresine sahip olanları tanımlamak amacıyla Weibull karma dağılımı kullanılmıştır (Quantin ve ark., 1999). Hastanede kalış süresinin modellenmesi amacıyla yapılan bir çalışmada, iki bileşenli poisson karma modeli kullanılmış ve istatistiksel olarak iyi bir uyum oluşturduğu görülmüştür (Wang ve ark., 2002). Obstetrik tanıya bağlı gruplara gamma karma modelinin uygulandığı bir çalışmada, doğum sonrası yeni doğum yapan kadınların hastanede kalış süresinin heterojenliğini modellenmiştir. Hastanede kalış süresinin farklılıkları, normal ya da sezaryen doğumdan dolayı ya da hastane türüne bağlıdır. Elde edilecek sonuçlar ile sigorta firmalarının vakaya göre kamu ya da özel sektöre ayıracakları bütçelere karar verilebilir (Lee ve ark., 2002). Bir çalışmada, hastanede kalış süresi değişkeni tanıya bağlı gruplar içinde sonlu karma dağılımlarla modellenmiştir. Modelde, tek bir bir dağılım karma modeli yerine Gamma, Weibull ve lognormal gibi asimetrik dağılımların bir karma modeli kullanılmış ve geçerliliği kanıtlanmıştır (Atienza ve ark., 2008). İsrail’de acil servis ve gelen telefon çağrılarının talep tahmininin yapıldığı bir çalışmada, 263 gün toplanan telefon çağrısı verisi 30 dakikalık parçalara bölünmüş ve gamma karma dağılım yaklaşımıyla modellenmiştir. Uyum iyiliği testleri sonucunda yapılan tahminin yüksek doğrulukta olduğu hesaplanmıştır (Maman, 2009). Hindistan’da yatan ülke çapında verisi kullanarak, üç bileşenli bir sonlu negatif binom karma modelinin, heterojen olan hastanede kalış süresi dağılımına iyi bir uyum sağladığı bulunmuştur (Singh ve Ladusingh, 2010). Hastanelerde ayakta tedavi gören poliklinik ziyaretleri için karma bir yapay zeka modeli geliştirmek amacıyla yapılan bir çalışmada, önerilen yaklaşımın literatürdeki diğer ilgili çalışmalara kıyasla yüksek doğruluğa sahip olduğunu gösterilmiş ve bu nedenle ayakta tedavi tahmin problemleri için uygun bir araç olarak değerlendirilmiştir (Hadavandi ve ark., 2012). Kanada’da yapılan bir çalışmada oldukça dinamik bir sistem olan ambulans talep planlaması günlük, aylık ve mevsimsel olarak karma dağılım yaklaşımıyla çözülmüştür. Önerilen karma dağılım modelinin yüksek istatistiksel tahmin doğruluğu sağladığı ortaya konulmuştur (Zhou ve ark., 2015). Monte Carlo simülasyon modeli ile hastanelerin eğitim kapasitesi, hastane kapasitesi, yatan hasta etkinliği ve hastanelerin eğitimindeki cerrahi bekleme listesi büyüklüğü arasındaki ilişkinin incelenmek istendiği bir çalışmada, yatan hasta etkinliği normal dağılıma uydurularak simüle edilmiştir (Antelo ve ark., 2015). Makine öğrenmesi ve geleneksel talep tahminleri yöntemleri karşılaştırmak için, bir tıp merkezinde ayakta tedavi randevu taleplerini analiz edilmiş ve hibrit (Geleneksel + Makine öğrenmesi) yöntemiyle, makine öğrenmesi yöntemi karşılaştırılmıştır. Hibrit yöntemin, hem

geleneksel hem de makine öğrenme yöntemlerinin güçlü yönlerini kullanarak oluşması sebebiyle, birçok farklı veri kümesi türünde iyi performans gösterdiği ortaya konmuştur (Klute ve ark., 2019).

Bu literatür çalışmaları ışığında, Pediatri ve Reanimasyon YBÜ'lerine gelen hastaların GAS ve bu ünitelerde SS verisini karma dağılım yardımıyla modellenmesi ve kurulan simülasyon modeline girdi parametresi olarak tanımlanması amaçlanmaktadır.



3. MATERYAL VE METOT

Materyal ve Metot bölümü, bir araştırmanın en önemli kısımlarından biridir ve araştırmanın amacına ulaşmasında kilit bir rol oynamaktadır. Bu bölümde, çalışmanın materyalleri ve araştırma yöntemleri ayrıntılı bir şekilde açıklanacaktır.

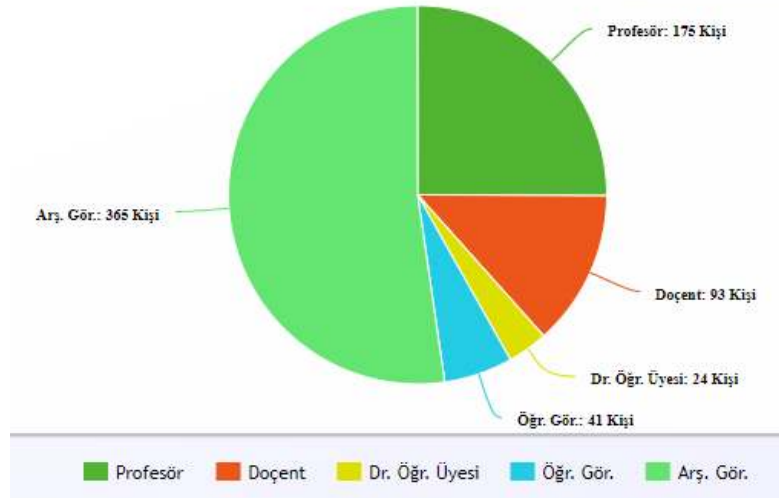
Materyal kısmında, veri toplama araçları ve ortamı gibi çalışmanın fiziksel bileşenleri tanımlanırken, metot kısmında araştırma sürecinde izlenen adımlar, veri toplama yöntemleri, analiz teknikleri ve istatistiksel yöntemler gibi çalışmanın yürütülmesinde kullanılan yöntemler açıklanacaktır. Bu bölüm, çalışmanın güvenilirliği, geçerliliği ve tekrarlanabilirliği açısından da son derece önemlidir ve çalışmanın yöntemleri hakkında ayrıntılı bilgi vererek, araştırmanın sonuçlarının değerlendirilmesine yardımcı olmaktadır.

Yoğun bakım üniteleri (YBÜ), kritik hastalık veya travma sonrası yaşamsal fonksiyonları desteklemek için tasarlanmış özel birimlerdir. Bu ünitelerde, yaşamı tehdit eden durumlar için acil tıbbi müdahaleler yapılır ve hastaların kritik parametreler yakından takip edilir. YBÜ'ler, özellikle ağır hastalık veya travma geçiren hastalar için hayati öneme sahiptir. Ayrıca, YBÜ'ler, hastaların ağrı, enfeksiyon ve diğer komplikasyonları gibi sorunlarının kontrol altına alınmasına da yardımcı olunan birimler olarak tanımlanabilir. YBÜ'lerde çalışan tıbbi personel, özellikle ağır hasta bakımı konusunda özel eğitim almış uzmanlar olmakla birlikte, kritik durumlarda hızlı bir şekilde müdahale eden ve hastaların yaşamsal fonksiyonlarını koruyan kişilerdir. Bu nedenle, YBÜ'lerin kalifiye elemanla olan varlığı, tıbbi müdahalelerin kalitesini ve hastaların sağlık durumunu olumlu yönde etkilemektedir.

Adana, Türkiye'nin güneyinde bulunan en önemli kentlerden birisidir ve pek çok sağlık kuruluşuna ev sahipliği yapmaktadır. Bunlardan en önemlisi Çukurova Üniversitesi Tıp Fakültesi bünyesinde yer alan Balcalı Eğitim ve Araştırma Hastanesi'dir. Üniversitenin yayınladığı 2022 yılı birim faaliyet raporuna göre, 1200 yatak kapasitesiyle bölgenin en büyük hastanesi unvanını elinde bulundurmaktadır. Hastane 130 bin m²'ye yakın bir alana kurulmuş olup 1987 yılından günümüze hizmet vermektedir. Hastanenin bünyesinde 698 kadrolu akademik personel görev almaktadır.

Şekil 3.1.'de Çukurova Üniversitesi Balcalı Eğitim ve Araştırma Hastanesi akademik personelin istidam şekline ilişkin bilgiler paylaşılmıştır. Akademik personelin yarısından çoğunu Araştırma Görevlileri oluşturmaktadır. Onları Profesörler ve devamında Doçent'ler takip eder. En düşük orandaki akademik personel Doktor Öğretim Üyesi (Dr. Öğr. Üyesi) olan personellerdir.

Akademik personelin istihdam şekline ilişkin bilgiler
Çukurova Üniversitesi Tıp Fakültesi 2022 Yılı Birim Faaliyet Raporu



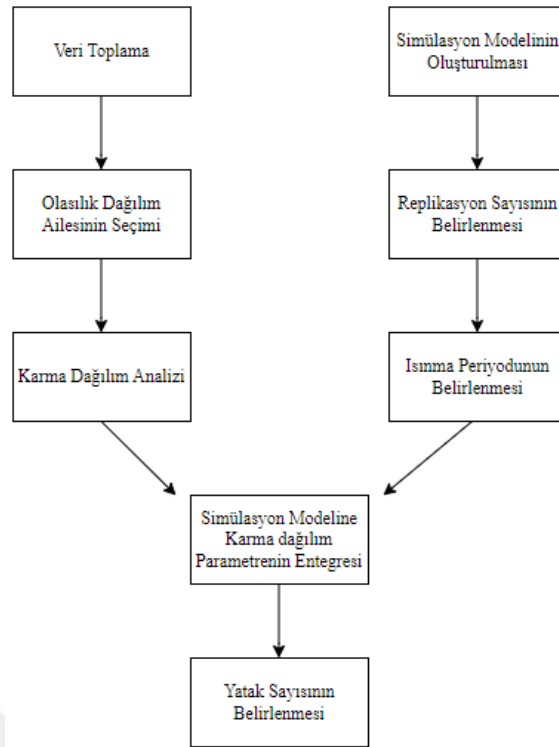
Şekil 3.1. Çukurova Üniversitesi Balcalı EAH akademik personelin istidam bilgileri

Çukurova Üniversitesi Balcalı EAH'nin imkanlarının geniş olması nedeniyle, çevre illerden önemli ölçüde hasta kabulü yapılmaktadır. En kritik durumdaki hastalara hizmet veren YBÜ'lerdeki kaynak sıkıntısı, hastaların ciddi bekleme sürelerine mazur kalmasına ve sistemde hasta taleplerine karşılık verilmesinde problemler yaşanmasına neden olabilmektedir.

Bu çalışmada, 7/24 çalışan Çukurova Üniversitesi Balcalı EAH Pediatri YBÜ ve Reanimasyon YBÜ ele alınmıştır. Veri hastanenin SQL sistemine kaydedilmektedir ve bu çalışmada, Ocak 2018-Aralık 2018 tarihlerini içeren bir yıllık veri kullanılmıştır. Veri seti tarih ve saat cinsinden hasta geliş ve ayrılış zamanlarını içermektedir. Her iki YBÜ için veri, geliş zamanı dikkate alınarak kronolojik olarak sıralanmıştır. Her bir hasta için ayrılış zamanından geliş zamanı çıkartılmış ve saat cinsine çevrilip Servis Süresi (SS), ardışık hastaların geliş zamanı birbirinden çıkartılarak ise Gelişler Arası Süre (GAS) hesaplanmış ve kaydedilmiştir. Hastane yönetiminden elde edilen bilgiye göre, Pediatri YBÜ'de 13, Reanimasyon YBÜ'de ise 8 yatak bulunmaktadır.

Bu bölümün devamında Materyal ve Metot başlığının Metot kısmı ele alınmış ve çalışmanın adımları ve bu adımlarda yararlanılacak yöntemler ayrıntılı bir şekilde açıklanmıştır.

Bu tezde çalışma, üç aşamada yürütülmüştür ve Şekil 3.2.'de akış olarak gösterilmiştir. İlk aşama, verinin elde edilmesi ve bu verinin girdi parametresi olarak tanımlanması için gereken istatistiksel prosedürleri içermektedir. Bu aşamanın paralelinde YBÜ'leri temsil edecek simülasyon modelinin geliştirilmesi ve replikasyon sayısı, ısınma periyodu gibi değerlerin belirlenmesi gerçekleştirilmiştir. Nihayetinde karma dağılım yardımıyla tahmin edilen girdi parametreleri simülasyon modeline entegre edilerek iki aşama birleştirilmiş ve hastaların bekleme süresine maruz kalmaması için gereken yatak sayısı belirlenmeye çalışılmıştır.



Şekil 3.2. Çalışmanın Aşamaları

Bu çalışmanın amacı, Çukurova Üniversitesi Balcalı EAH'nin Pediatri YBÜ ve Reanimasyon YBÜ yatak kapasitesinin yeterli olup olmadığını ve bekleme süresi olmadan hasta kabulü yapabilmek için gereken yatak kapasitesinin belirlenmesidir. Bu hedefle oluşturulan simülasyon modeli için gereken girdi parametreleri olan GAS ve SS verisinin istatistiksel dağılıma uydurma problemi ise karma dağılım yaklaşımıyla çözülmektedir.

3.1. Temel İstatistik Dağılımlar

Bu bölümde özellikle kuyruk modellemede en sık kullanılan dağılımlar ayrıntılı bir şekilde ele alınmıştır.

3.1.1. Normal Dağılım

Normal dağılım, pratik uygulamaları olan bir sürekli olasılık dağılım ailesidir ve Gauss dağılımı olarak da bilinmektedir. Carl Friedrich Gauss, astronomik verisi analiz ettiği sırada normal dağılım grubu ile ilgilenmiş ve bu dağılım için olasılık yoğunluk fonksiyonunu keşfetmiştir. Bu fonksiyonun grafiği, bir çanın şekline benzediği için "çan eğrisi" olarak da bilinmektedir. Bu, dağılımın uçlarının düşük, zirvenin yüksek ve orta noktanın birleştiği bir şekle sahip olduğu anlamına gelmektedir. Normal dağılım, bir simetriye sahiptir. Bu, dağılımın ortasında bir eksen olduğu ve dağılımın her iki tarafının da aynı şekilde düzgün şekilde dağıldığı anlamına gelmektedir.

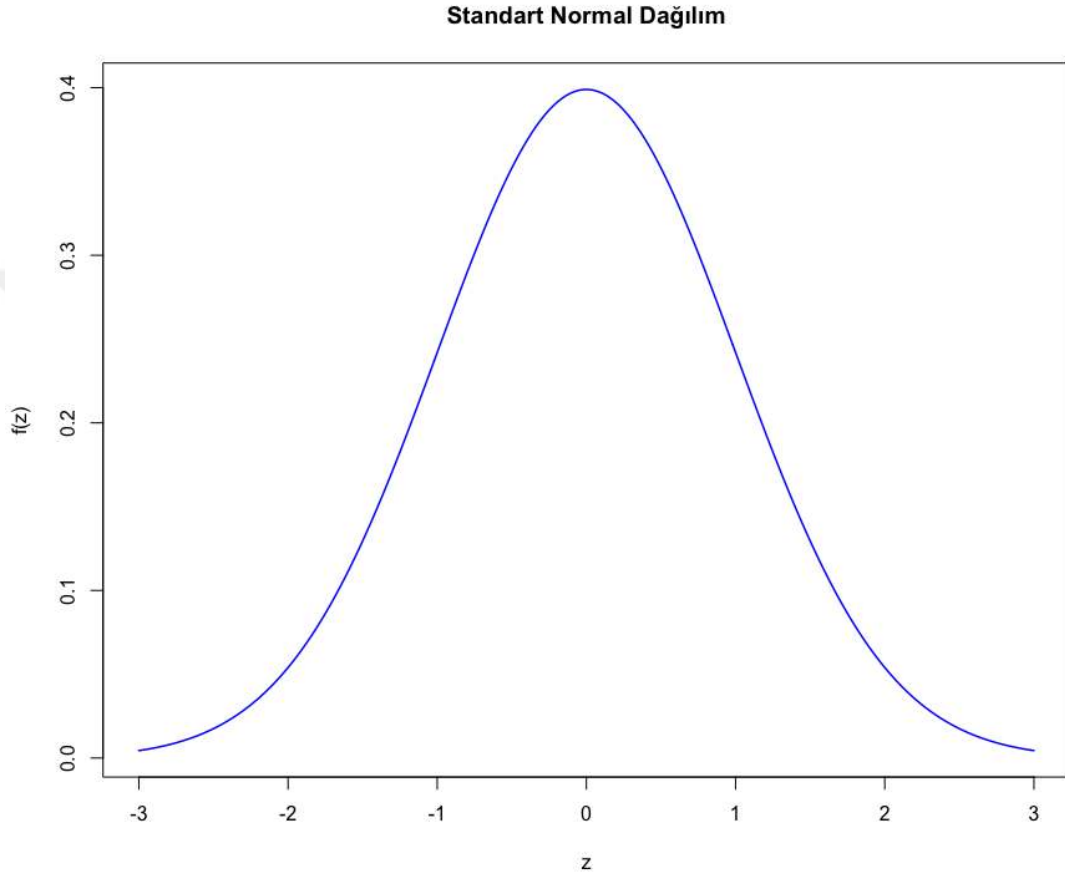
Normal dağılım, istatistik biliminde birçok alanda kullanılmaktadır. İstatistikte kullanılan dağılımlar arasında en yaygın olanı normal dağılımdır ve birçok istatistiksel test normal dağılımın varlığına dayanmaktadır. Olasılık kuramı içinde kullanılan dağılımlar arasında normal dağılım, sürekli olasılık dağılımları ve ayrık olasılık dağılımlarının limitleri de yer almaktadır.

Bu dağılım ailesi, ortalama (μ) ve varyans (σ^2) parametreleri ile tanımlanır. Ortalama, dağılımın merkezindeki nokta ile ifade edilirken, standart sapma, dağılımın ne kadar değişken olduğunu göstermektedir. Normal dağılımın olasılık yoğunluk fonksiyonu (OYF), $f(x) = (1/\sigma\sqrt{2\pi}) \exp(-(x - \mu)^2/2\sigma^2)$ şeklinde gösterilir. Normal dağılımın kümülatif dağılım fonksiyonu (KDF), bir değişkenin belirli bir değere kadar olan olasılığını ifade etmektedir. KDF, normal dağılım eğrisi altında kalan alanın toplamını temsil etmektedir. Genellikle Φ sembolü ile gösterilmektedir. KDF, bir normal dağılım için integral hesaplaması kullanılarak bulunmaktadır. Ancak pratikte, tablolar veya hesaplayıcılar gibi kaynaklar kullanılarak KDF değerleri hesaplanır. KDF, $F(x) = P(X \leq x) = \Phi((x - \mu) / \sigma)$ şeklinde ifade edilmektedir ve burada, X normal dağılımı; Φ sembolü standart normal dağılımın kümülatif dağılım fonksiyonunu temsil etmektedir. KDF'nin kullanılmasıyla, değişkenin belirli bir değer altında, üstünde veya bir aralıkta olma olasılığı hesaplanabilir.

Normal dağılım ile dağılmış X rassal değişkeni için $X \sim N(\mu, \sigma^2)$ şeklinde yazılarak gösterilmektedir. Normal dağılımın diğer bir parametreleme şekli, kesinlik parametresi olarak adlandırılan ve $1/\sigma^2$ değerine eşit olan τ kullanılarak gerçekleştirilmektedir. Bu parametreleme şekli, özellikle çok küçük varyans değerlerinde oluşan sifıra yakın bölmelerden kaynaklanan limit problemlerinden kaçınmayı sağlamaktadır. Ayrıca, normal dağılımın bir üstel ailesi bireyi olarak kullanılabilmesini mümkün kılmaktadır. Bu yöntem, normal dağılımın daha spesifik bir şekilde tanımlanması ve bazı uygulamalarda kullanılması için bir alternatif sunmaktadır.

Normal dağılımın en önemli karakteristik özelliklerinden birisi Merkezi Limit Teoremidir. Merkezi Limit Teoremi (Central Limit Theorem), istatistikte önemli bir teoremdir ve birçok istatistiksel analizin temelini oluşturmaktadır. Merkezi Limit Teoremi, bağımsız ve özdeş dağılımlı rassal değişkenlerin toplamının, toplamda yer alan değişken sayısı n arttıkça normal dağılım gösterdiğini ifade etmektedir. Bu teoreme göre, rassal değişkenlerin dağılımının şekli ne olursa olsun, eğer bu değişkenler bağımsız ve özdeş dağılıma sahipse, toplamlarının dağılımı normal bir dağılım şeklini almaktadır. Böylece, büyük sayıda rassal değişkenin toplamı için normal dağılım yaklaşımı kullanımı mümkün kılınır. Merkezi Limit Teoremi sayesinde, normal dağılım, doğa ve davranış bilimlerinde birçok fenomenin niceliksel modeli yapılmasında kullanılır. Normal dağılım modeli, psikolojik ölçümler ve fiziksel fenomenlerin açıklanmasında iyi bir yaklaşım sağlamaktadır. Merkezi Limit Teoreminde, rasgele değişkenlerin birbirinden bağımsız olması gerekmektedir. Değişkenlerin birbirinden bağımsız olması, her bir gözlem veya denemeni için sonuçların birbirinden etkilenmediği anlamına gelmektedir.

$X_1, X_2, \dots, X_n$ bağımsız ve aynı dağılıma sahip rasgele değişkenler olsun. Bu değişkenler için ortalama μ ve standart sapma σ olduğunu varsayalım. Toplamı $S_n = X_1 + X_2 + \dots + X_n$ olan bu değişkenlerin toplamı için merkezi limit teoremi göz önüne alındığında değişken sayısı n arttıkça dağılım yaklaşık olarak standart normal dağılıma yakınsar. Standart normal dağılım ise ortalama değeri $\mu, 0$ ve varyans değeri $\sigma^2, 1$ olan tek bir normal dağılımdır ve Şekil 3.3.'te OYF'si gösterilmiştir.



Şekil 3.3. Standart Normal Dağılım Fonksiyon Grafiği

3.1.2. Lognormal Dağılım

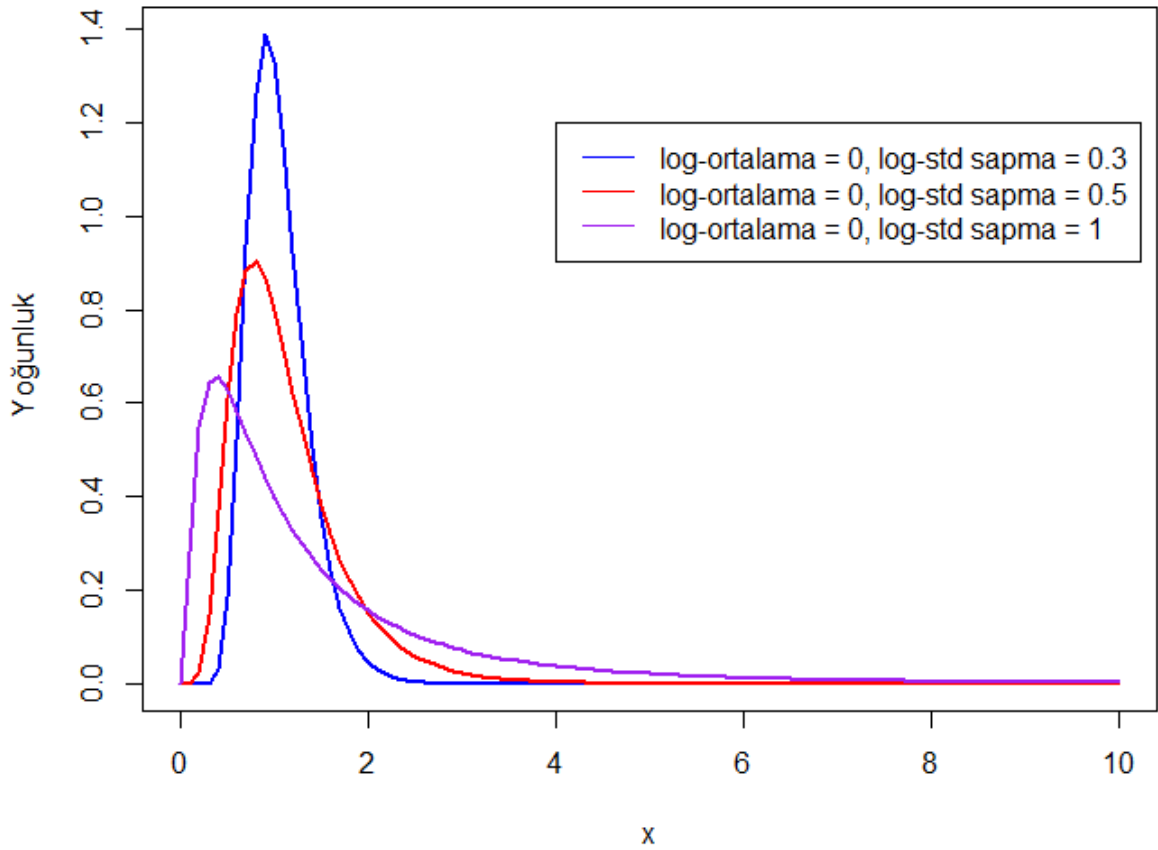
Lognormal dağılımı, bir değişkenin logaritmasının normal, $N(\mu, \sigma^2)$, olarak dağıldığı bir sürekli olasılık dağılımıdır. Lognormal dağılımı, iki parametreye sahiptir. Bu parametreler lognormal dağılımın ortalama (μ) ve standart sapma (σ) değerleri değildir. Bu parametreler dağılımın şeklini belirlemektedir. Lognormal dağılımı Francis Galton'dan dolayı Galton dağılımı olarak kaynaklarda isimlendirilebilir (Johnson ve ark., 1994).

Lognormal dağılımı, normal dağılımın bir dönüşümüdür. Yani, eğer rasgele değişken X lognormal dağılım gösteriyorsa, $Y = \ln(X)$ normal dağılıma sahiptir. Benzer olarak, Y normal dağılım gösteriyorsa, Y 'nin üstel fonksiyonu olan $X = \exp(Y)$, lognormal dağılım gösterir.

Lognormal dağılım, çarpan etkisi olarak da bilinen bir özelliğe sahiptir. Yani, lognormal dağılımın çarpımı, normal dağılımın toplamına yaklaşır.

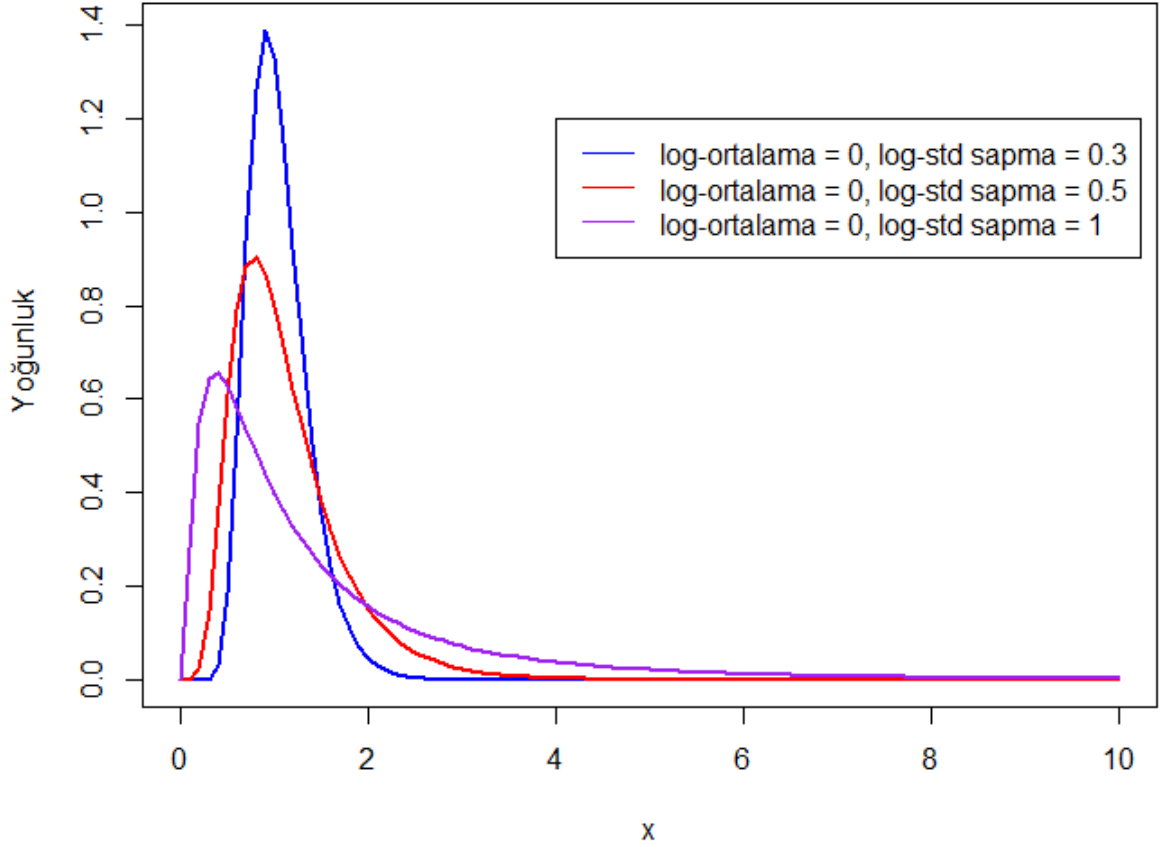
Lognormal dağılımın OYF'si $f(x) = \left(\frac{1}{x\sigma\sqrt{2\pi}}\right) \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)$, KDF'si ise $F(X) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)$ ise şeklinde ifade edilir. Lognormal dağılımın OYF'si, pozitif ve simetrik bir eğriye sahip normal dağılımın logaritmik bir dönüşümüdür. Lognormal dağılımın eğimi, ortalama ve standart sapma değerlerine bağlıdır. Lognormal dağılımın KDF'si, logaritmik bir normal dağılımı temsil eder. Erf terimi, normal dağılımın KDF'sini hesaplamak için kullanılan hata fonksiyonunu ifade eder.

Lognormal Dağılımı



Şekil 3.4.'te görüldüğü gibi, lognormal dağılımı pozitif sonsuz değerlere uzanan uzun kuyruklu ve asimetrik bir dağılımdır ve daha büyük ortalama ve standart sapma değerleri, dağılımın sağa doğru kaymasına neden olur ve daha yüksek değerlerin elde edilme olasılığı artmaktadır. Lognormal dağılımın sağa çarpık olması nedeniyle medyan, ortalama değerinin sağında yer almaktadır.

Lognormal Dağılımı



Şekil 3.4. Farklı Parametre Değerlerinde Lognormal Dağılımları

Lognormal dağılıma sahip bir rasgele değişken sadece pozitif reel değerler almaktadır. Normal dağılıma benzemeyen, pozitif ve sağa çarpık veri setlerini modellemek için sıklıkla kullanılır. Lognormal dağılım, birçok doğal fenomenin, istatistiksel dağılımını iyi bir şekilde açıklamaktadır. Bu nedenle, doğa bilimleri, mühendislik, tıp, ekonomi ve diğer alanlardaki ölçümler özellikle maliyet, gelir, büyüme oranları, stok fiyatları gibi pozitif değerlerin modellenmesi gereken durumlarda sıklıkla kullanılmaktadır. Lognormal dağılımda ortalama değer genellikle medyan değerden çok daha yüksektir ve ortalama değere yapılan katkının büyük kısmı, en yüksek değerli veri noktalarından oluşan küçük bir popülasyondan kaynaklanmaktadır.

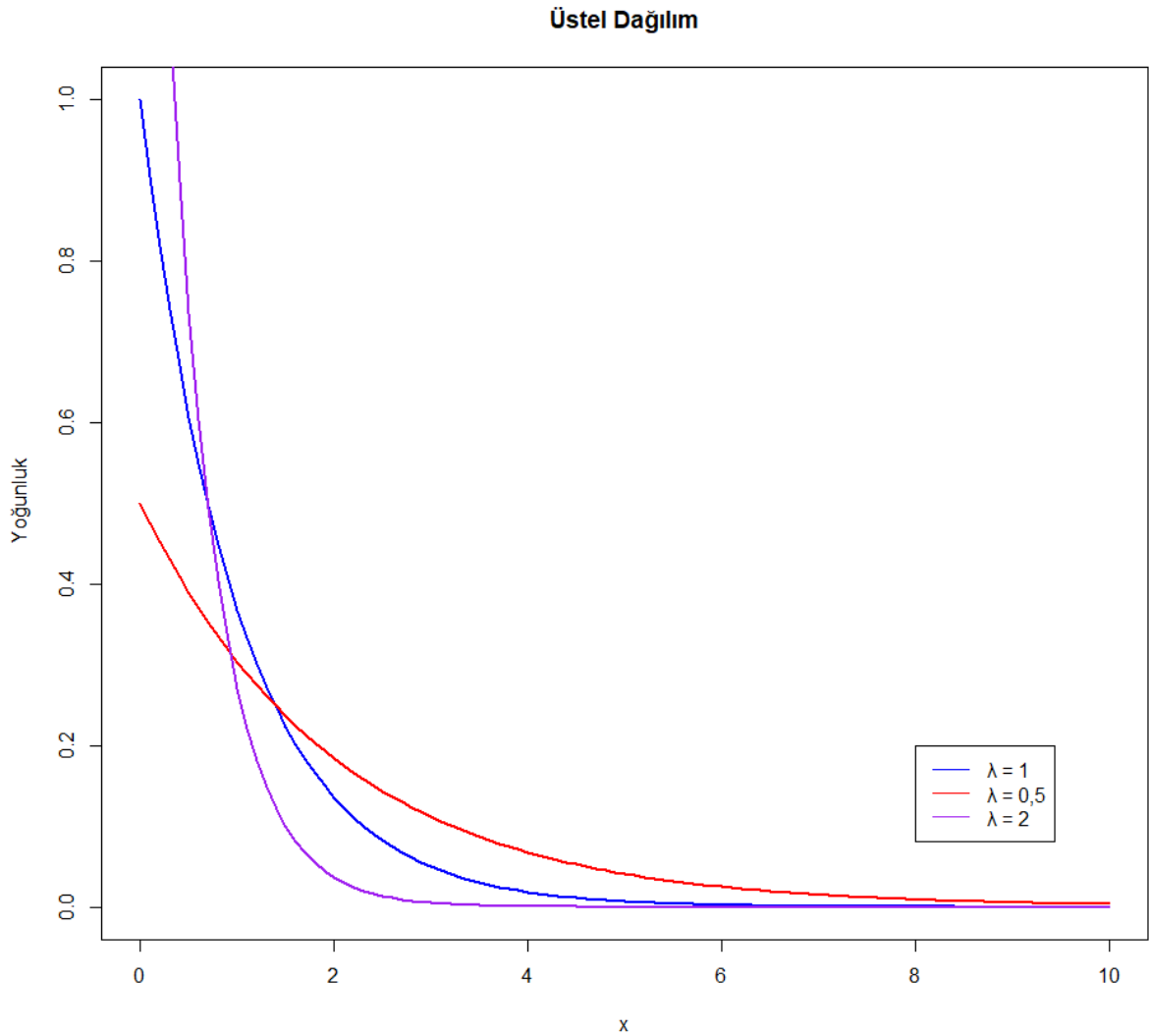
3.1.3. Üstel Dağılım

Üstel dağılım, sürekli olasılık dağılımları arasında önemli bir yere sahip olan bir dağılım türüdür. Bu dağılım, özellikle olayların zaman veya aralık gibi pozitif değerlerle ifade edildiği durumlarda yaygın bir şekilde kullanılmaktadır. Üstel dağılım, rasgele değişkenin pozitif gerçek değerler alır ve bu nedenle olasılık dağılımı sıfırdan büyük değerlerde tanımlıdır.

Üstel dağılımın OYF'si, $f(x; \lambda) = \lambda e^{-\lambda x}$; KDF'si ise $F(x; \lambda) = 1 - e^{-\lambda x}$ şeklinde temsil edilmektedir. Bu formüllerde, x pozitif bir gerçek sayıyı temsil ederken, λ ise pozitif bir

parametredir ve üstel dağılımın oran parametresi olarak adlandırılır. $1/\lambda$ değeriyle, olayların ortalama sayısı ifade edilmektedir. Örneğin, bir süpermarketin saatlik müşteri giriş sayısını modellemek için üstel dağılım kullanıldığında, $1/\lambda$ parametresi ortalama müşteri giriş hızı temsil edilmektedir.

Üstel dağılımın beklenen yani ortalama değeri $1/\lambda$ olarak hesaplanırken, varyansı ise $1/\lambda^2$ şeklinde ifade edilmektedir. Bu dağılımın şekli, oran parametresi λ 'ya bağlı olarak değişmektedir. Daha yüksek λ değerleri, daha sıkı bir dağılımı temsil ederken, daha düşük λ değerleri daha yayılmış bir dağılımı göstermektedir.



Şekil 3.5. Farklı Parametre Değerlerinde Üstel Dağılımlar

Şekil 3.5.'te λ parametresinin 0,5, 1 ve 2 olduğu üstel dağılım grafikleri çizilmiştir. Grafikten görüldüğü üzere, tüm değerler pozitif ve gerçek değerlerdir. Daha büyük bir λ değeri, olasılık yoğunluk fonksiyonunun daha hızlı bir şekilde düşer. λ parametresi arttıkça grafiğin sıkılaştığı ve değerlerin düşüşünün dramatikleştiği görülmektedir. Üstel dağılım oran parametresi λ 'nın değişiminin etkisini daha iyi görebilmek için bu parametre değeri 2'ye yükseltilmiştir. Açıkça

görüldüğü üzere, değerler başlangıç bölgesine doğru yüksek orana yığılmış ve değerlerin yayılımı düşmüştür.

Üstel dağılımın en önemli özelliği belleksizlik özelliğidir. Bu, bir olayın geçmişiyle bağlantısız olduğu anlamına gelmektedir. Yani, gelecekteki olaylar geçmiş olaylardan bağımsızdır. Örneğin, bir rasgele değişkenin beklenen değeri, tüm geçmiş değerlerin bilinmesine rağmen gelecekteki değerlere etki etmemektedir.

Üstel dağılımın birçok pratik uygulaması bulunmaktadır. Örneğin, rasgele olayların zaman aralıklarını modellemek, etkinlik sürelerini tahmin etmek veya düşük olasılıklı olayları modellemek için kullanılabilir. Ayrıca, kuyruk teorisi ve güvenilirlik analizi gibi alanlarda da yaygın olarak kullanılmaktadır. Üstel dağılım, pozitif sürelerin modellenmesi için önemli bir araç olarak tanımlanabilir ve gerçek hayatta birçok olayın doğasını etkili bir şekilde yansıtabilir. Bununla birlikte, dağılımın sadece pozitif değerleri kapsaması ve belirli bir olayın bağımsız olmasını varsayması, bazı durumlarda sınırlayıcı olmaktadır. Dolayısıyla, uygulama alanına bağlı olarak diğer dağılım modelleriyle karşılaştırılmalı ve en uygun olanı seçilmelidir.

Üstel dağılımın istatistiksel özelliklerini anlamak ve doğru bir şekilde kullanmak, gözlemlerin analizinde ve tahminlerin yapılmasında büyük önem taşımaktadır. Bu nedenle, istatistik ve olasılık teorisi alanında çalışan araştırmacılar ve analistler için üstel dağılımın temel prensiplerini ve uygulamalarını anlamak oldukça önemlidir.

3.1.4. Gamma Dağılımı

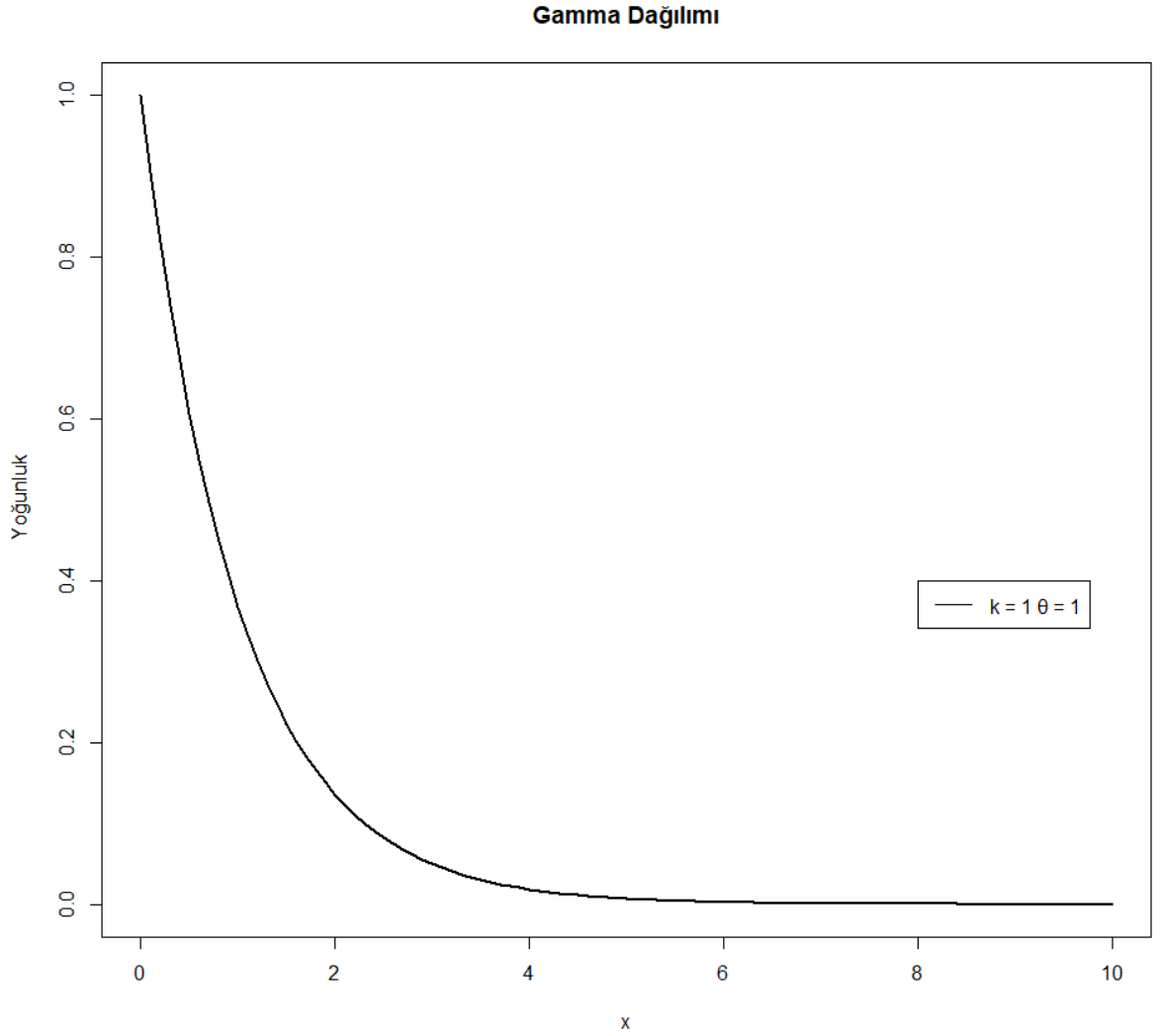
Gamma dağılımı, sürekli bir olasılık dağılımıdır ve genellikle sürelerin veya beklenen değerlerin modellenmesinde kullanılır. Gamma dağılımı iki parametre ile tanımlanır. İlki ölçek (θ), ikincisi ise şekil (k) parametresidir. Bu parametrelere alternatif olarak bazı kaynaklarda şekil parametresi α ile gösterilir. Ölçek parametresi ise tersi alınarak oran ($\beta = 1/\theta$) parametresine dönüşür. Şekil parametresi, dağılımın şeklini belirlerken, ölçek parametresi ise dağılımın genişliğini kontrol eder. Bu parametrelere bağlı olarak, gamma dağılımının ortalama değeri k/θ ya da $\alpha\beta$, varyans değeri ise, k/θ^2 veya $\alpha\beta^2$ şeklinde hesaplanır.

Gamma dağılımı, pozitif kesirli değerleri ve sağa çarpık veri setlerini modellemek için kullanılan sürekli bir dağılımdır. Bu, dağılımın sıfırdan küçük olamayacağı anlamına gelmektedir. Gamma dağılımının bu özelliği, özellikle servis süreleri gibi gözlemlerin tanımlanması için büyük bir avantaj sağlamaktadır ve pozitif sürelerin modellenmesi ve analizi için ideal bir seçenek olarak ortaya çıkmaktadır.

Gamma dağılımı, YBÜ'lerdeki hastaların servis sürelerinin tahmin edilmesi ve kapasite planlaması gibi alanlarda oldukça faydalı bir dağılımdır. Gerçek veri dağılımını daha iyi yansıtmaya ve tahmin etme yeteneği sayesinde, gamma dağılımının esnek yapısı önemli bir avantaj sağlamaktadır ve sürekli olasılık dağılımları arasında geniş bir uygulama alanına sahiptir. Özellikle, pozitif sürelerin analizi ve tahmini için ideal bir seçenek olarak kabul edilmektedir. Bu nedenle,

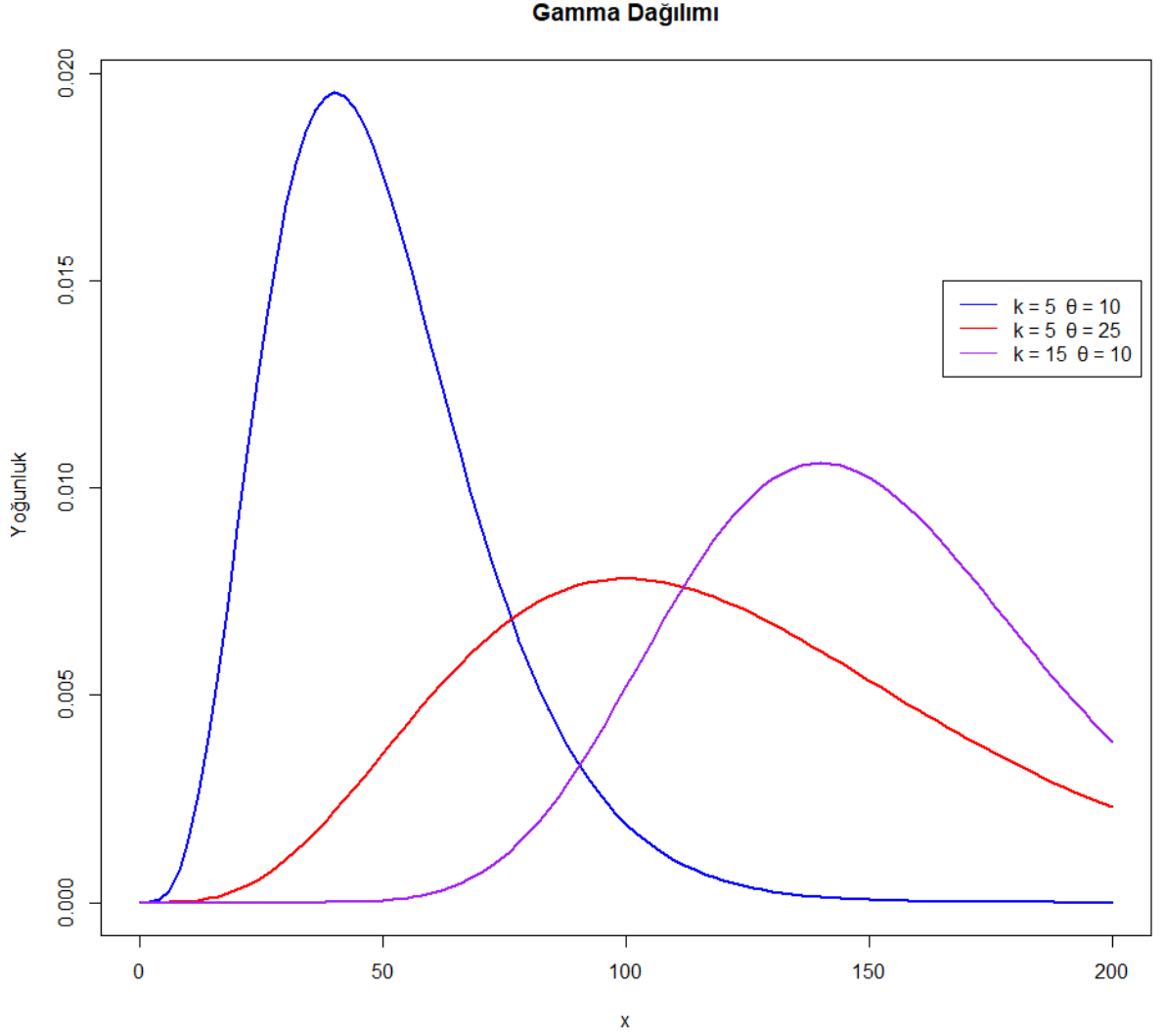
YBÜ gibi kritik alanlarda gamma dağılımının doğru şekilde kullanılması, hasta bakımının kalitesini artırmaya ve kaynakların verimli bir şekilde kullanılmasına yardımcı olabilir.

Gamma dağılımının yoğunluk fonksiyonu $f(x) = (\frac{1}{\Gamma(k)} \theta^k) x^{k-1} e^{-x/\theta}$ şeklinde ifade edilir. Burada k , şekil parametresidir ve θ ölçek parametresidir. Eğer $k = 1$ ise, gamma dağılımı üstel dağılıma indirgenir ve OYF'si Şekil 3.6.'daki gibi olur. Üstel dağılım elde edilmekte ve bu dağılım aynı zamanda; şekil ve ölçek parametreleri ile kontrol edilebilir, türevlenebilir bir fonksiyon olarak ifade edilebilir ve üstel dağılımdan hareketle Poisson dağılımı ile ilişkilendirilebilir.



Şekil 3.6. $k = 1, \theta = 1$ olan Gamma Dağılımı

Gamma dağılımının farklı parametre değerlerinde nasıl bir yoğunluk grafiği görünümünde olduğu incelenmek istenmiştir. Bu kapsamda, şekil parametresi k , ve ölçek parametresi θ sistematik olarak farklı değerler aldığı OYF'ler çizilmiş ve Şekil 3.7.'de paylaşılmıştır.



Şekil 3.7. Farklı Parametre Değerlerinde Gamma Dağılımları

Şekil parametresi k 'nın 5, ölçek parametresi θ 'nın ise 10 olduğu bir Gamma dağılımının OYF'si incelendiğinde, sağa çarpık olduğu görülmektedir.

Şekil parametresi k bir önceki durum ile aynı değer olan 5, ölçek parametresi θ ise 10 değerinden 25 değerine yükseltildiğinde, şekil parametresi k sabit kalarak, ölçek parametresi θ 'nın artmasıyla oluşan OYF'nin tepesi daha geniş hale geldiği ve zirve noktasının sağa kaydığı görülmüştür.

Şekil parametresi k 'nın değişmesiyle gamma dağılımı ile üretilen sayıların yoğunluk grafiği de incelenmiştir. Bu hedefle, şekil parametresi k 15, ve ölçek parametresi θ ise 10 olarak belirlenmiş OYF'si çizilmiştir. Şekil parametresi k arttıkça gamma dağılımın tepe noktası sağ tarafa doğru kaydığı ve sola çarpık bir dağılım oluştuğu görülmektedir.

Gamma dağılımı, özellikle istatistikte ve mühendislikte yaygın olarak pozitif sayılarla ilgili birçok problemin çözümünde kullanılan kullanılan bir dağılımdır.

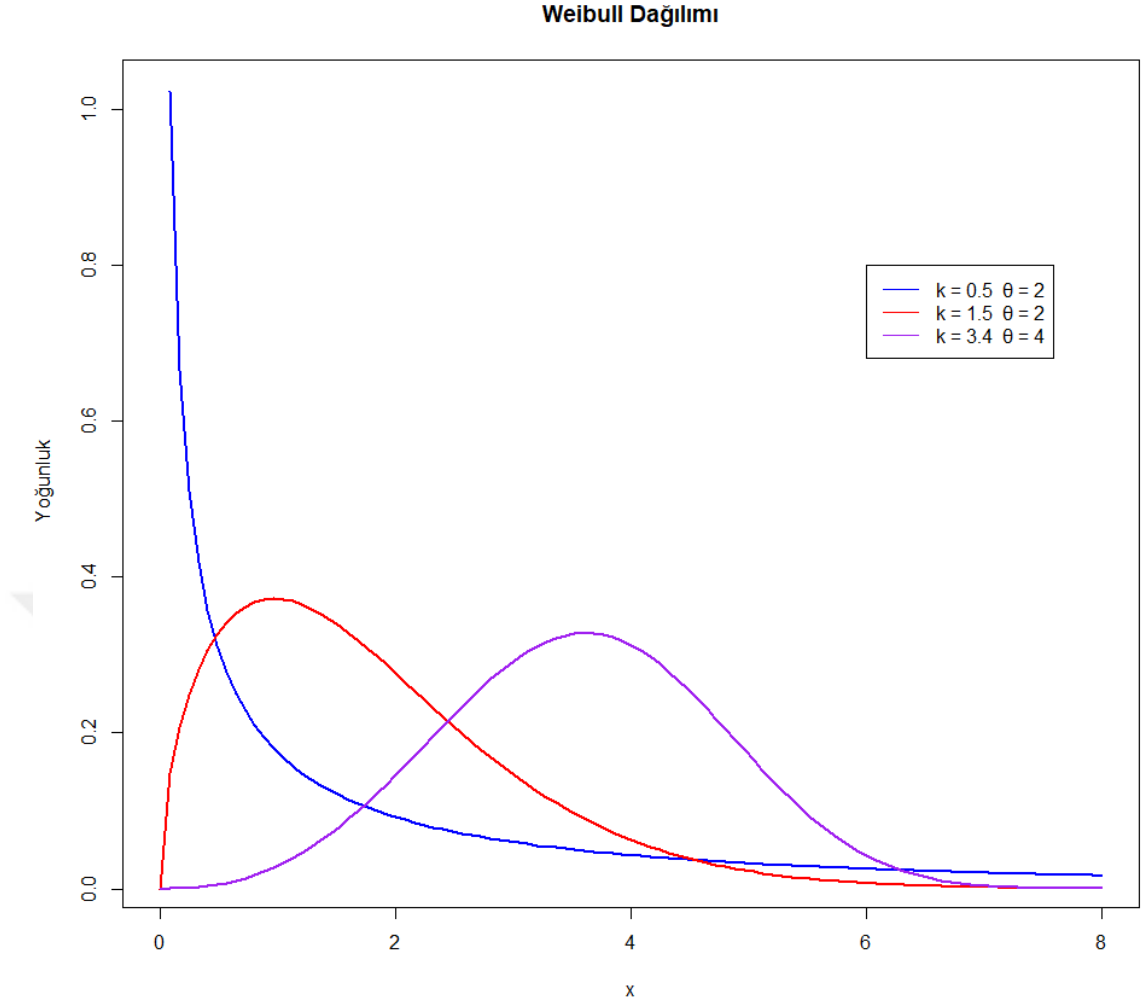
3.1.5. Weibull Dağılımı

Weibull dağılımı, olasılık kuramı ve istatistik bilim dallarında sıklıkla kullanılan bir sürekli olasılık dağılımıdır. Bu dağılım, Waloddi Weibull adlı bir bilim insanının anısına isimlendirilmiştir (Weibull, 1951).

Ölçek ve şekil olarak iki parametreyle tanımlanır. Ölçek parametresi (λ) pozitif bir sabittir ve verinin yayılımını kontrol eder. Şekil parametresi (k) ise dağılımın formunu belirler. Pozitif eksende tanımlı Weibull dağılımının OYF'si $f(x) = \left(\frac{k}{\lambda}\right)\left(\frac{x}{\lambda}\right)^{(k-1)} e^{-(x/\lambda)^k}$ ve KDF'si ise $F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$ şeklinde ifade edilmektedir.

Genel olarak, hastanede görülen servis süresi dağılımları asimetriktir ve genellikle gamma, lognormal ve Weibull dağılımlarıyla uyumludur (Atienza, ve ark., 2008; Griffiths, ve ark. 2006).

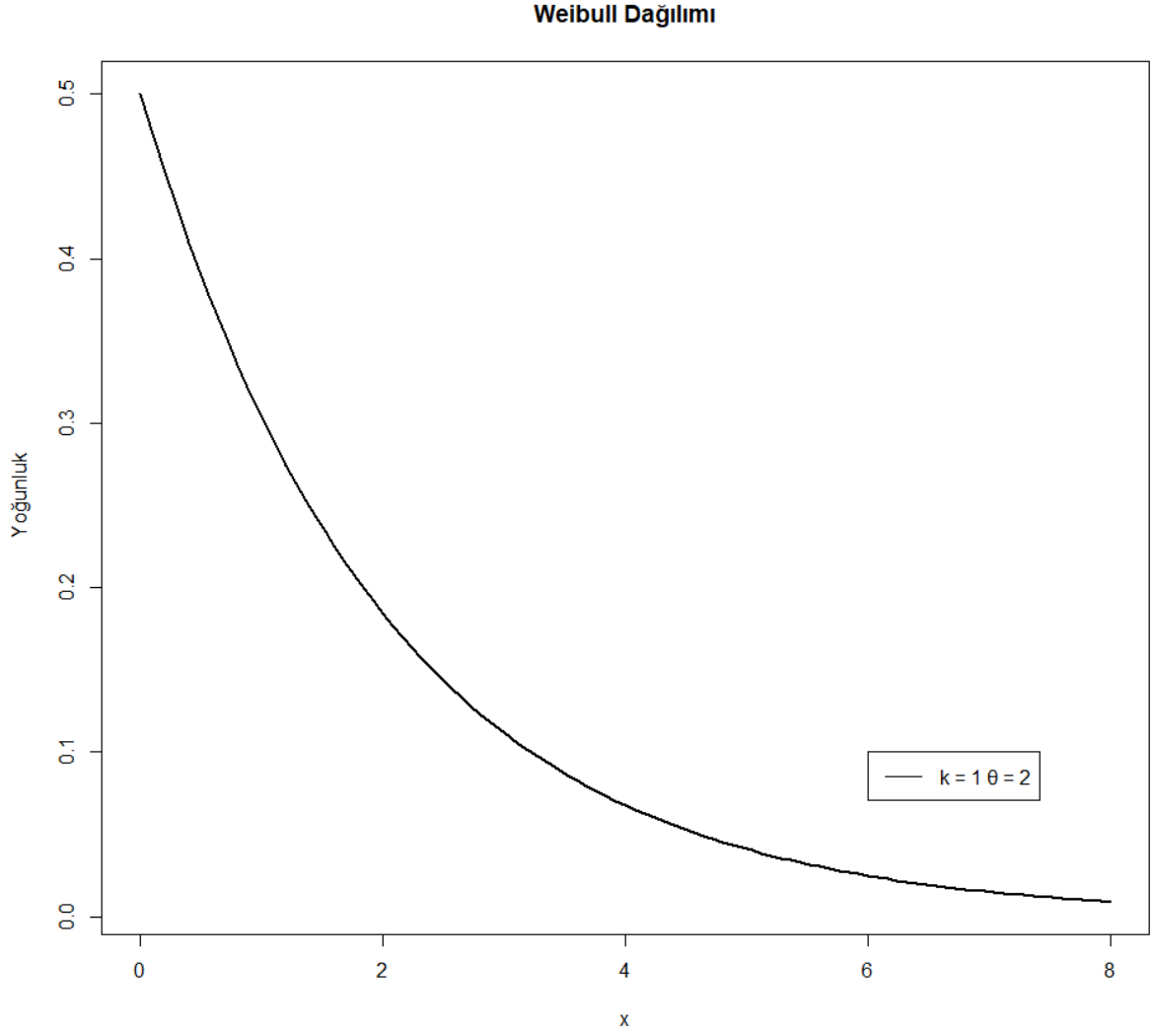
Weibull dağılımı, farklı şekil parametre değerleri için farklı şekiller almaktadır. Şekil parametresi, veri setinin dağılım şeklini belirler. $k > 1$ olduğunda, sağa çarpık bir dağılım elde edilmektedir. $k = 1$ olduğunda, Weibull dağılımı, üstel dağılıma dönüşmektedir. $k < 1$ olduğunda, sola çarpık bir dağılım elde edilmektedir. Ölçek parametresi, dağılımın genişliğini veya ölçeğini belirler. λ değeri büyüdükçe, dağılım genişler. λ değeri küçüldükçe, dağılım daralır. Bu durumları gözlemek için farklı parametre değerlerine sahip Weibull dağılımı grafikleri çizilmiş ve Şekil 3.8.'de paylaşılmıştır.



Şekil 3.8. Farklı Parametre Değerlerinde Weibull Dağılımları

Şekil parametresi, $k < 1$ parametresine sahip Weibull dağılımı azalan bir hızda artan bir yoğunluğa sahip olacak şekilde sağa kaymaktadır. $k > 1$ olan Weibull dağılımında grafik, artan bir hızda azalan bir yoğunluğa sahip olacak şekilde sola kaymıştır. Weibull dağılımında şekil parametresinin özel değerlerinden biri 3,4 değeridir. $k = 3,4$ değerini aldığı anda dağılım normal dağılıma oldukça benzerlik göstermektedir.

$k, 1$ olduğunda ise; Weibull dağılımı üstel dağılım gibi davranır. Bu parametreye sahip Weibull dağılımı ise Şekil 3.9’da gösterilmektedir.



Şekil 3.9. $k = 1$ Olduğu Durumda Weibull Dağılımı

Weibull dağılımının kullanım alanları oldukça geniştir ve birçok gerçek hayat veri seti Weibull dağılımı yardımıyla başarıyla modellenenebilir. Örneğin, ömrü, arızalanma süresi, kullanım süresi gibi süreçlerin modellenmesinde bu dağılım kullanılır. Weibull dağılımı, özellikle yüksek doğruluk oranları gerektiren dayanıklılık analizleri gibi alanlarda sıklıkla tercih edilmektedir. Bu dağılımın özellikleri arasında, iki parametrelili olması, çeşitli şekil parametreleri ile özelleştirilebilir olması ve güvenilir bir şekilde veri uyumunu test etme yeteneği yer almaktadır. Weibull dağılımı, üstel dağılım gibi belleksizlik özelliğine sahiptir. Yani, gelecekteki olaylar geçmiş olaylardan bağımsızdır. Bu özellikleri sayesinde Weibull dağılımı, geniş bir uygulama alanına sahiptir ve birçok istatistiksel analizde kullanılmaktadır.

3.1.6. Ampirik Dağılım

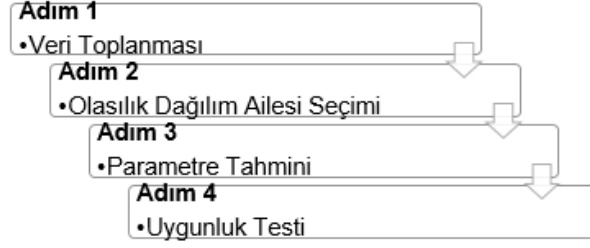
Çoğu gerçek yaşam durumunda, varış süreci ve hizmet süreci rasgele bir şekilde gerçekleşmektedir. Bu rasgeleliği modellemenin yollarından biri, olayın belirli bir olasılık dağılımı tarafından yönetilen rasgele bir değişken olarak tanımlamaktır. Örneğin, bankaya varışlar bir Poisson sürecine göre gerçekleşiyorsa, o zaman olasılık teorisinden varışlar arası süre dağılımının üstel bir dağılım olduğu bilinmektedir. Genel olarak, müşterilerin nasıl geldiğine ilişkin bilgiler, sistemin doğrudan gözlemlenmesi yoluyla veya geçmiş veri kullanılarak güvence altına alınmalıdır. Hiçbir bilgi kaynağı mevcut değilse, rasgele süreci bir olasılık modeliyle açıklamak için bazı varsayımlar yapılmalıdır.

Ampirik dağılım, teorik bir dağılım gibi bir fonksiyon olmasa da, veri kümesinin dağılımını tanımlamak için sıklıkla kullanılmaktadır. Bu sayede, veri kümesinin özellikleri hakkında daha fazla bilgi elde edilebilir ve gelecekteki analizler için bir referans noktası sağlanabilir. Ancak, ampirik dağılımın kullanımı teorik bir dağılıma göre daha kısıtlı, az güvenilir ve veriyi temsil etmede hassasiyeti düşüktür. Ampirik dağılımlarda herhangi bir parametre tahmini yapılmaması nedeniyle parametrik olmayan bir yöntemdir. Ayrıca ampirik dağılımlar veri setindeki her bir gözlem için bir çubuk veya nokta kullanması sebebiyle çok büyük veri kümeleri için uygun olmayabilir. Ampirik dağılım, veri setine dayanır ve veri setinin dağılımını temsil etmektedir. Ancak, bu dağılım, teorik dağılımlardan elde edilen parametrelerle tam olarak eşleşmeyebilir. Ampirik dağılım, veri setinin özelliklerini yakalamaya yönelik bir yaklaşımdır ve gerçek dağılımın bir tahmini olarak kullanılmaktadır.

Karma dağılım modelleri, ampirik dağılımların aksine, çeşitli teorik dağılımların birleşiminden oluşmaktadır. Bu sayede, veri kümesinin farklı bölümlerinin farklı dağılımlarla modellenebileceği durumlarda bile uygun bir modelleme sağlanmaktadır. Ayrıca, karma dağılım modelleri genellikle daha fazla parametreye sahip oldukları için daha esnek bir yapıya sahiptir. Bu da, daha fazla veriye uyum sağlamalarını ve daha doğru sonuçlar üretmelerini sağlamaktadır. Ampirik dağılımların bahsi geçen dezavantajlarından dolayı, bu çalışmada karma dağılım yönteminin kullanılması fikri ortaya atılmaktadır.

3.2. Veri Toplama

Simülasyon modeli oluşturulurken girdi parametrelerinin belirlenmesi en önemli aşamadır ve Şekil 3.10.'da gösterildiği dört adımda ileriye alınmıştır. Yapılan çalışmanın ilk adımı veri toplama olarak planlanmıştır.

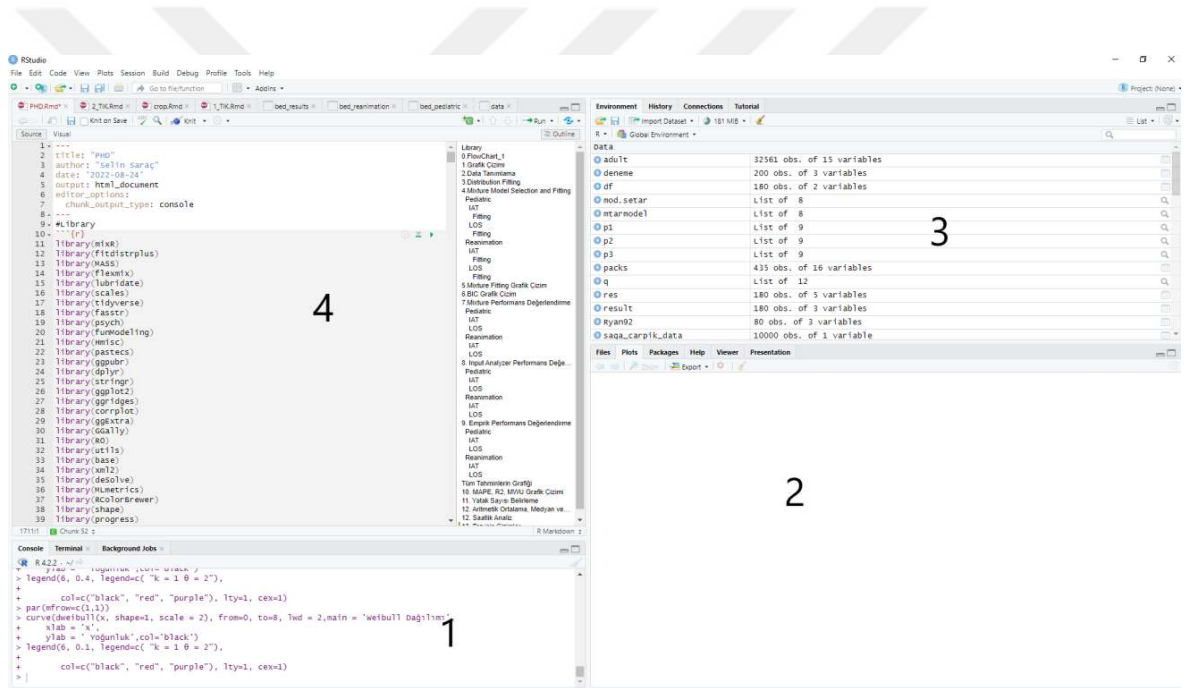


Şekil 3.10. Olasılık Dağılımı Seçilmesi

Bir yıl boyunca Çukurova Üniversitesi Tıp Fakültesi Balcalı EAH Pediatri ve Reanimasyon YBÜ'lerinde kaydedilen hasta kayıt sisteminden temin edilmiştir. Microsoft Excel tipindeki veri seti, hastanın yatış yaptığı YBÜ'yu, tarih ve saat cinsinden geliş ve ayrılış zamanını içermektedir. Tüm veri gerçek hastalardan oluşmasına rağmen, hiçbir hastanın özel ve kişisel bilgisini içermemektedir. Bu veri seti, geliş zamanları kronolojik olarak tasnif edilmiştir. Ardışık hastaların geliş zamanı birbirinden çıkartılıp saat cinsine çevrilerek gelişler arası süre (GAS) ve her bir hastanın ayrılış zamanından geliş zamanı çıkartılarak servis süresi (SS) hesaplanmıştır. Nihayetinde, Çukurova Üniversitesi Tıp Fakültesi Balcalı Eğitim ve Araştırma Hastanesi Pediatri YBÜ'den 759, Reanimasyon YBÜ'den ise 255 hasta verisi olmak üzere toplamda 1014 hastanın verisi elde edilmiştir.

Yapılan tüm istatistiksel analizlerde RStudio programı kullanılmıştır. RStudio programı, günümüzde veri bilimcileri, istatistikçiler ve diğer veri odaklı profesyoneller tarafından sıklıkla kullanılan bir araçtır. Bu program başlangıçta Auckland Üniversitesi İstatistik bölümünde geliştirilmiştir. Açık kaynak kodlu bir entegre geliştirme ortamı olan RStudio, R programlama dilinin kullanımını kolaylaştırarak, veri analizi, veri görselleştirme, makine öğrenmesi, istatistik ve diğer birçok alanda kullanılabilecek araçlar sunar. RStudio, veri bilimcilerinin ve istatistikçilerin işlerini kolaylaştırmayı amaçlamaktadır. Bunun için R programlama dilinde kullanılacak tüm araçları bir araya getirerek, veri analizi sürecini daha verimli hale getirir. RStudio, kullanıcılara ihtiyaçlarına uygun olarak farklı sürümler sunar. RStudio Desktop, bir masaüstü uygulaması olarak kullanılabilecek tam özellikli bir sürümdür. Bu sürüm, kullanıcıların R programlama dili ile çalışmalarını yapmalarına ve veri analizlerini gerçekleştirmelerine olanak tanır. RStudio Server, bir sunucu uygulaması olarak kullanılabilir ve uzaktan erişim için idealdir. Bu sürüm, birden fazla kullanıcının aynı sunucuya bağlanarak, aynı kaynakları paylaşarak çalışmalarını yapmalarına olanak sağlar. RStudio programı, veri analizi sürecinde kullanılabilecek birçok araç sunar. R kodu düzenleme, hata ayıklama, ve otomatik tamamlama özellikleri, kullanıcıların R programlama dili ile çalışırken hızlı ve doğru bir şekilde işlemlerini gerçekleştirmelerine olanak tanır. Ayrıca R Markdown ile dokümantasyon ve raporlama yapmak mümkündür. Grafik ve veri görselleştirme araçları da RStudio'nun sunduğu özellikler arasındadır. Bu özellikler sayesinde, kullanıcılar veri analizlerini görselleştirebilir ve daha anlaşılır hale getirebilirler.

SPSS, Excel, MATLAB gibi birçok veri analizi yazılımı bulunmasına rağmen, bu çalışmada RStudio tercih edilmiştir. Bunun ilk sebebi, RStudio'nun açık kaynak kodlu ve ücretsiz bir yazılım olmasıdır. Dünya genelinde 2 milyondan fazla kullanıcısı bulunması bu programın sürekli gelişmesini sağlamaktadır. Kullanım kolaylığı ve uygulama esnekliği sayesinde de tercih edilen bir yazılım haline gelmiştir. Grafik ve görsel üretimdeki başarısı, paket kullanım imkanı, kolay raporlama ve RMarkdown sayesinde sunum yapma kolaylığı da bu programın artıları arasında yer almaktadır. Bunun yanında birkaç dezavantajları da bulunmaktadır. Örneğin, sık paketler ve programlar güncellenme gereksinimi duymaktadır. Ayrıca uzmanlaşmanın diğer programlara göre biraz daha zor olduğu söylenebilir. Bununla birlikte, RStudio yine de veri analizi, veri görselleştirme, makine öğrenmesi, istatistik ve diğer birçok alanda kullanımı kolay ve popüler bir yazılımdır. Bu sebeple bu tezde yapılan tüm istatistiksel analizler ve grafikler RStudio kullanılarak elde edilmiştir. RStudio'nun arayüzü Şekil 3.11'de gösterildiği gibidir.



Şekil 3.11. RStudio Ara yüzü

Birinci kısım, konsol bölümüdür. Tüm kodlar bu bölüme yazılır ve sonuçlar görünür. İkinci kısımda, Dosyalar, Grafikler, Paketler ve Yardım sekmeleri bulunur. Dosyalar sekmesinde içinde çalışılan klasörün içindeki dosyalar listelenir. Grafikler sekmesinde, çizilen tüm grafikler görüntülenir. Bu grafikler hem Plot sekmesi hem de kod yazılarak pdf, PNG gibi formatlarda kaydedilebilir. Paketler sekmesinde, programda yüklü olan kütüphaneler takip edilebilir. Kullanılacak kodlara göre bulunduğu kütüphaneler aktif edilmeli ve güncellikleri sıklıkla kontrol edilmelidir. Mevcut paketler kullanılabileceği gibi, kullanıcılar kendi kodlarını yaratarak yeni paketler oluşturabilir. Yardım sekmesi bu programın en güçlü yönlerinden biridir. Kullanılacak kod Yardım sekmesine yazılarak tüm komponentleri ve çeşitli örneklerle kodun kullanımı ayrıntılı bir

şekilde açıklanır ve ilgili kodla ilgili daha fazla bilgiye sahip olunur. Üçüncü kısım ise Ortam ve Geçmiş sekmelerini içerir. Ortam sekmesinde oluşturulan değişkenlerin listesi ve bilgileri listelenir. RStudio'da kodlar yukarıdan aşağıya doğru okunur. Bu, bir değişkene en son hangi değer verilmişse o değere atama yapılacağı anlamına gelmektedir. Geçmiş sekmesinde önceden çalıştırılan kodlar takip edilebilir ve bu bölümün oluşturduğu .RHistory dosyasıyla ortamlar arası geçmiş kod bilgisi aktarılabilir. Dördüncü ve son kısım ise Editör ekranı olarak adlandırılır. Bu ekranda kodları yazıp çalıştırabilir, kullanıcı yorumları ve açıklamaları eklenebilir. Mevcut dört panelin, yeri, sırası ve/veya mevcudiyeti kullanıcının istediğine göre değiştirilebilir. Aynı zamanda, ara yüzün arka plan rengi, yazı tipi gibi genel özellikleri yine kullanıcı seçimine göre kişiselleştirilebilir.

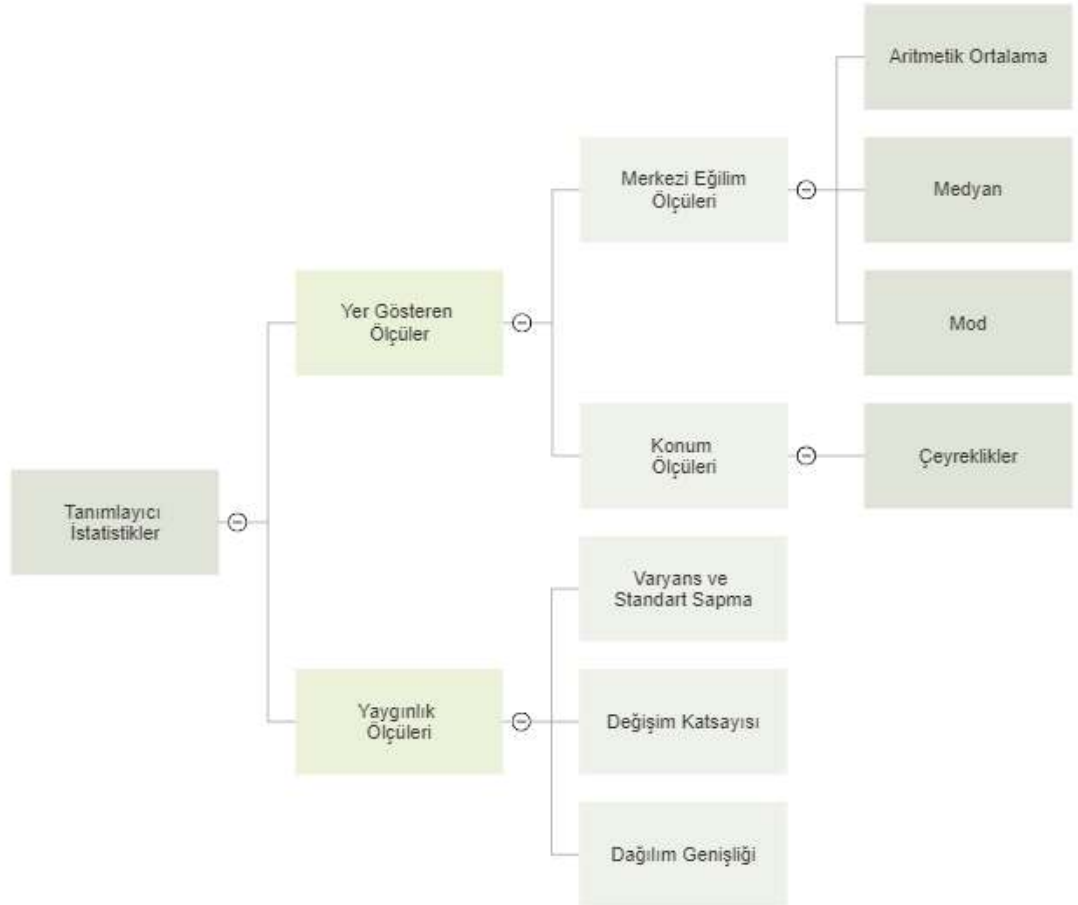
3.3. Olasılık Dağılım Ailesinin Seçimi

Rasgele faktörler hayatımızın tüm alanlarını etkiler. Günümüzün son derece rekabetçi ortamında başarılı olmaya çalışan sistemleri, risk ve belirsizlikle başa çıkmak için bir araca ihtiyaç duymaktadır. Olasılık dağılımlarını kullanmak, belirsizlikle baş etmenin ve bilinçli iş kararları vermenin bilimsel bir yolu olarak tarif edilebilir. Uygulamada olasılık dağılımları, aktüerya bilimi ve sigortacılık, risk analizi, yatırım, pazar araştırması, iş ve ekonomik araştırma, madencilik, kimya mühendisliği, hidroloji, görüntü işleme, fizik, tıp gibi çok çeşitli alanlarda uygulanmaktadır.

Simülasyon modelleri sistemi olasılıksal olarak tanımlayan sistemlerdir (Yavuz, 2008). İşlem zamanı, iş tipleri, işlemler arası süreler, tamir süreleri gibi rassallık kaynakları sistemin deterministik olarak tanımlanmasını engellemektedir. Gerçek hayatta hemen hemen tüm sistemler bir ya da birden fazla rassallık içeren kaynağa sahiptir. Bu nedenle GAS ve SS gibi stokastik girdileri simülasyonda modellemek için öncelikle bu girdilerin en iyi uyan olasılık dağılım ailesinin seçilmesi gerekmektedir. Simülasyon modeline bilinen bir olasılık dağılım ailesi olarak seçildiğinde, çalışma zamanı boyunca, tahmin edilen girdi parametrelerinden üretilen rassal sayılar kullanılmaktadır. Olasılık dağılım ailesini gerçek sistemi ne kadar iyi temsil ederse, bu girdiler yardımıyla oluşturulacak parametreler ve hipotez testleri de o kadar iyi uyumda olur. Böylece simülasyon modeli, gerçek sistemi en iyi şekilde temsil eder (Gürsakal, 2018). Olasılık dağılımları, belirsizlikle başa çıkmak için bir yöntem olarak görülebilir. Ancak veriye tam olarak uymayan bir dağılım seçip uygulanırsa, sonraki hesaplamalarınız yanlış olacak ve bu kesinlikle yanlış kararlara yol açacaktır. Yanlış modellerin kullanılması, görevlerin veya projelerin zamanında tamamlanamaması; gereksiz ya da fazla ekipman planlaması gibi ciddi zaman ve para kaybına, yanlış mühendislik tasarımının pahalı ekipmanların hasar görmesine neden olması gibi ciddi sonuçlara yol açabilir. Doğru olasılık dağılım ailesinin seçimi, rassal süreçlerin geçerli modellerini geliştirmeye olanak tanımakta, geçersiz model seçimi nedeniyle doğabilecek potansiyel zaman ve para kaybından korumakta ve daha iyi iş kararları verilmesini sağlamaktadır.

İlk adımda veri sistemi en iyi temsil edebilecek süre boyunca toplanmalıdır. Bu nedenle veri toplanmadan önce sistem ayrıntılı şekilde incelenmeli ve gerekli formlar hazırlanmalıdır. Mevsimsellik olduğu düşünülen durumlarda en az 12 aylık veri toplamak gerekir. Analiz edilmek istenen çıktılara göre toplanacak gözlemlerin ne olduğunun tespiti de önemli bir faktördür. Örneğin, cinsiyet, gelir durumu, yaş gibi değişkenlerin herhangi bir yönelime olan etkisi incelenmek isteniyorsa, bu bilgiler de ayrıca kaydedilmelidir. Yeterli sayıda örneklemin olması girdi olasılık dağılım ailesinin doğru olarak belirlenebilmesi için önemli unsurlardan biridir. Bu nedenle yeteri kadar örnek toplandığından emin olunmalıdır. Verilerin homojen olarak toplanması göz önünde bulundurulması gereken bir başka unsurdur. Yani, veri ardışık zaman periyotlarında ve aynı ortamda kaydedilmelidir. Veri toplama aşamasında uç veri üreten durumlar söz konusu olduğunda not edilmeli ve gerekirse veri temizleme aşamasında bu veri göz ardı edilmelidir.

İkinci adım olasılık dağılım ailesinin seçimidir. Çoğu durumda, iki veya daha fazla dağılımı seçip, sonuçları karşılaştırmak ve en geçerli modeli seçmek gerekir. Seçilen "aday" dağılımlar, gözlemlerin doğasına bağlı olarak seçilmelidir. Örneğin, teknik cihazların arızaları arasındaki süreyi analiz ediliyorsa, arıza süresi negatif olamayacağından, üstel veya Weibull gibi negatif değer almayan dağılımlar seçilmelidir.



Şekil 3.12. Tanımlayıcı İstatistik Sınıflandırılması

Veriyi oluşturan olasılık dağılım ailesini doğru modelleyebilmenin birçok avantajı vardır. Bu dağılım ailesini seçebilmenin en kolay yollarından biri histogramları incelemektir. Bununla birlikte bir verinin hangi dağılıma uyacağına karar verirken tanımlayıcı istatistikler oldukça faydalıdır.

Histogram, veri kümesinin dağılımının grafiksel bir temsili olan veri analizinde iyi bilinen bir çizimdir. Histogram, kutulara düşen gözlem sayısını özetler. Kutu sayısı aralığını değiştirmek, yoğunluğun histogramın şekline göre ortak bir olasılık dağılımı gibi görünüp görünmediğini belirlemeye yardımcı olmaktadır. Ayrıca gözlemlerin simetrik, çarpık olması, tepe noktasına veya aykırı değerlere sahip olup olmadığı konusunda da ipuçları vermektedir.

Dağılım ailesi seçiminde yer gösteren ölçüler, yaygınlık ölçüleri ve görsel teknikler kullanılabilir. Verilerin çarpıklık, basıklık, aykırı değerler, dağılım şekli, tek değişkenli, çift modlu vb. gibi temel özelliklerini keşfetmek dağılım ailesi seçiminde oldukça önemlidir. Bu özelliklere dayanarak hangi modellerin en iyi olduğuna karar vermek daha kolaydır. Tanımlayıcı istatistikler, verisi özelliklerine göre karakterize etmeye olanak tanımaktadır. Tanımlayıcı istatistikler Yer Gösteren Ölçüler ve Yaygınlık Ölçüleri olmak üzere ikiye ayrılır. İlk olarak Yer Gösteren Ölçüler ele alınmıştır.

3.3.1. Yer Gösteren Ölçüler

Yer gösteren ölçüler, merkezi eğilim ölçüleri ve konum ölçüleri olmak üzere iki ana başlığa ayrılabilir. Merkezi Eğilim Ölçüleri, kitlenin tüm farklı değerlerinin merkezini gösteren ölçülerdir. Aritmetik ortalama, medyan (ortanca), mod (tepe değeri) ve çeyreklikler bu grubun altında yer almaktadır. Matematikte, özellikle istatistikte birkaç tür ortalama vardır. Her ortalama, genellikle belirli bir veri kümesinin genel değerini daha iyi anlamak için belirli bir veri grubunu özetler.

Bir veri kümesi için, aritmetik ortalama, sonlu bir sayı kümesinin merkezi eğiliminin bir ölçüsüdür, veri kümesindeki tüm sayıların toplamının veri sayısına bölünmesiyle $\bar{x} = \frac{1}{n}(\sum_{i=1}^n x_i)$ hesaplanır. Aritmetik ortalama hem tüm kitle hem de seçilen bir örneklem için ayrıca hesaplanabilir. Kitlenin aritmetik ortalaması μ , örneklemin aritmetik ortalaması ise $\bar{x}$ sembolüyle gösterilir.

Medyan, anlaşılması basit, hesaplanması kolay ve aynı zamanda ortalamaya sağlam bir yaklaşım olduğu için, tanımlayıcı istatistiklerde popüler bir özet istatistiktir. Medyan değeri, bir ana kütle ya da örneklem veri serisinin küçükten büyüğe sıralandığında tam ortada yer alan değeri olarak tanımlanır, bir başka deyişle popülasyonun veya olasılık dağılımının üst yarısını alt yarısından ayıran değerdir. Medyan, tek boyutlu veri için iyi tanımlanmıştır ve mesafe ölçüsünden bağımsızdır. Bu nedenle sıralı ve sayısal olmayan veri için uygulanamaz. Aşırı uç değerlerden etkilenmez (Yüzer, 2006). Bu sebeple yanlış bir göstergedir. Buna rağmen veri setinin “tipik” değerinin ne olduğunu anlamak için iyi bir temsil sağlamaktadır. Simetrik olmayan dağılımlar için

medyan, %50'lik kırılma noktasına sahip ve dirençli bir istatistik değeri olması nedeniyle, ortalama değerden daha uygun bir merkezi eğilim ölçüsüdür. Medyan, tipik olarak bir dağılımın çarpık olduğu, uç değerlerin bilinmediği veya aykırı değerlerin güvenilir olmaması, yani ölçüm/transkripsiyon hataları olduğu durumlarda bir merkezi eğilim ölçütü olarak kullanılabilir. Aynı zamanda veri sayısının çok fazla olduğu durumlarda verisi iyi özetleyen bir merkezi eğilim ölçüsüdür. Tam ortada olan değer olan medyan bulunurken, veri sayısının tek ya da çift olmasına bağlı olarak farklı şekilde hesaplanır.

$$Medyan = \begin{cases} x_j, j = \frac{n+1}{2} & n \text{ tek,} \\ \frac{x_j + x_{j+1}}{2}, j = \frac{n}{2} & n \text{ çift} \end{cases}$$

$$x_1 \leq x_2 \leq \dots \leq x_n$$

Mod, bir anakütlede en sık görülen ya da tekrarlanan değer olarak tanımlanır, tepe değer olarak da adlandırılır. Bir veri dizisi içerisinde tekrar eden değer bulunmuyorsa, mod değeri olmayacağı gibi; eşit sayıda tekrar eden değerler çoklu mod durumunu ortaya çıkartabilir ve bu durumda dağılımlar unimodal, bimodal veya multimodal olarak sınıflandırılır. Medyan değerine benzer şekilde, mod değeri aykırı veriden etkilenmez. Ancak veri sayısının az olduğu durumlarda güvenilir bir merkezi eğilim ölçüsü değildir, cebirsel işlemlere uygun değildir ve tanımlanması zor olabilir.

Tanımlayıcı istatistiklerde, ortalama, medyan, mod veya orta aralık ile karıştırılabilir, çünkü bunlardan herhangi biri "ortalama" olarak adlandırılabilir. Bir dizi gözlemin ortalaması, değerlerin aritmetik ortalamasıdır; ancak çarpık dağılımlar için ortalama, ortadaki değer (medyan) veya en olası değer (mod) ile aynı olmak zorunda değildir. Örneğin, ortalama gelir, çok yüksek gelire sahip az sayıda insan tarafından tipik olarak yukarı doğru çarpılır, böylece çoğunluğun geliri ortalamanın altında olur. Buna karşılık, medyan gelir, nüfusun yarısının altında, yarısının üstünde olduğu seviyedir. Mod geliri, en olası gelirdir ve daha düşük gelirli daha fazla sayıda kişiyi temsil eder. Medyan ve mod genellikle bu tür çarpık veri için daha sezgisel ölçümler olsa da, üstel ve Poisson dağılımları da dahil olmak üzere birçok çarpık dağılım en iyi ortalama değer ile tanımlanır. Simetrik bir dağılımda mod, medyan ve aritmetik ortalama birbirine eşittir (NIST/SEMATECH, 2022).

Olasılık teorisi ve istatistikte çarpıklık, bir rasgele değişkenin ortalamasına göre olasılık dağılımının asimetrisinin bir ölçüsüdür. Çarpıklık değeri pozitif, sıfır, negatif veya tanımsız olabilir. Tek modlu bir dağılım için, negatif çarpıklık genellikle kuyruğun dağılımın sol tarafında olduğunu, pozitif çarpıklık ise dağılımın kuyruğunun sağda daha uzun olduğunu gösterir. Bir kuyruğun uzun, diğerinin kalın olduğu durumlarda çarpıklık basit bir kurala uymaz. Örneğin, sıfır değeri, ortalamanın her iki tarafındaki kuyrukların genel olarak dengelendiği anlamına gelir; bu

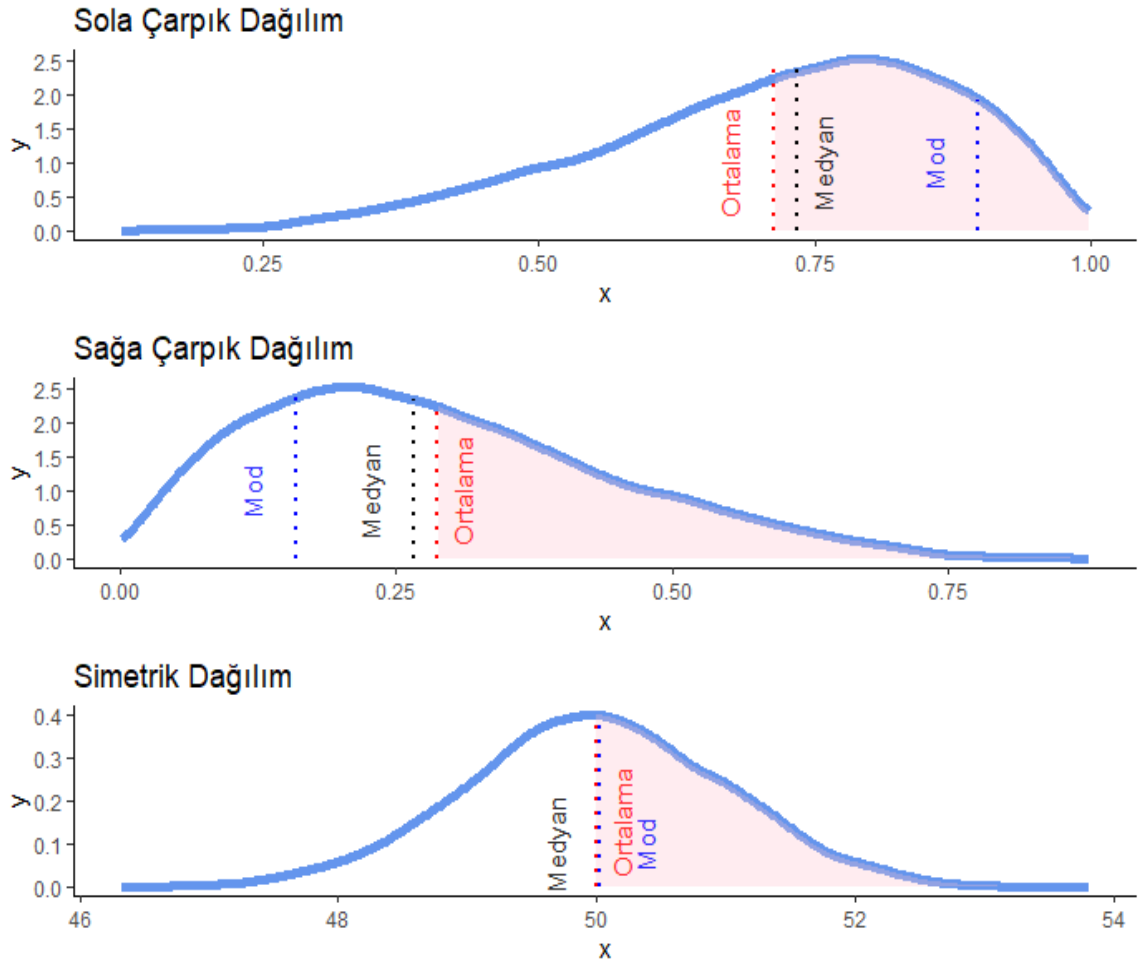
simetrik bir dağılım için geçerlidir, ancak bir kuyruğun uzun ve ince, diğerinin kısa ama kalın olduğu asimetrik bir dağılım için de geçerli olabilir. Simetrik olmayan dağılımlarda veriler sola (negatif) ya da sağa (pozitif) çarpıktır.

Bir dağılımın çarpıklığı, çarpıklık katsayısı ile ölçülür. Bu katsayı dağılımın çarpıklık yönünü göstermektedir. Çarpıklık katsayısı değeri 0'a yakınsa, dağılım simetrik kabul edilir. Negatif bir değer ise sola çarpıklığı (sol kuyruğun daha ağırlıklı olduğunu) ifade ederken, pozitif bir değer sağa çarpıklığı (sağ kuyruğun daha ağırlıklı olduğunu) gösterir.

Çizelge 3.1. Çarpıklık Katsayısı Yorumu

Çarpıklık Katsayısı	Çarpıklık Yorumu
-0,5 ile +0,5 arası	Simetrik
-1 ile -0,5 veya +0,5 ile +1 arası	Orta
1'den küçük veya +1'den büyük	Yüksek

Çizelge 3.1.'de çarpıklık katsayısının nasıl yorumlanacağı verilmiştir. Çarpıklık katsayısı -0,5 ile +0,5 arasındaysa, dağılım simetriktir ve çarpıklığı yoktur. Bu durumda, veri merkezi bir eğilim etrafında dengeli bir şekilde dağılmıştır. Çarpıklık katsayısı -1 ile -0,5 veya +0,5 ile +1 arasındaysa, dağılım ortalama bir çarpıklığa sahiptir ve orta düzeyde çarpıktır. Bu durumda, dağılım hafif bir şekilde sola veya sağa çarpık olabilir. Ancak, çarpıklık katsayısı 1'den küçük veya +1'den büyükse, dağılım yüksek bir çarpıklığa sahiptir. Bu durumda, dağılım belirgin bir şekilde sola veya sağa çarpıktır ve kuyruklarından biri diğerine göre daha uzun veya ağır basmaktadır.



Şekil 3.13. Sola, Sağa Çarpık ve Simetrik Dağılım Grafikleri

Negatif çarpık dağılımlarda, sol kuyruk daha uzundur ve dağılım grafiğın sağ tarafında yoğunlaşmıştır. Pozitif çarpık dağılımlarda ise sağ kuyruk daha uzundur, ortalama değeri grafiğın sağ tarafındadır. Çarpıklık değeri sıfır olması, dağılımın mutlaka simetrik bir dağılım olduğunun ispatı değildir ancak simetrik dağılımların çarpıklığı her zaman sıfıra eşittir. Sola, sağa çarpık ve simetrik dağılım grafikleri Şekil 3.13.'te ayrıntılı şekilde çizilmiştir.

Bir dağılımın sağa çarpık olduğu durumlarda, dağılımın aritmetik ortalaması, ortanca değerinden daha büyüktür. Ortanca değeri ise tepe değerinden daha büyük olur. Yani, bu dağılımda gözlemlerin çoğu sağ tarafta toplanmış ve bu nedenle aritmetik ortalama sağa doğru kayar. Ancak, dağılımın tamamına bakıldığında, ortanca değeri genellikle ortada ve tepe değeri ise solda yer alır. Eğer dağılım sola çarpıksa, tepe değeri ortalama ve ortanca değerden daha büyüktür. Bu durumda, verinin çoğu soldadır ve bu nedenle aritmetik ortalama da sol tarafa doğru kaymıştır. Yine de, dağılımın tamamına bakıldığında, ortanca değeri genellikle ortada, tepe değeri ise sağ tarafta yer alır.

Bir dağılımın simetrik ise, dağılımın ortasında bulunan üç merkezi eğilim ölçüsü olan aritmetik ortalama, ortanca değeri ve tepe değeri birbirine eşittir. Bu dağılımın sol tarafı ile sağ tarafı birbirine aynı şekilde yansıtılır ve böylece dağılımın tam orta noktası bulunur. Bu dağılımda

bulunan verinin yarısı orta noktanın solunda, yarısı da sağında bulunur. Aritmetik ortalama, ortanca değeri ve tepe değeri bu durumda aynı noktada bulunduğu için, bu üç merkezi eğilim ölçüsü dağılımın merkezinde bulunur.

Çeyreklikler, verinin küçükten büyüğe sıralandıktan sonra dört eşit parçaya bölünmesiyle oluşan değerlerdir. Birinci çeyreklik değeri verinin %25'lik ilk parçasını gruplar ve Q1 olarak temsil edilir. Q2 değeri medyan yani ortanca değere denk gelmektedir. Benzer şekilde Q3 olarak adlandırılan üçüncü çeyreklik değeri verinin %75'ini sağında, kalan %25'ini ise solundadır. Q1 değeri, medyan değerinden küçük, Q3 ise büyük olan verinin ortalaması olarak hesaplanmaktadır.

3.3.2. Yaygınlık Ölçüleri

Merkezi eğilim ölçüleri veri setinde bulunan gözlemlerin ortalama etrafında toplanma özelliklerini yansıtırken, yaygınlık ölçüleri ise gözlemlerin nasıl dağıldığını göstermektedir. Bir anlamda, dağılım ölçüleri veri setinde bulunan gözlemler arasındaki değişimi göstermektedir. Aritmetik ortalama, medyan ve mod değeri genellikle dağılımın şekli hakkında bilgi vermekte yetersiz kalmaktadır. Bu durumlarda değişim genişliği, varyans, standart sapma ve değişim katsayısı gibi dağılım ölçüleri kullanılmaktadır. Değişim genişliği (range), varyans, standart sapma ve değişim katsayısı gibi değerler Yaygınlık Ölçüleri başlığı altında incelenebilir.

Değişim genişliği, bir verinin en kolay anlaşılan yaygınlık ölçüsüdür. Değişim genişliği, verinin en büyük ve en küçük değeri arasında farktır ve değişim aralığını gösterir. Sadece uç iki değer göz önüne alınması nedeniyle diğer verinin hiçbir önemi yoktur. Bu nedenle tek başına bir dağılım ölçüsü olarak kullanılması sağlıklı değildir. Değişim genişliği, farklı veri setlerini karşılaştırmada sıkça kullanılan bir ölçüttür. Ancak bu karşılaştırmalar yapılırken, veri setlerinde bulunan gözlem sayıları birbirine eşit olmalı ve aynı birimi taşımalıdır. Farklı sayıda ve farklı birimlerdeki gözlemlerin değişim genişliği ile karşılaştırılması doğru sonuçlar veremez. Bu nedenle, değişim genişliği ölçüsünü kullanırken dikkatli olmak gerekmektedir ve karşılaştırma yapılacak veri setlerinin özelliklerine uygun bir şekilde kullanılmalıdır.

Olasılık teorisi ve istatistikte, varyans, rasgele bir değişkenin popülasyon ortalamasından veya örneklem ortalamasından sapmasının kareler toplamının birim sayısına bölünmesiyle hesaplanır. Bir veri setinin ortalama değerlerinden ne kadar uzağa yayıldığıнын bir ölçüsüdür. Aritmetik ortalamaya benzer şekilde tüm kitle hem de seçilen bir örneklem için hesaplanabilir. Kitlenin varyansı σ^2 , örneklemin varyansı ise s^2 sembolüyle gösterilir ve $\sum_{j=1}^n (x_j - \bar{x})^2 / (n - 1)$ şeklinde hesaplanır. Varyans değerinin karekökü standart sapmayı vermektedir. Kitlenin standart sapma değeri σ , örneklemin standart sapma değeri ise s sembolüyle gösterilir.

Değişim katsayısı ($c_v = \sigma/\mu$), normalize edilmiş ve birimsiz bir dağılım ve sürekli dağılımlarda kullanılan bir ölçüdür. Standart sapma değerinin aritmetik ortalamaya bölümü sonucunda elde edilir. Böylece farklı ölçeklerden oluşan veri setlerinin değişkenlik ölçüleri

karşılaştırılabilir. Örneğin, üstel dağılım için değişim katsayısı; $c_v = \sigma/\mu = \sqrt{1/\beta^2/1/\beta} = 1$ olarak hesaplanır. Değişim katsayısının 1'e yakın olduğu durumlarda üstel dağılım seçilebilir. Değişim katsayısının hesaplanabilmesi için, formülü tanımı gereği aritmetik ortalama değerinin sıfırdan farklı değerde olması gerekir. Değişim katsayısının değeri arttıkça, gözlemlerin değişkenliğinin yüksek olduğu yorumu yapılır. Değişim katsayısının küçük olduğu durumlarda, veri aritmetik ortalama değeri etrafında homojen bir yayılım gösterir.

Yaygınlık ölçüleri genel olarak şöyle tanımlanabilir. Değişim genişliği, en küçük ve en büyük veri arasındaki farkı gösterirken, varyans ve standart sapma gözlemlerin ortalama etrafında ne kadar dağıldığını göstermektedir. Değişim katsayısı ise standart sapmayı ortalama etrafında normalize ederek, farklı birimlerdeki gözlemlerin karşılaştırılmasını kolaylaştırır.

Yer gösteren ve yaygınlık ölçülerine göre, veriyi temsil edebilecek aday dağılımlar seçilebilir. Ele alınan dört dağılımın tüm özellikleri Çizelge 3.2.'te ayrıntılı olarak gösterilmiştir.

Çizelge 3.2. Dağılım Özellikleri

Dağılım	Normal	Gamma	Lognormal	Weibull	Üstel
Parametre	μ, σ^2	k, θ	μ, σ	λ, k	λ
Ortalama	μ	$k\theta$	$e^{\mu+\sigma^2/2}$	$\lambda\Gamma(1 + 1/k)$	$1 / \lambda$
Medyan	μ	Kapalı form yok	e^μ	$\lambda\ln(2)^{1/k}$	$\frac{\ln 2}{\lambda}$
Mod	μ	$(k - 1)\theta$	$e^{\mu-\sigma^2}$	$\lambda(k - 1/k)^{1/k}$	0
Varyans	σ^2	$k\theta^2$	$(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$	$\lambda^2\Gamma(1 + 2/k) - \mu^2$	λ^{-2}
Çarpıklık	0	$2/\sqrt{k}$	$(e^{\sigma^2} + 2)\sqrt{e^{\sigma^2}} - 1$	$\Gamma(1 + 3/k)\lambda^3 - 3\mu\sigma^2 - \mu^3/\sigma^3$	2

3.3.3. Grafiksel Yöntemler

Verilerin frekans dağılım histogramları, kutu grafikleri, olasılık dağılım grafiği, Quantile-Quantile (Q-Q) ve Percentile- Percentile (P-P) grafikleri seçilecek aday dağılıma dair güvenilir bilgiler verir.

Histogram, gözlemlerin dağılımını görselleştirmek için kullanılan bir grafiksel yöntemdir. Bu grafik, verinin frekans dağılımını çubuklar halinde gösterilir. Yatay ekseninde değer aralıkları ve dikey ekseninde ise bu aralıklara düşen frekanslar yer almaktadır. Histogram çizimi için ilk adım verinin aralıklara ayrılmasıdır. Bu aralıkların büyüklüğü, verinin dağılımına ve veri sayısına bağlı olarak belirlenir. Daha sonra her aralık için verinin frekansı hesaplanarak, bu frekanslar grafik üzerinde çubuklar halinde gösterilir. Literatürde, örneklem sayısının n olduğu bir veri seti $\sqrt{n}$ aralığa bölünmesi gerektiği önerilmiştir (Montgomery, 2009). Başka çalışmaya göre, gözlem sayısı $n > 50$ ise, aralık sayısı 10 - 20, aksi durumda 6 - 10 arasında tayin edilebileceği belirtilmiştir (Blank, 1968). Histogramlar, gözlemlerin aday dağılıma uyumlu olup olmadığını belirlemek için sıklıkla kullanılan ancak tam anlamıyla yeterli olmayan bir yöntemdir. Histogramlar, aykırı

değerlerin belirlenmesi ve verinin dağılımının simetri veya çarpıklığı hakkında bilgi edinmek için kullanılır. Ayrıca, farklı grupların veri karşılaştırılması için de kullanılabilir.

Kutu grafikleri, veri dağılımını görselleştirmek için kullanılan bir yöntemdir. Bu grafik, verinin çeyreklikler, minimum ve maksimum değerler, aykırı değerler ve medyan gibi istatistiksel özelliklerini gösterir (DuToit ve ark., 1986). Bu sayede veri ortalamalarından, dağılımlarına kadar birçok istatistiksel özellikleri hakkında bilgi edinilebilir. Kutu grafikleri, ilk olarak 1970'lerde John Tukey tarafından geliştirilmiştir (Tukey, 1977). Bu grafik, genellikle tek bir değişkenin dağılımını göstermek için kullanılır, ancak birden fazla değişkenin karşılaştırılması için de kullanılabilir. Ayrıca, gruplar arasındaki farkları da göstermek için kullanılabilir. Kutu grafikleri, beş temel bileşen içerir,

- Minimum Değer: En küçük değer kutunun sol ucunda yer alır.
- Alt Çeyrek (Q1): Verilerin en küçük %25'ini temsil eder.
- Medyan (Q2): Verilerin ortanca değerini temsil eder.
- Üst Çeyrek (Q3): Verilerin en büyük %25'ini temsil eder.
- Maksimum Değer: En büyük değer kutunun sağ ucunda yer alır.

Ayrıca, kutu grafikleri aykırı değerleri de gösterir. Aykırı değerler, genellikle diğer değerlerden önemli ölçüde farklı olan değerlerdir ve grafikte daire veya üçgen şeklinde gösterilirler. Verilerin dağılımı hakkında bilgi edinmek, aykırı değerlerin tespiti ve gruplar arasındaki farkların belirlenmesi gibi birçok amaçla çok sık kullanılan bir yöntemdir. Python, RStudio ve Excel gibi birçok istatistiksel yazılım paketi, kutu grafiklerini çizmek için fonksiyonlar içerir.

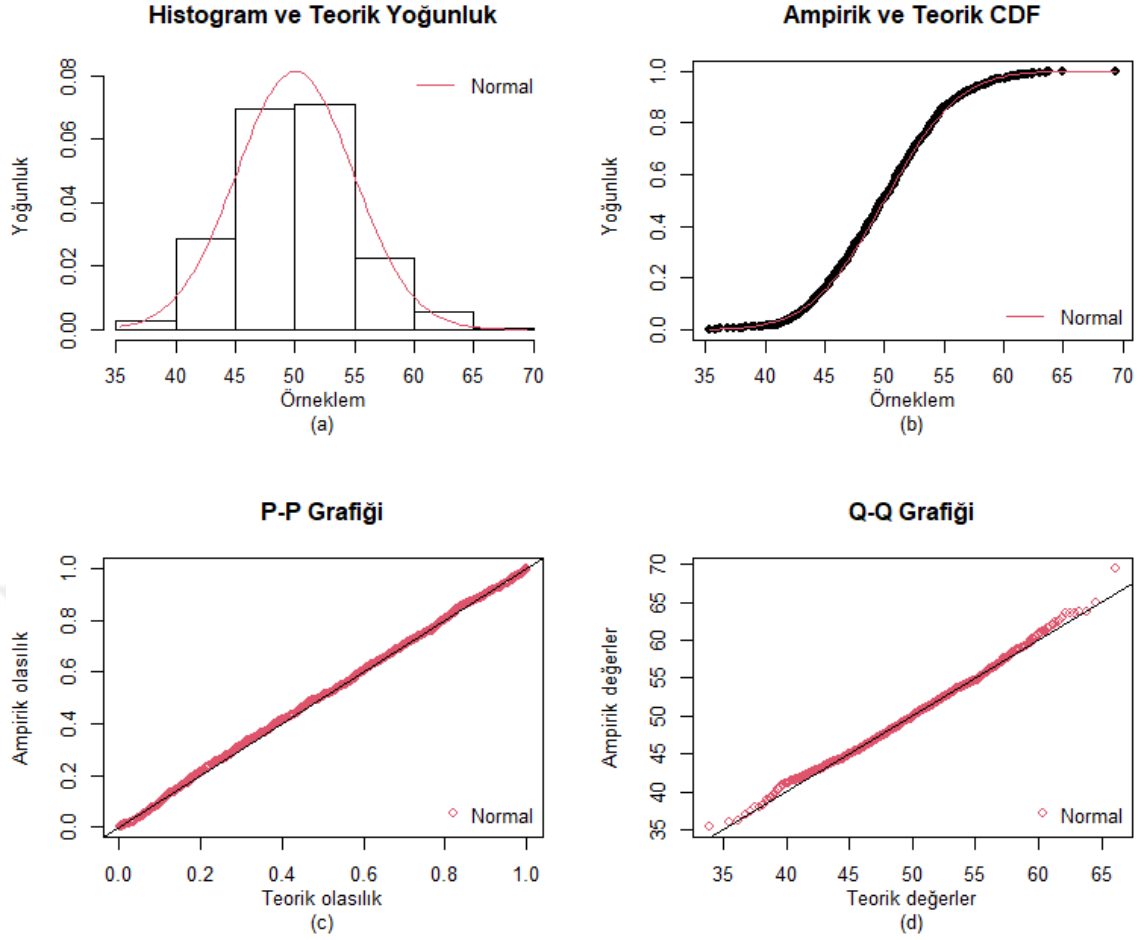
Olasılık yoğunluk fonksiyonu (OYF), olasılık teorisinde kullanılan bir fonksiyondur ve sürekli rassal değişkenlerin dağılımını açıklamak için kullanılır. OYF, rassal değişkenin alabileceği tüm olası değerlerdeki olasılıkların yoğunluğunu ifade etmektedir. Bu yöntemde, bir dağılımın OYF'si çizilir ve dağılımın karakteristik özellikleri bu bir çizgi grafik üzerinde gösterilir. Yatay eksen de rassal değişkenin değerleri, dikey eksen de ise bu değerlerin olasılık yoğunluğu yer almaktadır. Grafik altında kalan alan, rassal değişkenin tüm olası değerlerindeki olasılık toplamı olan 1 değerine eşittir. OYF, sürekli rassal değişkenlerin dağılımının incelenmesi için önemli bir araçlardan bir tanesidir. Bu fonksiyon, rassal değişkenin alabileceği olası değerler ve bu değerlerin olasılık yoğunlukları hakkında bilgi sağlamaktadır. Ayrıca, OYF'ler aracılığıyla farklı grup verilerinin karşılaştırılması da mümkündür.

Quantile-Quantile (Q-Q) grafiği, teorik bir dağılım ve gerçek veri kümesinin dağılımını karşılaştırmak ve bu veri kümelerinin seçilen aday dağılıma uygunluğunu test etmek için kullanılan bir grafik yöntemidir. Grafikte, bir eksen teorik dağılımın OYF'den elde edilen değerleri gösterirken, diğer eksen gerçek gözlemlerin değerlerini gösterir. Quantile, verinin belirli bir

yüzdesini ifade eder ve bu yüzde, gözlemlerin küçükten büyüğe sıralanması sonucu belirlenir. Eğer gerçek veri teorik dağılıma uygunsa, Q-Q grafiği üzerinde noktalar teorik çizgi üzerinde yer almaktadır. Eğer gerçek veri teorik dağılıma uymuyorsa, noktalar teorik çizgiye yakın olmaz veya teorik çizgiyi aşar. Q-Q grafiği, normal dağılımın yanı sıra diğer dağılım türleri için de kullanılır. Ayrıca, Q-Q grafiği, aykırı değerlerin tespiti ve gözlemlerin seçilen aday dağılıma uygunluğunun test edilmesi gibi istatistiksel analizlerde de kullanılır.

Bir P-P grafiği, bir veri setinin ampirik kümülatif dağılım fonksiyonunu belirtilen teorik kümülatif dağılım fonksiyonu (KDF) ile karşılaştırır. Bir Q-Q grafiği ise, bir veri dağılımının niceliklerini belirli bir dağılım ailesinden standartlaştırılmış bir teorik dağılımın nicelikleriyle karşılaştırır. P-P grafiklerinin ve Q-Q grafiklerinin oluşturulma ve yorumlanma biçiminde üç önemli fark vardır:

- Bir Q-Q grafiğinin oluşturulması, teorik KDF'nin konum veya ölçek parametrelerinin belirtilmesini gerektirmez. Teorik nicelikler, belirtilen aile içindeki standart bir dağılımdan hesaplanır. Doğrusal bir nokta modeli, belirtilen ailenin veri dağılımını uygun bir şekilde tanımladığını ve konum ve ölçek parametrelerinin, doğrusal modelin kesişme noktası ve eğimi olarak görsel olarak tahmin edilebileceğini gösterir. Bununla, bir P-P grafiğinin oluşturulması, sıralı veri değerlerinde KDF'yi değerlendirmek için teorik kümülatif dağılım fonksiyonunun konum ve ölçek parametrelerini gerektirir.
- Bir Q-Q grafiğindeki nokta deseninin doğrusallığı, konumdaki veya ölçekteki değişikliklerden etkilenmez. Bir P-P grafiği çiziminde, konumdaki veya ölçekteki değişiklikler mutlaka doğrusallığı korumaz.
- Bir Q-Q çiziminde, belirli bir teorik OYF'yi temsil eden referans çizgisi, konum ve ölçek parametrelerine eşit kesişim ve eğime sahip olan bu dağılımın konum ve ölçek parametrelerine bağlıdır. Bir P-P grafiğinde, herhangi bir dağılım için referans çizgisi her zaman $y = x$ düzlem çizgisidir.



Şekil 3.14. $N(50, 5)$ Dağılımının Normal Dağılıma Uyum Grafikleri

Şekil 3.14.'te, ortalama değeri 50 ve standart sapma değeri 5 olan bir normal dağılım ele alınmıştır. Bu dağılımın, histogram, OYF, KDF, P-P ve Q-Q grafikleri farklı dağılım aileleri ile elde edilerek nasıl bir görünüme sahip olduğu incelenecektir.

Şekil 3.14. (a), normal dağılım teorik OYF'si ile gözlenen değerlerin histogramıdır. Gözlenen değerlerin normal dağılıma uygun bir yayılım gösterdiği ilk bakışta söylenebilir.

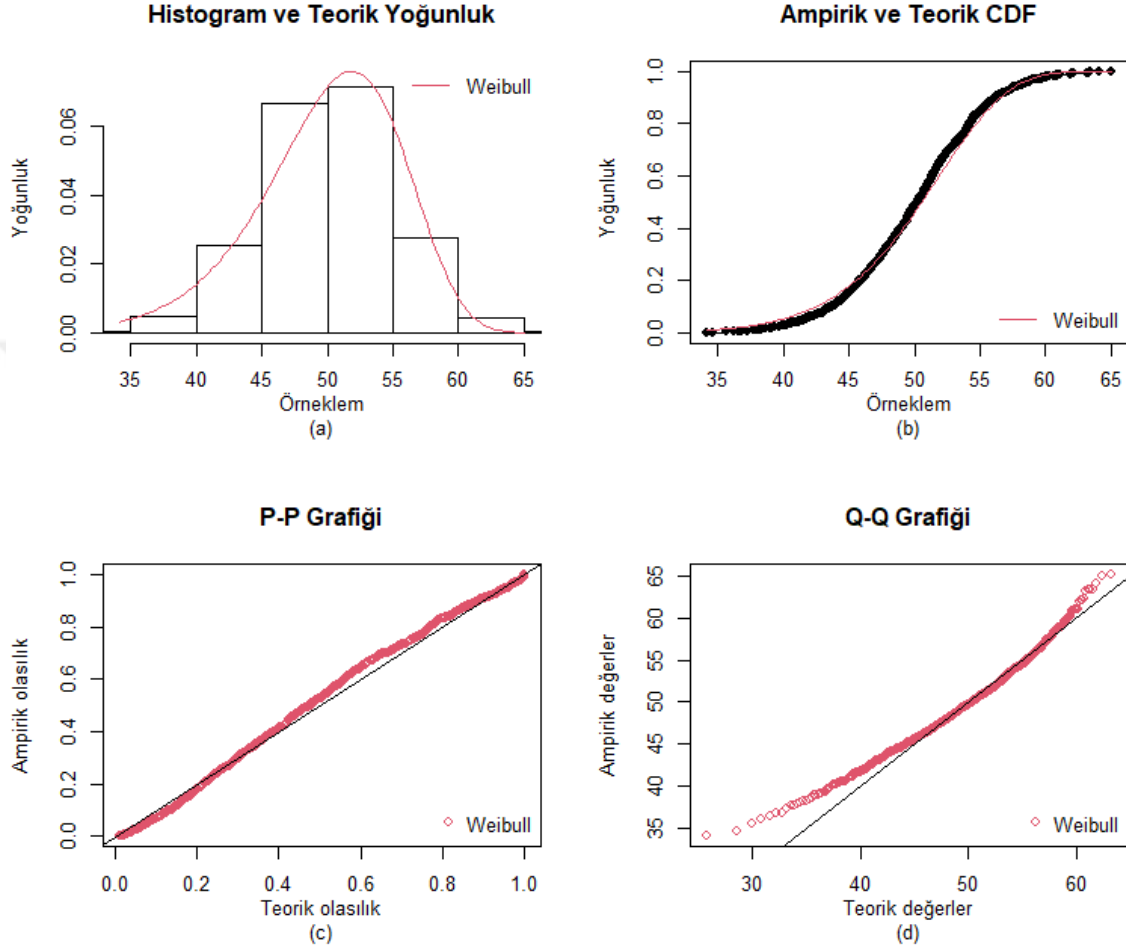
Şekil 3.14. (b), gözlenen verinin ampirik KDF'si ile, seçilen normal dağılımın KDF grafiğidir. Gözlenen veri KDF'sinin seçilen normal dağılım KDF'sine iyi bir uyum sağladığı görülmektedir.

Şekil 3.14. (c), teorik olasılık değerinin ampirik olasılık değeriyle kıyaslandığı P-P grafiğidir. Gözlenen verinin seçilen dağılıma uyması durumunda, veri P-P grafiğindeki lineer çizgi üzerine konumlanır. Görüldüğü üzere, verisi temsil eden noktalar düz bir çapraz çizgi üzerinde yer almıştır. Böylece gözlenen verinin normal dağılıma uyduğu söylenebilir.

Şekil 3.14. (d) ise, teorik değerlerin, seçilen dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. Bu grafik, bir veri kümesinin potansiyel olarak seçilen teorik dağılımlardan gelip gelmediğini değerlendirmek için kullanılır. P-P grafiğine benzer şekilde, veri seçilen dağılıma

uyarsa, Q-Q grafiğinde noktalar düz bir çapraz çizgi üzerinde yer alır. Şekil 3.14. (d) incelendiğinde, gözlemlerin seçilen normal dağılımı takip ettiği yorumu yapılabilir.

Genel olarak Şekil 3.14. değerlendirildiğinde, analiz edilen verinin aday dağılım olarak seçilen normal dağılıma uygun olduğu görünmektedir.



Şekil 3.15. N(50, 5) Dağılımının Weibull Dağılıma Uyum Grafikleri

Ortalama değeri 50 ve standart sapma değeri 5 olan bir normal dağılımdan üretilen sayıların, Weibull dağılıma uyumu Şekil 3.15.'de çizilen grafiklerle izlenmek istenmiştir.

Şekil 3.15. (a), Weibull dağılım teorik OYF'si ile gözlenen değerlerin histogramıdır. Gözlenen değerlerin Weibull dağılıma uygun bir yayılım gösterip göstermediği net görülememektedir.

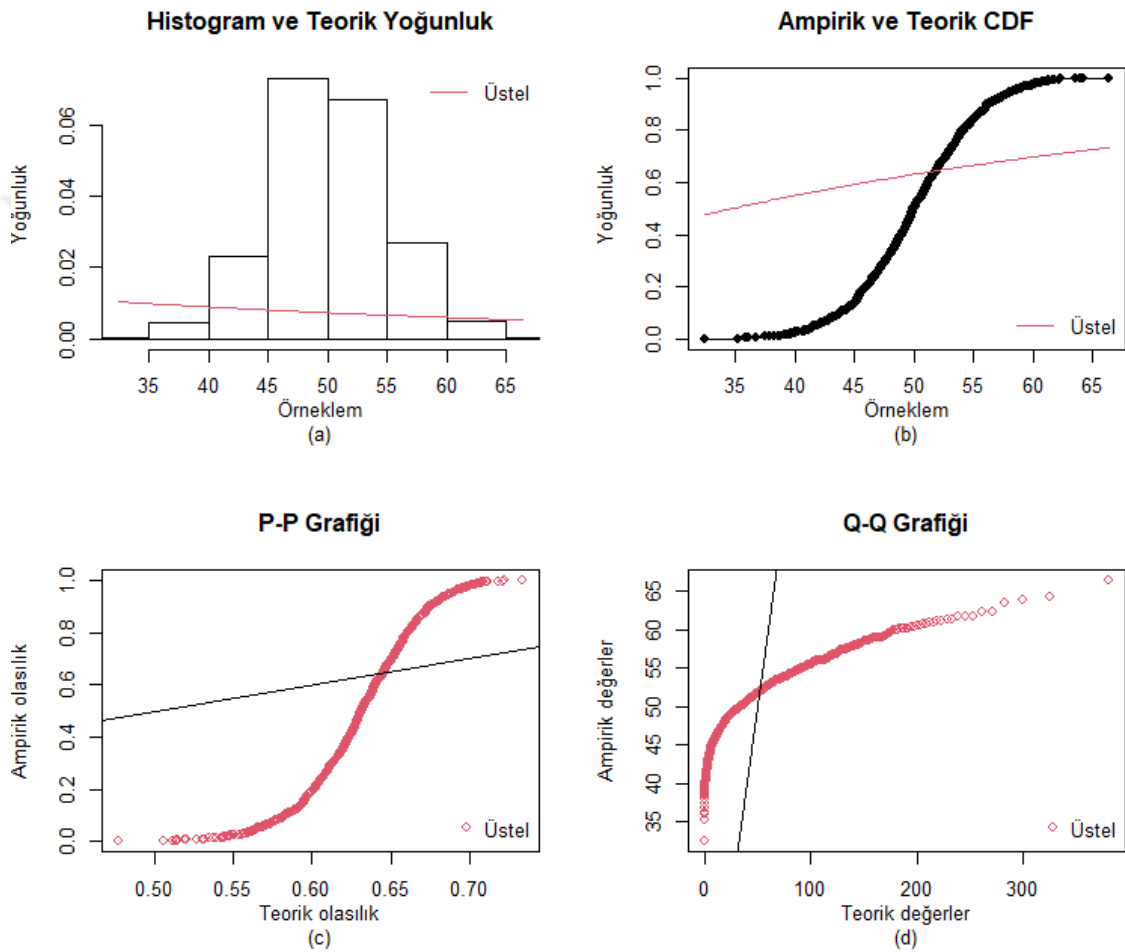
Şekil 3.15. (b), gözlenen verinin ampirik KDF'si ile, seçilen Weibull dağılımın KDF'si grafiğidir ve karşılaştırıldığında sapmalar olduğu görülmektedir.

Şekil 3.15. (c), teorik Weibull olasılık değerinin ampirik olasılık değeriyle kıyaslandığı P-P grafiğidir. Görüldüğü üzere, verisi temsil eden noktalar düz bir çapraz çizgi üzerinde yer almak

yerine, salınımlar yapmıştır. Böylece gözlenen verinin Weibull dağılıma uyum göstermediği söylenebilir.

Şekil 3.15. (d) ise, teorik değerlerin, seçilen dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. Q-Q grafiğinde gözlenen bariz sapmalar sayesinde, gözlenen verinin Weibull dağılımını izlemediği yorumu rahatlıkla yapılabilir.

Genel olarak Şekil 3.15. değerlendirildiğinde, analiz edilen verinin $N(50, 2)$; aday dağılım olarak seçilen Weibull dağılıma uygun olmadığı görülmektedir.



Ortalama değeri 50 ve standart sapma değeri 5 olan bir normal dağılımdan üretilen sayıların, üstel dağılıma uyumu Şekil 3.16'de görüldüğü gibi grafiklerle test edilmiştir.

Şekil 3.16. (a), üstel dağılım teorik OYF'si ile gözlenen değerlerin histogramıdır. Gözlenen değerlerin üstel dağılıma uygun bir yayılım göstermediği çok net bir şekilde görülmektedir.

Şekil 3.16. (b), gözlenen veri KDF'sinin seçilen üstel dağılım KDF'siyle karşılaştırıldığında çok büyük sapmalar olduğu görülmektedir. Böylece verilen üstel dağılıma uymadığı yorumu rahatlıkla yapılabilir.

Şekil 3.16. (c), teorik üstel olasılık değerinin ampirik olasılık değeriyle kıyaslandığı P-P grafiğidir. Görüldüğü üzere, veri hiçbir şekilde P-P grafiğine uyum göstermemektedir.

Son olarak Şekil 3.16. (d) ise, teorik değerlerin, seçilen dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. Q-Q grafiğinde, P-P grafiğine benzer olarak gözlenen verinin üstel dağılıma kesinlikle uymadığı rahatlıkla söylenebilir.

Histogram ve yoğunluk grafikleri, KDF, P-P ve Q-Q grafikleri, bir veri setinin aday bir dağılıma uyup uymadığını bazı durumlarda çok net bir şekilde görülebilse de her zaman yeterli bir yaklaşım değildir. Güvenilir bir tahmin için grafikler yardımıyla sadece uygun olabilecek aday istatistiksel dağılımlar seçilmelidir. Aday istatistiksel dağılımlar seçildikten sonra bu dağılımlara ait parametre tahminlerinin yapılması gerekir. Parametre tahminleri yapıldıktan sonra, tahmin edilen parametrelere sahip sayılar üretilir ve istatistiksel uygunluk testleriyle sınanarak en uygun parametreye sahip dağılım modeli seçilir.

3.4. Parametre Tahmini

Üçüncü adımda seçilen "aday" dağılımlara, veriye dayalı olarak dağılım parametrelerini tahmin etmeye olanak tanıyan istatistiksel yöntemler uygulanmalıdır. Parametre tahmininde bilgisayar programları kullanılabilir. Parametre tahmininde en sık kullanılan yöntem en çok olabilirlik (maximum likelihood ML) yöntemidir.

En çok olabilirlik yöntemi, istatistiksel bir modelin parametrelerini tahmin etmek için kullanılan istatistiksel bir yöntemdir ve 1922’de Fisher tarafından ortaya atılmıştır (Fisher, 1922). Gözlenen verinin model parametrelerinin bir fonksiyonu olan olabilirlik fonksiyonunu maksimize eden parametre değerlerinin bulunması prensibine dayanır. Olabilirlik işlevi, belirli bir parametre değerleri kümesi verilen verisi gözlemeleme olasılığıdır. Parametrelerin maksimum olabilirlik tahmini, bu olasılığı en üst düzeye çıkaran değerler kümesidir, yani gözlemlenen verisi en olası kılın parametre değerleridir. En çok olabilirlik yönteminin diğer parametre tahmin yöntemlerine göre birçok avantajı bulunmaktadır ve çok çeşitli istatistiksel modellere uygulanabilen esnek bir yöntemdir (Hogg ve Craig, 1989). Aynı zamanda, tutarlılık ve asimptotik verimlilik gibi iyi istatistiksel özelliklere sahip tahminler sağlar. Bu yöntem, doğrusal regresyon, lojistik regresyon ve hayatta kalma analizi modelleri dahil olmak üzere birçok farklı istatistiksel modelin parametrelerini tahmin etmek, hipotez testi ve model seçimi için olduğu kadar güven aralıkları oluşturmak ve çıkarım yapmak için kullanılır (Larson, 1982). Bu yöntemin dezavantajı ise, bazı durumlarda maksimizasyon problemi çözümü sırasında yakınsama sıkıntısı probleminin yaşanmasıdır (Mendenhall ve ark., 1990).

$x_1, x_2, \dots, x_n$ veri setinin, olasılık yoğunluk fonksiyonu $f(x; \theta)$ olan bir örneklem grubundan geldiği varsayımı yapılsın. Bu durumda θ parametresi için olabilirlik fonksiyonu, $L(\theta; x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$ şeklinde gösterilir. Bu denklemi

maksimum yapan değer, θ 'nın en çok olabilirlik tahmin değeridir. Daha genel bir gösterimle θ 'nın en çok olabilirlik tahmin edicisi $\hat{\theta}(x_1, x_2, \dots, x_n) = \max L(\theta; x_1, x_2, \dots, x_n)$ şeklinde hesaplanır.

Çizelge 3.3. farklı dağılım türleri için en çok olabilirlik tahmin edicilerini göstermektedir. Bu tahmin ediciler, veriye dayalı en olası parametre değerlerini hesaplamak için kullanılır. Bu tahmin ediciler, istatistiksel analizde önemli bir rol oynar ve veriye uygun parametre değerlerini sağlar.

Çizelge 3.3. İstatistiksel Dağılımların En Çok Olabilirlik Tahmin Edicileri (Karaman, 2019)

Dağılım Türü	Parametresi	ML Tahmin Edici
Binom	n : Deneme sayısı p : Başarı gerçekleşme olasılığı k : Seçilen özelliğin gözlenen sayısı	$\hat{p} = \frac{\sum_{i=1}^n k_i}{\sum_{i=1}^n n_i}$
Poisson	n : Deneme sayısı λ : Gerçekleşen ortalama olay sayısı	$\hat{\lambda} = \frac{\sum x_i}{n}$
Normal	μ : Ortalama σ : Standart sapma	$\hat{\sigma}^2 = \frac{\sum u_i^2}{n}$
Üstel	x : Ortalama süre θ : Süre oranı	$\hat{\theta} = \frac{\sum x_i}{n}$

3.5. Uygunluk Testi

Dağılımlar parametreleri tahmin edildikten sonra, seçilen dağılımların veriye ne kadar uyduğu belirlenir. Bu, belirli uyum iyiliği testleri kullanılarak veya ampirik (örnek veriye dayalı) ve teorik (uydurulmuş) dağılım grafiklerini karşılaştırarak görsel olarak yapılabilir. Sonuç olarak, veriyi açıklayan en geçerli model seçilmelidir. İstatistiksel uygunluk testleri, bir süreç veya süreçler hakkında nicel kararlar almak için bir mekanizma sağlamaktadır. Amaç, süreçle ilgili bir varsayımı veya hipotezi "reddetmek" için yeterli kanıt olup olmadığını belirlemektir. Varsayım sıfır hipotezi (H_0) olarak adlandırılır.

Ki-kare (χ^2) uyum iyiliği, bir değişkenin belirli bir aday dağılıma uyma olasılığının olup olmadığını belirlemek için kullanılan istatistiksel bir hipotez testidir. Bu test hem kesikli hem de sürekli dağılımlar için kullanılabilir. Örnek veri, H_0 hipotezi altındaki varsayımsal dağılıma uymaktadır. χ^2 uyum iyiliği testi için p -değeri anlamlılık düzeyinden (α) küçükse, H_0 hipotezi reddedilir (Snedecor ve Cochran, 1989). χ^2 uyum iyiliği hesaplaması için, test istatistiği O_i gözlenen, E_i beklenen frekans olmak üzere $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ şeklinde tanımlanır.

χ^2 uyum iyiliği testinde kullanılan veri gözlenen frekanslardır. Bu, her kategori için yalnızca bir veri değeri olduğu anlamına gelmektedir. Kategorilerin sırası değiştirilirse test istatistiğinin değeri değişmez. Hesaplanan ağırlıklı kare sapmaların toplamı küçükse, gözlemlenen frekanslar beklenen frekanslara yakındır ve o dağılımdan geldiği iddiasını reddetmek için hiçbir sebep yoktur. Sadece toplam büyük olduğunda dağılımı sorgulamak için bir nedendir. Bu nedenle, χ^2 uyum iyiliği testi her zaman bir sağ kuyruk testidir.

Kolmogorov-Smirnov (K-S) testi, parametrik olmayan bir istatistik testidir ve Andrey Kolmogorov ve Nikolai Smirnov adlı iki Sovyet bilim insanı tarafından geliştirilmiştir. χ^2 uyum iyiliği testi genel olarak gözlenen dağılımın varsayılandan önemli ölçüde farklı olup olmadığını test ederken, K-S testi tüm verisi kullanarak hipotezi test eder. Bu nedenle K-S testi, χ^2 uyum iyiliği testine göre daha güçlü bir alternatiftir ve daha genel uygulanabilirliğe sahiptir (Law, 2007).

Dağılımların uyum iyiliğini test etmek için kullanılan bir başka test ise Anderson-Darling (A-D) testidir ve χ^2 ve K-S uyum iyiliği testlerine bir alternatif olarak kullanılabilir (Stephens, 1976). Anderson-Darling (A-D) istatistiği, verinin belirli bir dağılımı ne kadar iyi takip ettiğini ölçmektedir. Belirli bir veri seti ve dağılımı için, dağılım gözleme ne kadar iyi uyarsa, bu istatistik o kadar küçük olur. Örneğin, veri kümesinin bir t testi için normallik varsayımını karşılayıp karşılamadığını belirlemek için A-D istatistiği kullanılabilir. Benzer şekilde A-D istatistiği, bir veri örneğinin belirli bir dağılıma sahip bir popülasyondan gelip gelmediğini test etmek için kullanılır. A-D testi için hipotezler şunlardır:

H_0 : Veri seçilen dağılımı takip eder.

H_1 : Veri seçilen dağılımı takip etmez.

Verilerin seçilen dağılımdan gelip gelmediğini test etmek için karşılık gelen p -değeri kullanılır. p -değeri seçilen anlam düzeyinden küçükse, gözlemlerin bu dağılımdan geldiği H_0 hipotezi reddedilir.

A-D testi kuyruk farklılıklarını tespit eder ve bazı istatistiksel dağılımlar için K-S testinden daha yüksek bir güce sahiptir. K-S testi seçilen dağılımlardan bağımsız olarak hesaplanırken, A-D testi, kritik değerlerin hesaplanmasında her dağılım için oluşturulmuş özel tablolardan yararlanır. Bu durum, daha duyarlı bir teste izin verme avantajına ve her dağılım için ayrıca kritik değerlerin hesaplanması gerektiği dezavantajına sahiptir.

Genel olarak; seçilen her bir aday dağılım için hesaplanan uyum iyiliği testlerinin p -değeri kritik değer ile karşılaştırılarak hipotezin kabulü ya da reddi yapılır. Düşük bir p -değeri, varsayımın yanlış olduğu ve gözlemlerin dağılıma uymadığı, yüksek bir p -değeri, varsayımın doğru olduğu ve dağılıma uyduğu anlamına gelmektedir. Tüm bu testlerin uygulanmasında RStudio programı kullanılabilir.

3.6. Karma Dağılımlar

Tahmin yöntemleri nitel ve nicel olmak üzere iki gruba ayrılır. Nitel yöntemler kişisel görüş ve deneyime dayalıyken, nicel yöntemler sayısal verisi kullanır. Talep tahmin yöntemlerinde en kritik noktalardan biri, farklı modellerin farklı durumlarda daha iyi sonuçlar verebilmesidir. Başka bir deyişle, farklı sistemlerde farklı tahmin yöntemleri uygulanabilir (Stekler, 2007). En sık kullanılan girdi belirleme yöntemlerinden biri olan teorik dağılıma uydurma, girdi verisine her zaman iyi uymayabilir. Sistem birden fazla faktör nedeniyle daha karmaşık hale geldikçe, giriş

verisinin bilinen bir saf dağılıma sığdırmak zorlaşır. Bu gibi durumlarda sistem ampirik bir dağılımla temsil edilmeye çalışılır. Girdi verisi büyük olduğunda veya belirli bir aralıkta kümelandığında, veri tam olarak temsil edilemeyebilir. Diğer bir yol olan zaman serisi analiz yöntemi, zaman içinde birden fazla faktörden etkilenen girdi verisinde doğru tahminler yapamayabilir. Uygulanan istatistik, gözlemlerin tek bir popülasyonda toplandığı varsayımına dayanmaktadır. Bu homojenlik varsayımı genellikle gerçekçi değildir ve girdi verisini doğru bir şekilde temsil etmek için yeterli değildir. Bu noktada, karma dağılım modelleri, farklı parametrelere sahip heterojen alt popülasyonların varlığını dikkate alan bir yaklaşım sundukları için veri girdisini temsil etmede önemli bir rol oynamaktadır.

Karma dağılım modelleri, doğal olaylardaki rasgelelik içeren özelliklerin ölçüm değerlerini istatistiksel olarak modellemek için matematiksel bir yaklaşım sunar. Bu yöntem, biyoloji, genetik, tıp, psikiyatri, astronomi, ekonomi, mühendislik ve pazarlama gibi birçok alanda kullanılmaktadır. Karma dağılım modelleri, sınıflandırma yöntemleri olarak da bilinir ve yığılma analizi, ayrıştırma analizi ve görüntü analizi gibi istatistiksel teknikleri içermektedir. Karma dağılımlar, farklı bileşen dağılımların ağırlıklı kombinasyonunu kullanarak daha karmaşık verisi modellemek için kullanılır. Bu dağılımlar, genellikle gerçek hayattaki karmaşık olayları ve verisi daha iyi açıklamak için kullanılır. Bu nedenle, karma dağılım modelleri teorik ve pratik uygulamalarda önemli bir yer tutmaktadır ve kullanımı giderek artmaktadır. Karma dağılımlar, zaman serisi modellerini temsil etmek için iyi bir araç olarak tanımlanabilir ve doğrudan kısa ila orta vadeli tahmin için kullanılabilir (Eirola ve Lendasse, 2013).

Karma dağılımların olasılık yoğunluk fonksiyonu, karma dağılımı oluşturan diğer dağılımların olasılık yoğunluk fonksiyonlarının ağırlıklı toplamı olarak ifade edilir. Her bir bileşenin ağırlığı, o dağılımın karma dağılıma ne kadar etkili olduğunu belirtir. Olasılık yoğunluk fonksiyonu, farklı bileşenlerin etkileşimini ve ağırlıklarını dikkate alarak veri setinin dağılımını modellemeye çalışır.

Karma dağılımların kümülatif olasılık yoğunluk fonksiyonu, olasılık yoğunluk fonksiyonuna benzer şekilde, karma dağılımı oluşturan diğer dağılımların kümülatif olasılık yoğunluk fonksiyonlarının ağırlıklı toplamı olarak ifade edilir ve veri setindeki bir değer için belirli bir değeri aşma olasılığını ifade eder.

Sonlu bir karma dağılımı, iki ya da daha fazla durumunu modellemek için uygulanabilecek bir dışbükey dağıtım işlevi kombinasyonu olarak düşünülebilir (Titterton ve ark., 1985). Daha basit bir tanımla iki ya da daha fazla bileşenden oluşan dağılımları karma dağılım olarak tanımlanır.

Karma oranı $(\pi_1, \pi_2, \dots, \pi_k)$, bileşen sayısı (k) ve dağılım parametreleri $f_1(x; \theta_1), f_2(x; \theta_2), \dots, f_k(x; \theta_k)$ olmak üzere üç komponent ile tanımlanır.

$$f(x) = \pi_1 f_1(x; \theta_1) + \pi_2 f_2(x; \theta_2) + \dots + \pi_k f_k(x; \theta_k)$$

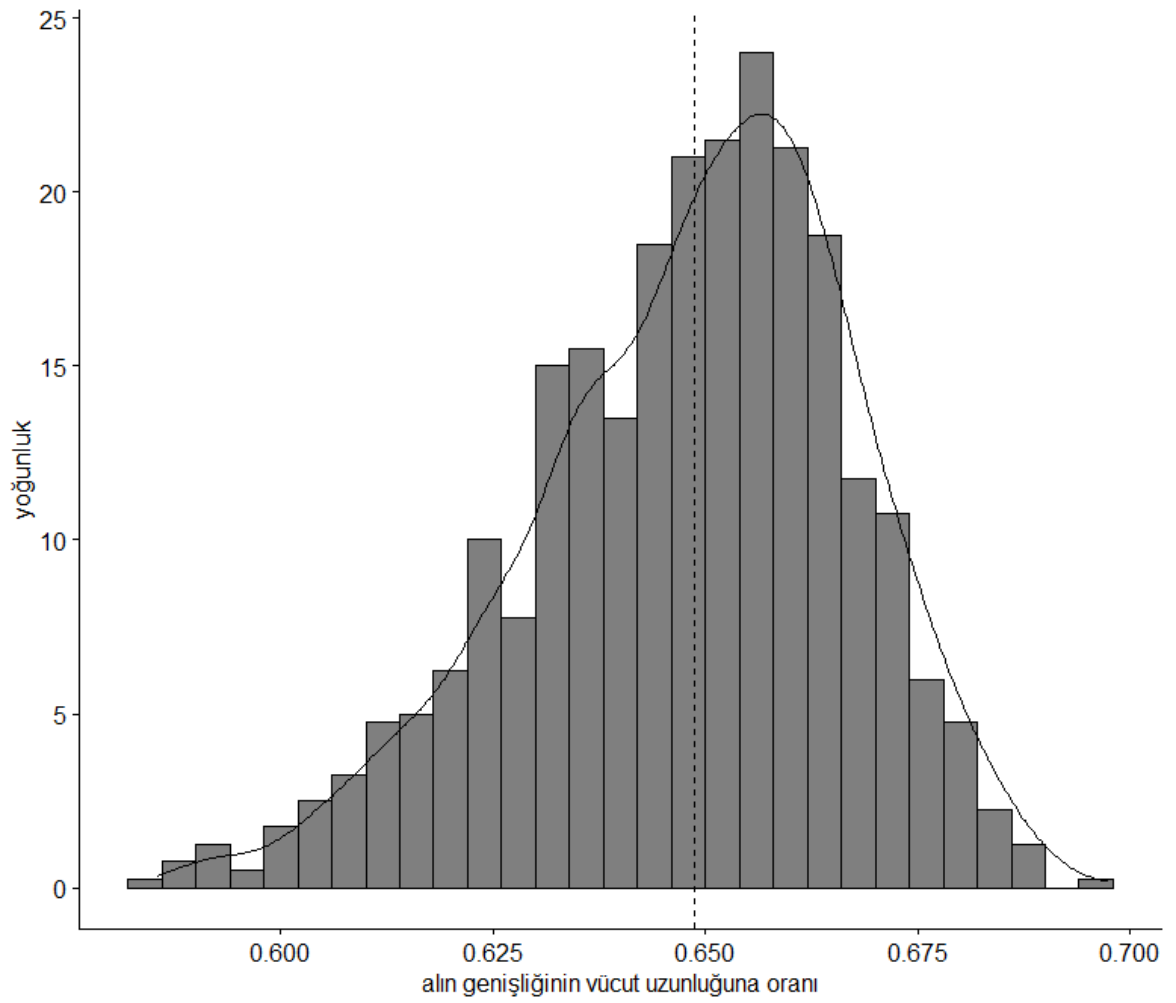
$$0 \leq \pi_i \leq 1 \quad i = 1, 2, \dots, k$$

$$\sum_{i=1}^k \pi_i = 1$$

Karma dağılımların genel bir gösteriminde parametreler için ψ sembolünü kullanılır.

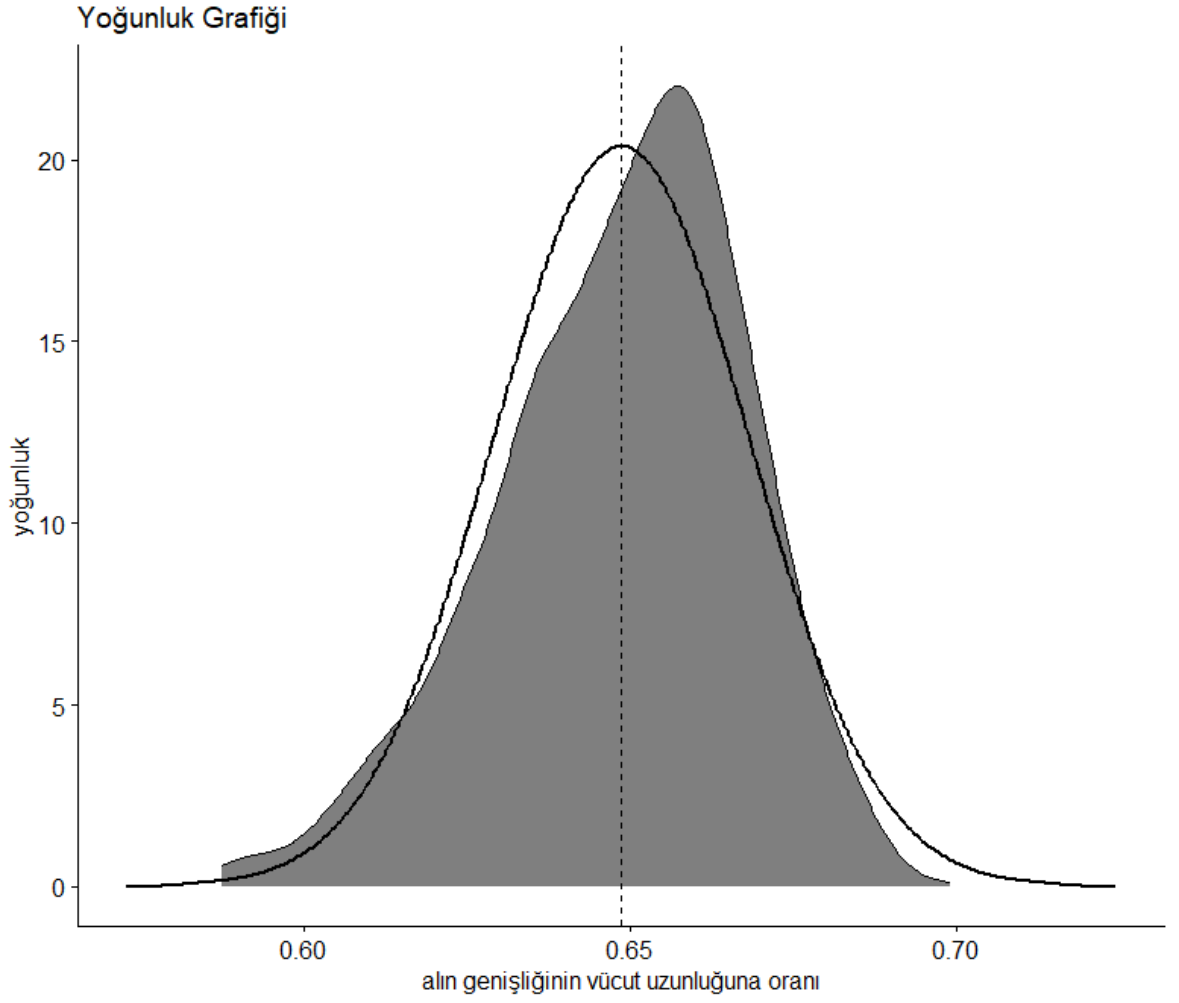
$$f(x, \psi) = \sum_{i=1}^k \pi_i f(x; \theta_i)$$

1892'de Weldon ve eşi, yengeç *Carcinus Moenas*'ın popülasyonları hakkında kapsamlı veri toplamış ve Napoli körfezindeki toplamda 1000 yengeçten topladığı alın genişliğinin vücut uzunluğuna oranının normal dağılımdan oldukça çarpık olduğunu gözlemlemişlerdir (Weldon, 1892). Weldon, bu dağılımın, popülasyonun aynı yörede iki farklı ırktan veya "tipten" oluşması ihtimaliyle birlikte, popülasyonun iki farklı normal dağılımın bir karması olmasının bir sonucu olup olmadığını araştırmak istemiştir. Verilerin histogramının asimetric olduğunu ya da simetrik olmadığını belirlediği için yengeç kitlesinin iki farklı türden oluştuğunu düşünen Weldon matematiksel yönünün eksik olması nedeniyle, Pearson'dan yardım istemiş ve böylece karma dağılım kullanımı fikri ortaya çıkmıştır. Pearson, tek değişkenli iki bileşenli karma dağılım modelinde ortalama, varyans ve karma oranlarına sahip özelliklerin incelenmesine odaklanmıştır. Pearson ilk beş momentin dokuzuncu dereceden bir polinom türetmiş ve bu polinomun gerçek köklerini bulmuştur. Her kök, ilk beş momentle eşleşen bir aday karma dağılımı vermektedir. Tüm bu karmaşık algortima hesaplamaları o zamanın şartlarıyla, herhangi bir bilgisayar makinesinin yardımı olmadan Pearson tarafından elle çözülmüştür.



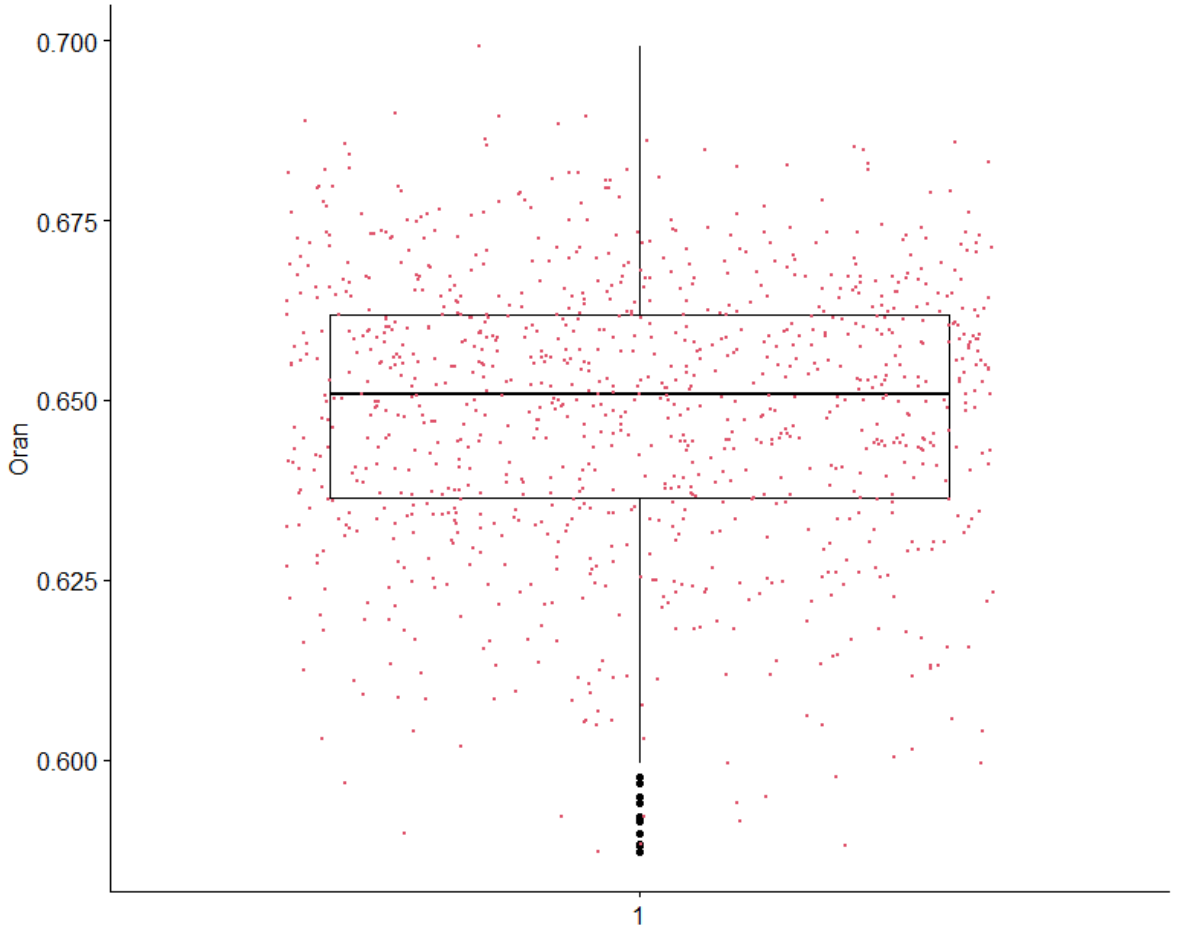
Şekil 3.17. Pearson'un Yengeç Verisi Histogramı ve Yoğunluk Grafiği

Şekil 3.17.'de Pearson'un kullandığı Weldon'un yengeç verisi histogramı ve yoğunluk grafiği çizilmiştir. Kesikli ortalama değer çizgisinin, normal dağılım ortalama değerinin solunda olduğu ve normal dağılıma uyum sağlamadığı grafiksel olarak görülmektedir.



Şekil 3.18. Pearson'un Yengeç Verisi Normallik Kontrolü

Şekil 3.18.'de, gözlenen yengeç verisi OYF'si ile aynı ortalama ve varyansa sahip bir normal OYF arasında karşılaştırma yapılmıştır. Bu grafikte; sapma, gözlenen veri eğrisinin asimetrik olduğunu vurgulamaktadır ve bu durum veri setinin normal dağılıma uymadığına dair bir ipucudur.



Şekil 3.19. Pearson'un Yengeç Verisi Kutu Grafiği

Şekil 3.19. Pearson ve Weldon'ın yengeç verisiyle oluşturulmuş kutu grafiğini göstermektedir. Kutu grafiği, medyanın asimetrisini ve birçok aykırı değer varlığını göstermektedir. Kutu grafiğine bakarak, bu verinin normal dağılımla modellenemeyeceği ihtimali giderek güçlenmektedir.

Genel olarak, hastanede kalış süresi ve hastaneye gelişler arası süre dağılımları asimetriktir. Bu veri sıklıkla gamma, lognormal ve Weibull dağılımları ile temsil edilmektedir (Griffiths ve ark., 2006; Atienza ve ark. 2008). Asimetriklik ve birden fazla grupların oluşturduğu faktörlerin mevcut bulunduğu durumlar söz konusu olduğunda, karma dağılım modeli vasıtasıyla tanımlanabilir (Kessler ve McDowell, 2012). Bu nedenle, tek bir saf dağılım kullanmak yerine, daha iyi sonuçlar almak için kalış süresi ve gelişler arası sürenin tanımlanacağı çalışmalarda, karma dağılım kullanılması literatürde önerilmektedir (Atienza ve ark., 2008).

Karma dağılım modellemesi, çalışılan popülasyonun grup yapısını tanımlar ve küme analizi gibi önemli istatistiksel alanlarda esnek ve güçlü bir olasılık yaklaşımı sağlamaktadır (Ng ve ark., 2019). Kümeleme analizi uygulamalarında, verisi birkaç kümeye gruplamak, karma dağılım yoluyla sağlanır. Böylece kümeler arasında en az, küme için en fazla benzerliğe sahip veri grupları elde edilmeye çalışılır. Karma dağılıma dayalı kümeleme yaklaşımları, verinin grup yapısı

hakkında önceden bilgi sahibi olmadan, her bir varlıkla ilişkili değişkenlere dayalı olarak gözlemlenen özelliklerin ölçülmesiyle gerçekleştirilir. Bu duruma denetimsiz sınıflandırma denir. Karma dağılım modelleri, kümeleme için sağlam bir matematiksel çerçeve sağlar ve bileşen sayısını belirlemek ve sınıflandırma belirsizliğini değerlendirmek için istatistiksel çıkarımlar sağlanmaktadır (McLachlan ve Peel, 2000). Karma dağılımları, çarpık dağılımlardan oluşan veri modellenmesinde de oldukça etkilidir (Papi ve Pontecorvi, 2015).

Karma dağılım modeli oluşturulurken karşılaşılan problemlerden en önemlisi, veri setinin birden fazla dağılımın karmasından oluşup oluşmadığının tespiti ve bu dağılımların belirlenebilmesidir (Atienza ve ark., 2008). Bileşenlerin tespiti problemi, dağılım parametrelerini tahmin etme problemine yol açmaktadır. Karma dağılımı bileşen sayısını tahmin etmek henüz tamamen çözülmemiş zorlu bir süreçtir (McLachlan ve Peel, 2000). Bu konudaki belirsizlik bir takım parametrik varsayımlar ve yığılma analizi uygulamasıyla çözülebilmektedir.

Bir veri setinin karma dağılım olarak modellenebilmesi için ilk olarak bileşen sayısının tahmin edilmesi gerekmektedir. İkinci aşamada karma oranları, üçüncü ve son aşama olarak dağılım parametreleri tahmin edilir.

Optimal bileşen sayısı (k) sayısını bulmanın ana amacı, verisi iyi ayrılmış kümelere bölmektir (Arabie ve ark., 1996). Bir başka anlatımla, heterojen veri setini homojenlik gösteren anlamlı parçalara bölmektir (Fraley ve Raftery, 1998). En uygun karma dağılımı ve optimal bileşen sayısı bulunduğundan sonra karma dağılımının parametreleri tahmin edilmelidir.

Karma dağılımla temsil edilecek veri setlerindeki dağılım türleri ve sayıları hakkında bilgi sahibi olmamak, bileşen sayısının belirlenmesinde en önemli zorluklardan biridir. Bu nedenle, bileşen sayısı belirlenirken bilgi kriteri analizinin de yapılması gereklidir (Bozdoğan, 1994). Bilinen bileşen sayısına sahip karma dağılımlı veri setlerinin parametre tahminleri, farklı yöntemler kullanılarak gerçekleştirilir.

Karma dağılımlı veri setlerinde bileşen sayısı tahmini için, OYF, P-P ve Q-Q grafikleri, loglikelihood değeri, Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) gibi yöntemler kullanılır. Bu yöntemler, bileşen sayısının doğru şekilde belirlenmesinde önemli rol oynar (Çalış, 2005).

3.6.1. Karma Dağılımlarda Bileşen Sayısı Belirleme

Veri madenciliği, son yıllarda oldukça popüler hale gelmiştir ve büyük veri kümelerindeki önemli bilgileri keşfetmek ve anlamak için kullanılır. Veri madenciliğinde kullanılan algoritmalar, veri kümesindeki gizli kalıpları ve ilişkileri keşfeder. Birçok veri madenciliği problemi, kümeleme gibi denetimsiz öğrenme teknikleri ile çözülebilir.

Kümeleme, benzer veri noktalarını gruplara ayırmak için kullanılan bir veri madenciliği tekniğidir. Kümeleme, müşteri segmentasyonu, görüntü işleme, biyoinformatik, coğrafi bilgi sistemleri gibi birçok farklı alan için sıklıkla kullanılmaktadır. Bir kümeleme algoritması, veri

kümesindeki benzer veri noktalarını bir araya getirerek kümeler oluşturur. Karmaşık veri kümelerinde, bu kümelerin sayısı belirlemek zor olabilmektedir. Bu nedenle, kümeleme algoritmaları genellikle veri kümesindeki kümelerin sayısını otomatik olarak belirleyebilen bir mekanizma ile donatılmıştır.

Çok değişkenli teknikler olan ayırıştırma analizi ve sınıflandırma, nesnelerin veya gözlemlerin farklı gruplara ayrılması ve önceden belirlenmiş grupların yeniden düzenlenmesiyle ilgilendir. Ayırıştırma analizi, genellikle sıradan ilişkilerin anlaşılması durumunda gözlenen farklılıkları araştırmak için kullanılır. Ayırıştırma ve sınıflandırmanın amacı ise sırasıyla; farklı özelliklere sahip gözlemleri grafiksel veya işlemsel olarak tanımlamak ve gözlemleri iki veya daha fazla isimlendirilmiş sınıflara ayırmaktır. Bu nedenle, ayırıştırma analizi ve sınıflandırma, çok değişkenli analizlerde önemli bir yere sahip tekniklerdir.

Karma dağılım modelleri temelde bir kümeleme problemi olarak adlandırılabilir. Karma dağılım modellerinin bileşenlerine ayırma probleminin tam olarak çözümlenememiş olduğu belirtilmektedir (McLachlan ve Peel 2000). Bu problemin zorluğu, farklı varyanslara sahip tek değişkenli iki dağılımın bir karmasını momentler metodunu kullanarak bileşenlerine ayırırken ortaya çıkmaktadır. Bu çözüm, Pearson tarafından önerilmiştir ve dokuzuncu dereceden bir polinomun köklerini bulmayı gerektirmektedir (Pearson, 1894). Cohen (1967), bu problemin tek bir negatif kök için çözülebileceğini göstermiş ve bunun için bir iterasyon metodu kullanmıştır. Ancak, Tan ve Chang (1972) ile Fryer ve Robertson (1972), momentler metodunun likelihood tahminine göre daha az başarılı olduğunu göstermişlerdir. Bileşenlerin varyansları eşit kabul edildiği koşullarda, bileşenlere dair tahminler sadece ilk dört momentten oluşturulan kübik eşitliğin negatif bir köküne dayanmaktadır (Rao, 1952). Karma dağılımların ortaya çıktığı ilk yıllarda, bilgisayar hesaplama sistemlerinin az gelişmiş olması nedeniyle karma dağılım bileşenlerini belirleyecek köklerin bulunması zorlu bir işlem olarak nitelendirilmekteydi. Ancak gelişen algortimalar, bilgisayarlar ve bilgi kriterleri sayesinde, günümüzde karma dağılım bileşen sayısı ve parametre tahmini çok zor bir problem olmaktan çıkmıştır.

Karma dağılımlarda bileşen sayısı belirleme, bir kümeleme algoritmasının hangi sayıda kümelerin var olduğunu belirlemesi için kullanılan bir tekniktir. Bu teknik, veri kümesindeki her bileşen için ayrı bir dağılım olduğunu varsayar. Daha sonra, farklı bileşen sayıları için farklı modeller oluşturulur ve her modelin uygunluğu ölçülür. En iyi model, veri kümesindeki en iyi uyumu veren modeldir. Bileşen sayısını belirlemedeki temel amaç oluşturulan kümeler içindeki gözlemlerin birbirine olan uzaklığın en küçük, farklı kümelerdeki gözlemlerin arasındaki uzaklığın ise en büyük olacak şekilde kümelenmesidir (Arabie ve ark. 1996). Bir başka anlatımla, heterojen veri setini homojenlik gösteren anlamlı parçalara bölmektir (Fraley ve Raftery, 1998).

1900'lü yılların ortalarına kadar karma dağılımların bileşenlerine ayırıştırılmasında moment tahmin edicilerin zorlu elde edilebilmesi ve modern teknoloji eksikliği nedeniyle grafiksel yöntemlerin kullanımına odaklanılmıştır (Fowlkes, 1979). Bu yöntemin öncü çalışmaları Harding

(1948) ve Cassie (1954) tarafından yapılmıştır ve Bhattacharya (1967) ile Wilk ve Gnanadesikan (1968) tarafından geliştirilmiştir. Chhikara ve Register (1979) verisi görselleştirmek amacıyla sayısal bir sınıflandırma tekniği geliştirirken, Tarter ve Silvers (1975) tarafından ise iki değişkenli normal yoğunluk fonksiyonunun özelliklerine dayalı bir grafiksel yöntem sunulmuştur.

Karma dağılımın bileşen sayısının belirlenmesi kolay olmayan bir süreçtir. Bileşen sayısının belirlenmesi bir karma dağılım analizin ilk aşamasıdır ve doğru bir istatistiksel temsil için bu aşamanın en uygun şekilde yönetilmesi gerekmektedir. Bileşen sayısı belirlendikten sonra karma dağılım homojen dağılımlara ayrılabilir ve alt sınıflar oluşturulabilir. Bileşen sayısı belirlemek için literatürde önerilen birden fazla yöntem vardır. Bunların ilki grafiksel yöntemlerdir. Grafiksel yöntemlerde OYF, P-P ve Q-Q grafiklerinden faydalanılır. Bu yöntem, karma dağılım bileşen sayısı hakkında kesin bir bilgi vermese de yardımcı olmaktadır.

Karma dağılım bileşen sayısı belirlemede bir diğer yöntem ise likelihood oran test istatistiğidir. Farklı bileşen değerlerinde likelihood oran test istatistiği izlenerek en iyi model seçilmeye çalışılır. Akaike ve Bayesci bilgi kriterleri (AIC, BIC); bileşen sayısı belirlenmesinde kullanılan bir diğer yöntemdir. Aday bileşen sayıları kullanılarak hesaplanan AIC ve BIC değerleri kıyaslanarak en uygun karma dağılım bileşen sayısı belirlenebilir. Karma dağılımlarda bileşen sayısının belirlenmesinde kullanılan bu yöntemler ayrıntılı bir şekilde açıklanmıştır.

Grafiksel Yöntemler

Karma dağılımlarda bileşen sayısı (k) belirleme, bir kümeleme algoritmasının hangi sayıda kümelerin var olduğunu belirlemesi için kullanılan bir tekniktir. Karma dağılımların zorluğu, veri kümesindeki gizli kalıpların ve yapıların keşfedilmesinde yatmaktadır.

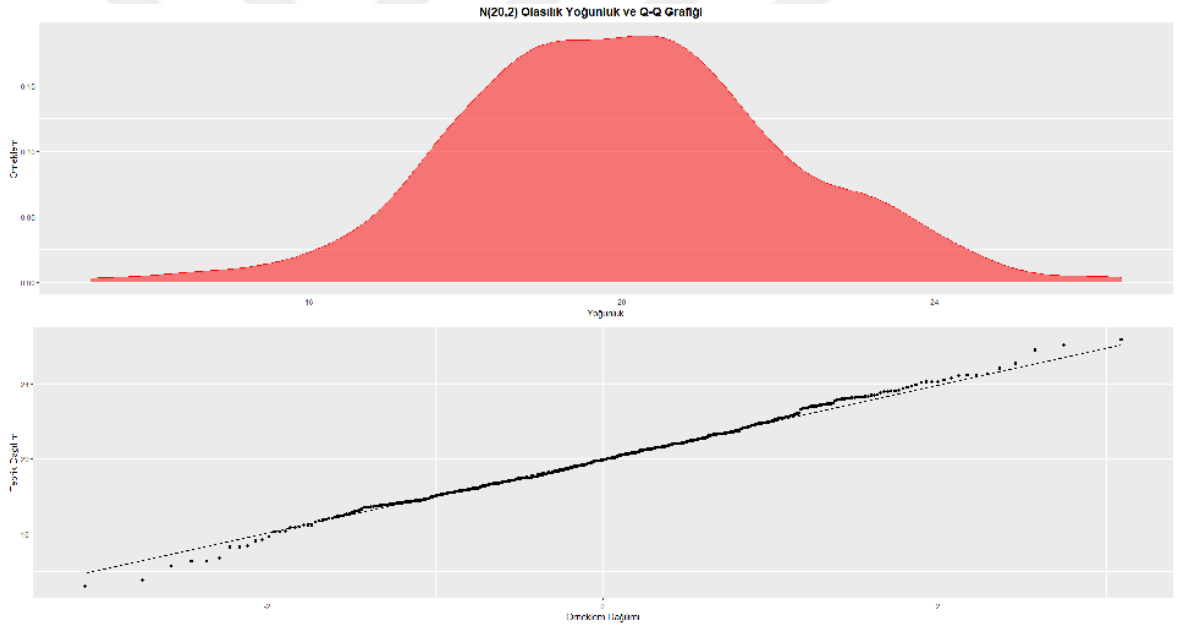
Karma dağılımlarda bileşen sayısı k 'nın belirlenmesi için en temel yaklaşımlardan biri grafiksel yöntemlerdir. Grafiksel yöntemlerde Q-Q ve OYF'den faydalanılabilir.

OYF, kümelenme yapısının belirlenmesi için tek boyutlu veri setlerinde ilk adım olarak kullanılabilir. Veri setindeki tek tepeli bir dağılım, herhangi bir kümeleme göstermeyen bir homojen yapının temsiliyken, birden fazla modun olduğu bir OYF heterojen bir veri setini tanımlayabilir. Birden fazla tepenin olduğu durumlar heterojenlik için çok önemli bir kanıt niteliğindedir (Servi, 2009). Bileşen sayısı tahmini için OYF'ye bakılarak veri setinin karma dağılıma sahip olup olmadığı, varsa başlangıç için bir bileşen sayısı belirlenmesi yapılabilir.

Q-Q grafiği, veri setinin nicelik çizimidir, seçilen bir dağılıma karşılık veri setinin dağılım niceliğini durumları eşit büyüklükte gruplara bölerek inceler. Bir Q-Q grafiği, gözlem değerlerinin sıralanması ve teorik bir dağılımdan elde edilen beklenen değerlerin karşılaştırılmasını gösterir. Q-Q grafikleri, karma dağılımlı veri setlerinde bileşen sayısının belirlenmesinde kullanılan bir yöntemdir. Bu yöntemde, gözlem ve teorik dağılımlar arasındaki farklılıkların görsel olarak incelenmesiyle bileşen sayısı belirlenir. Bir veri setinin karma normal dağılımdan oluşup oluşmadığını tespit etmek için Q-Q grafikleri yardımcı olabilir. Eğer veri seti tek bir dağılım

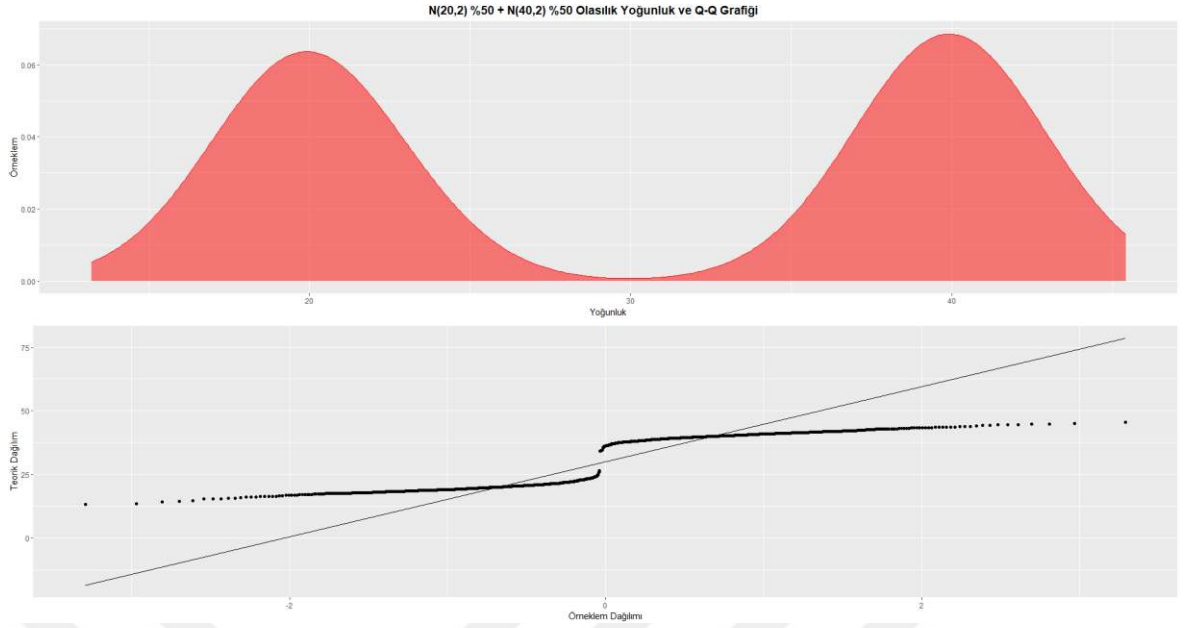
modeliyle açıklanabiliyorsa, Q-Q grafiği düz bir çizgi olarak gözükecektir. Ancak birden fazla dağılım modelinin varlığı durumunda, Q-Q grafiği düzensiz bir yapıya sahip olacaktır. Eğer seçilen değişken test dağılımına uyarsa, noktalar Q-Q grafiği üzerindeki beklenen normal değişkenlerin oluşturduğu çizgi etrafında kümelenir (Pallant, 2005). Bileşen sayısı belirlemek için, Q-Q grafiği incelenir ve düz bir çizgi oluşup oluşmadığına bakılır. Eğer düz bir çizgi oluşuyorsa, tek bir dağılım modelinin yeterli olduğu yorumu yapılabilir. Ancak düzensiz bir yapı varsa, birden fazla dağılım modelinin kullanılması gerektiği düşünülür ve bileşen sayısı bu yapıya göre belirlenir. Q-Q grafiğinde oluşan salınımlar veri setinin birden fazla bileşen içermesine kanıt olarak gösterilebilir. Bu yöntem, görsel bir yöntem olmakla birlikte, bileşen sayısının belirlenmesinde oldukça etkili ancak her zaman yeterli değildir.

Karma dağılım bileşen sayısı tahmininde OYF kullanıldığında araştırılması gereken durum, birden fazla mod değerinin görülüp görülmediğidir. Bu bölümün devamında farklı karma dağılımlar için OYF ve Q-Q grafikleri incelenmiştir.



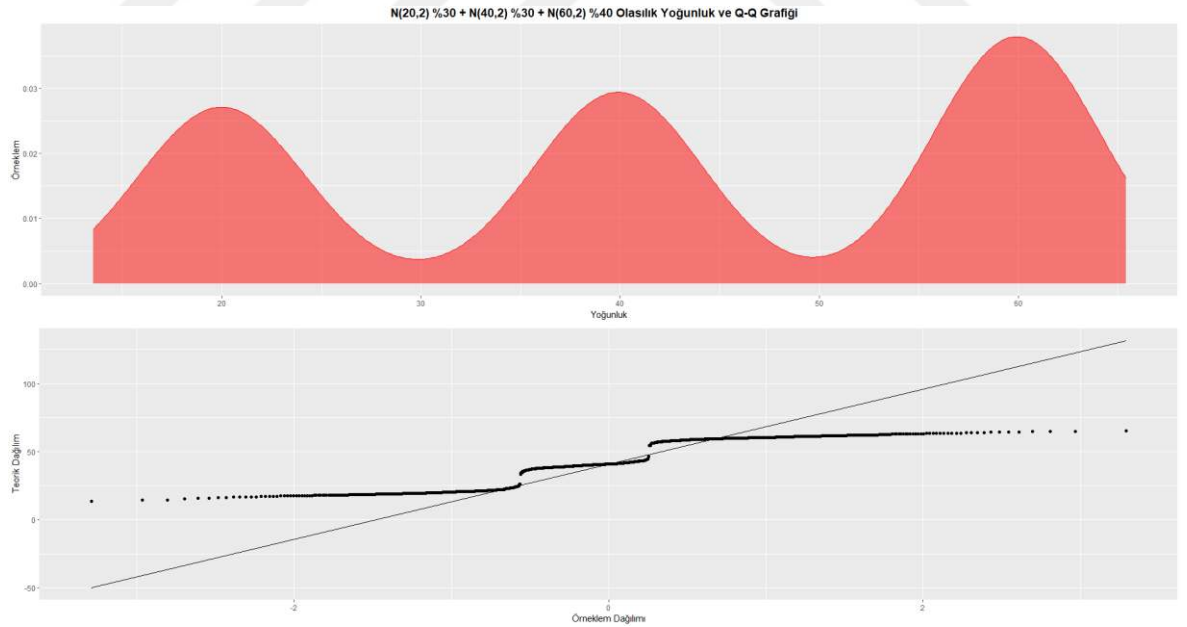
Şekil 3.20. $N(20, 2)$ OYF ve Q-Q Grafiği

Şekil 3.20., $N(20, 2)$ dağılımının OYF ve Q-Q grafiği çizilmiştir. OYF incelendiğinde, normal dağılım şeklinde bir davranış gösterdiği görülür. Benzer şekilde Q-Q grafiğinde de herhangi bir salınım yoktur. Bu grafikler yardımıyla, incelenen dağılımın herhangi bir karma dağılım bileşeni içermediği yorumu yapılabilir.



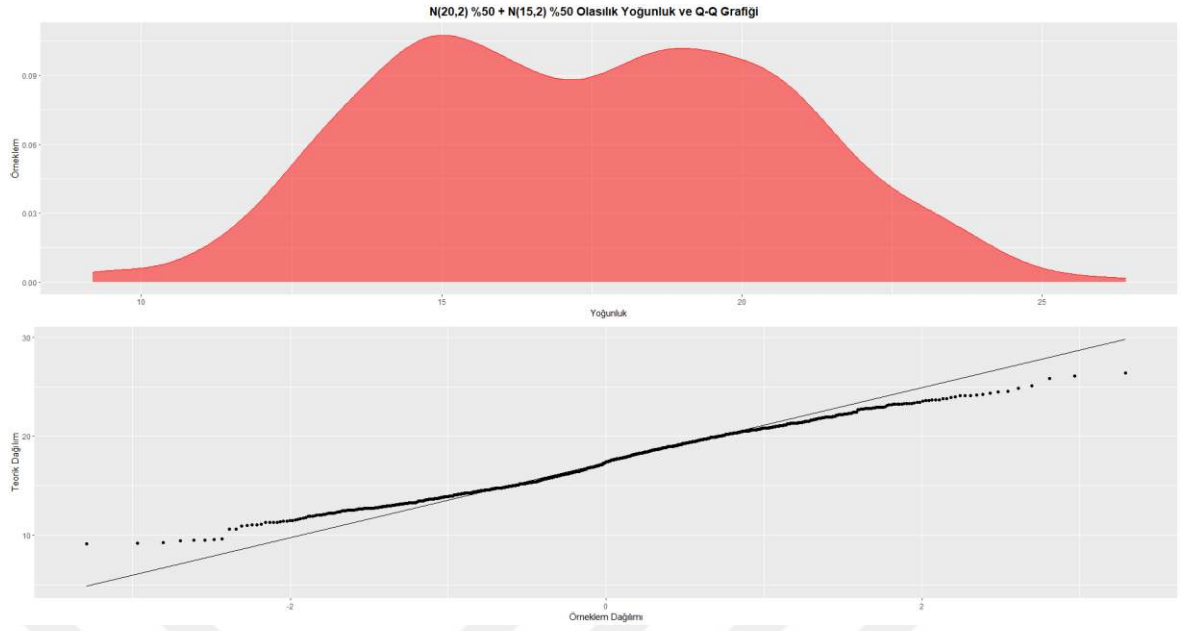
Şekil 3.21. $N(20, 2)$ %50 + $N(40, 2)$ %50 OYF ve Q-Q Grafiği

Şekil 3.21., %50'si $N(20, 2)$, kalan %50'sinin ise $N(40, 2)$ olan karma dağılımın grafiğidir. Ortalama değerlerin birbirine yeterli uzaklıkta olmasından dolayı iki tepe grafikte rahatlıkla görülebilmektedir. Benzer şekilde Q-Q grafiğindeki salınım, karma dağılım bileşeninin iki olduğunu destekler. Bu grafiklerden hareketle karma dağılım varlığı düşünülebilir.



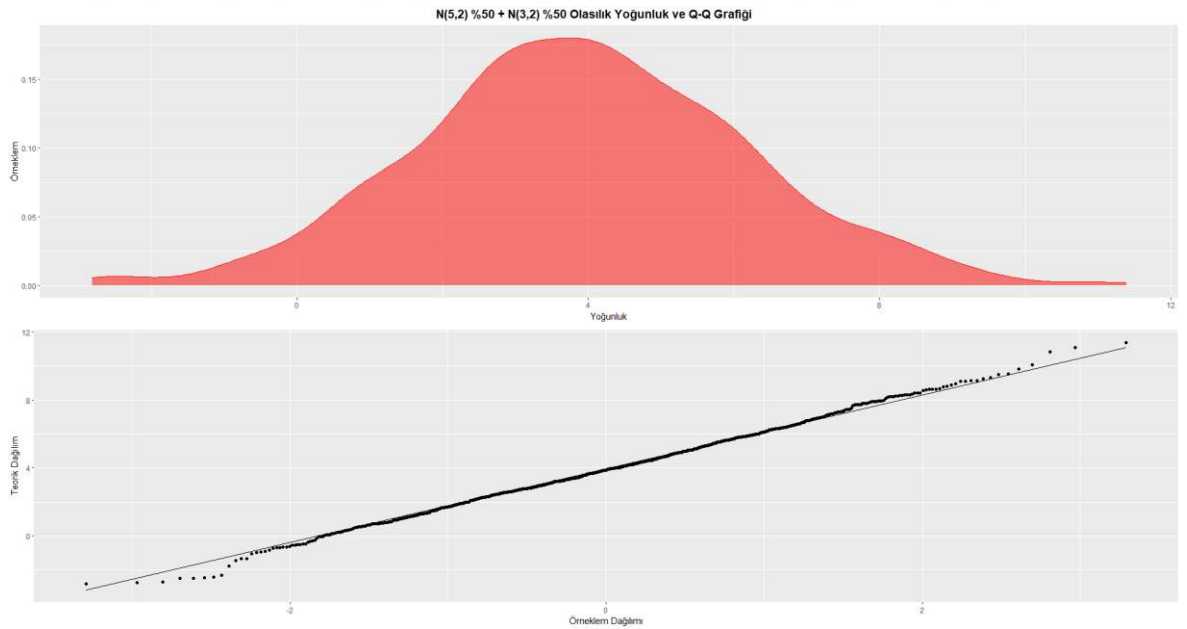
Şekil 3.22. $N(20, 2)$ %30 + $N(40, 2)$ %30 + $N(60, 2)$ %40 OYF ve Q-Q Grafiği

Şekil 3.22.'de, üç bileşenli bir karma dağılım ele alınmıştır. Şekil 3.21.'ye benzer olarak, ortalama değerler birbirinden yeterli kadar uzaktır ve üç tepe noktası net bir şekilde izlenir ve Q-Q grafiğinde üç bileşenli bir salınım görülmektedir.



Şekil 3.23. N(15, 2) %50 + N(20, 2) %50 OYF ve Q-Q Grafiği

Şekil 3.23.'da %50'si N(15, 2) dağılımından, %50'sinin ise N(20, 2) dağılımından oluştuğu bir karma dağılım ele alınmıştır. Bu karma dağılımda ortalama değerler birbirine yakındır ve OYF'de görüldüğü gibi tepe değerleri birbirine yaklaşmıştır. Q-Q grafiğinde ise, salınımlar net bir şekilde görülememeye başlamıştır.

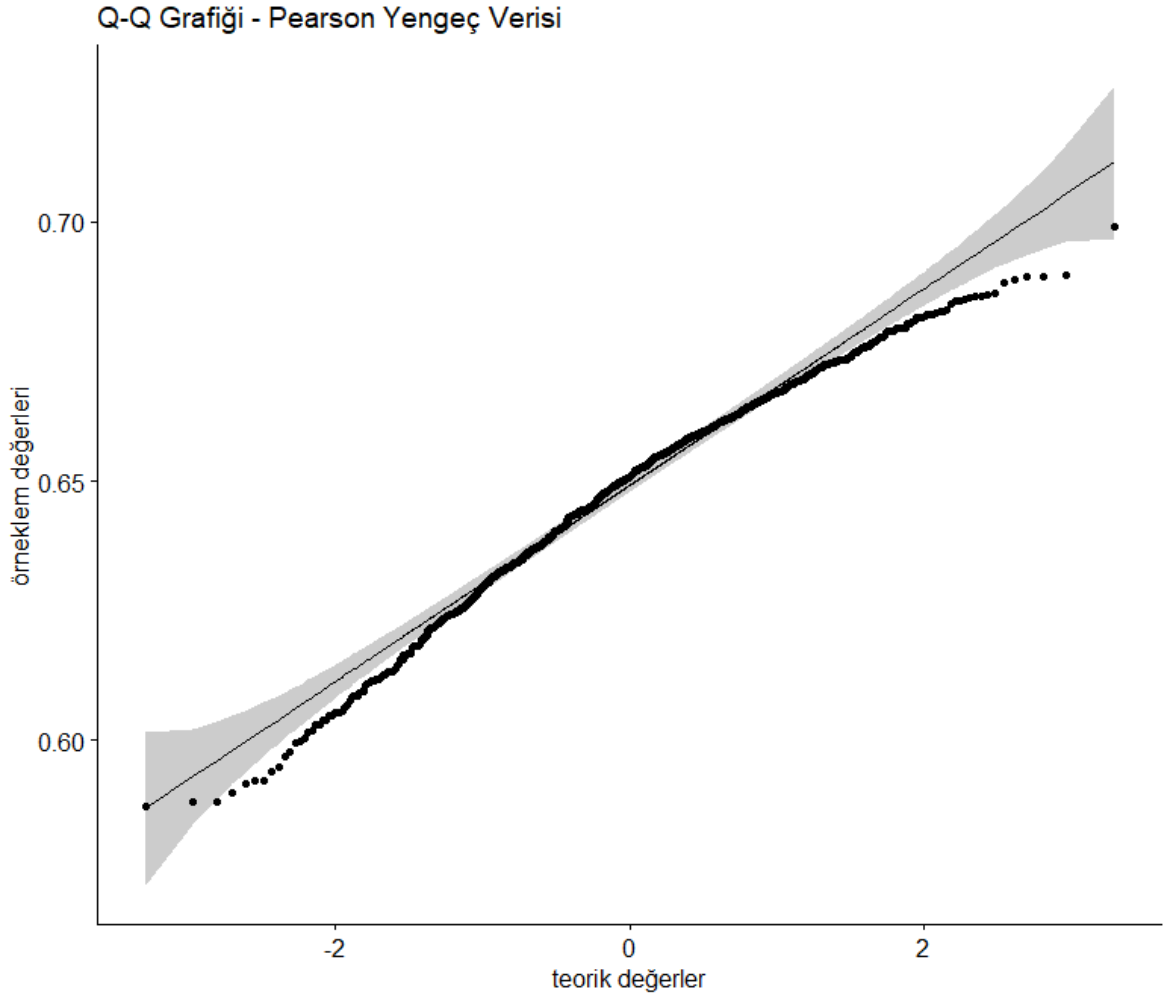


Şekil 3.24. N(3, 2) %50 + N(5, 2) %50 OYF ve Q-Q Grafiği

N(3, 2) ve N(5, 2) dağılımların eşit oranda oluşturduğu karma dağılımın ilgili grafikleri

Şekil 3.24.'te verilmiştir. Bu karma dağılımda her iki dağılımın değerleri iç içe geçmiş ve tek tepeli bir OYF görünümü oluşmuştur. Q-Q grafiğinde bileşen sayısını tespit edecek salınımlar görülmemektedir.

Pearson ve Weldon'un yengeç verisine geri dönecek olursak, bu verinin Q-Q grafiği Şekil 3.25.'de çizilmiştir. Q-Q grafiği üzerindeki dalgalanma, gözlemlerin normal dağılıma uymadığını göstermektedir.



Şekil 3.25. Pearson Yengeç Verisi Q-Q Grafiği

Örneklerden görüldüğü üzere, grafiksel yöntemler bileşen sayısı tahmininde, bileşenlerin birbirine uzak olduğu karma dağılımlarda oldukça iyi çalışırken, bileşenlerin ortalama değerlerinin birbirine yakın olduğu karma dağılımda çok etkili bir yöntem olmadığı görülmektedir. Bu nedenle, karma dağılım bileşen sayısı k 'yı belirleyebilmek için farklı yöntemlere ihtiyaç duyulur. Ancak yine de grafiksel yöntemlerin karma dağılım bileşen sayısı belirlemede tehditleri bilinmesi açısından ilk adım olarak kullanılması metotsal olarak bir problem yaratmamaktadır.

Model uygunluğunu ölçmek için kullanılan diğer yöntemler arasında olabilirlik oran test (LTR) istatistiği, Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) bulunur.

Olabilirlik Oran Test (LTR) İstatistiği

Karma dağılım modellerinde bileşen sayısı k 'nın tahmin edilmesi için kullanılan yöntemlerden biri, olabilirlik oran test (LTR) istatistiğidir. Bu istatistiksel test, farklı karma dağılım modelleri arasındaki uyumun derecesini ölçmektedir. Bu test, daha basit bir modele karşı daha karmaşık bir modele dayalı olarak, veri kümesine hangi karma dağılım modelinin daha iyi uyduğunu belirlemek için kullanılır. Böylece oluşturulan iki karma dağılım modelinin olabilirlik değerleri arasındaki farkın istatistiksel olarak anlamlı olup olmadığını belirler. LTR istatistiği, genellikle ki-kare (χ^2) dağılımına yakın bir dağılıma sahiptir (Johnson ve Wichern, 1982).

LTR istatistiği, birbirine bağlı olan hiyerarşik karma dağılım modelleri arasında karşılaştırma yapılırken geçerlidir. Karmaşık model, daha basit modele bir ya da daha fazla parametre eklenerek elde edilir. Karma dağılım modeline yeni parametreler eklendiğinde, olabilirlik değeri artar. Bu nedenle, LTR istatistiği, oluşturulan tüm modeller arasından bir seçim yapmak için kullanılabilir.

Bu test, iki modelin olabilirlik sonuçlarının karşılaştırılmasıyla elde edilir. Bu sayede, hangi modelin veri kümesine daha iyi uyduğunu belirlemek için bir kriter olarak kullanılabilir. LTR, karma dağılım modellerinde bileşen sayısı k 'nın tahmin edilmesi gibi birçok alanda kullanılabilir ve veri analizi için önemli bir araçtır. L_1 basit, L_2 karmaşık modele bir olabilirlik fonksiyonu olmak üzere, $LTR = 2(\ln L_1 - \ln L_2)$ olarak hesaplanır.

LTR istatistiğinin serbestlik derecesi, karmaşık modele eklenen parametre sayısına eşittir. Bu sayede, herhangi bir karma dağılım modelindeki ek parametrelerin etkisi test edilebilir. Test sonuçlarının yorumlanması için istatistiksel tablolar kullanılır ve test istatistiğinin kritik değeri belirlenir. LTR istatistiği, birçok alanda kullanılan güçlü bir istatistiksel araçtır ve veri analizinde önemli bir rol oynar. Test sonuçları, farklı karma dağılım modelleri arasındaki en iyi uyumu belirlemek ve gözlemlerin daha iyi anlaşılmasına yardımcı olmak için kullanılır.

Bilgi Kriterleri

Karma dağılım modellerinde, uygun modelin seçimi ve bileşen sayısının belirlenmesi, gözlenen veriye en iyi uygunluk sağlayacak modelin seçilmesiyle sağlanır. Bu amaçla, bilgi kriterleri karma dağılım modellerinde bileşen sayısı belirlemek için kullanılan bir diğer yöntemdir.

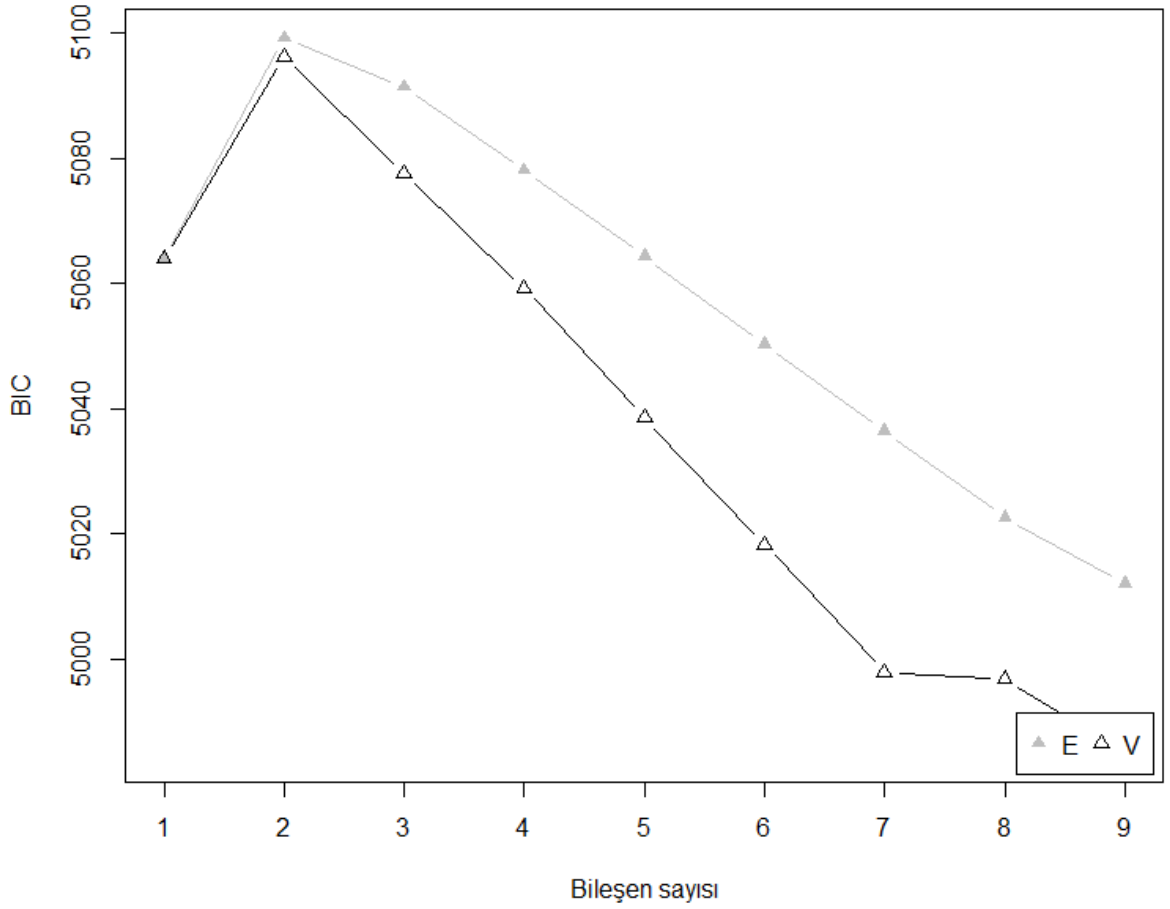
Bilgi kriterleri, farklı karma dağılım modellerinin uygunluğunu değerlendirmek için kullanılan istatistiksel yöntemlerdir. Bu kriterler arasında, Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC), model doğruluğunu ve karmaşıklığı dengeleyerek optimal k 'yi seçmek için yaygın olarak kullanılmaktadır (Garg ve ark., 2009), (Titterington ve ark., 1985), (Sin ve White, 1996). Bu kriterler, modelin uygunluğunu hesaplamak için kullanılan matematiksel formüllere dayanır ve her iki kriterde dedaha düşük bir değer, daha iyi bir uygunluğu gösterir. Bilgi kriterleri yöntemi, matematiksel hesaplamalara dayanması nedeniyle daha objektif bir yöntemdir. Karma dağılımlarındaki bileşen sayısını belirlemek için olabilirlik oran test istatistiği, AIC ve BIC gibi

model tabanlı yöntemler entegre bir şekilde kullanılabilir. Bu bilgi kriterleri veri setine uyacak en iyi karma modeli seçebilmek için kullanılır (Çalış, 2005).

Akaike Bilgi Kriteri (AIC), bir veri kümesi için kullanılan istatistiksel modellerin kalitesini göreceli olarak ölçen bir yöntemdir. Bu kriter, Japon istatistikçi Akaike Hirotugu tarafından geliştirilmiştir. İlk kez 1971 yılında bir sempozyumda duyurulan AIC, daha sonra 1973 yılında yayınlanmıştır (Akaike, 1974). AIC, bir veri kümesi için birden fazla model varsa, her modelin kalitesini diğer modellerin kalitesiyle karşılaştırarak tahmin eder. Bu nedenle, AIC model seçimi için bir yol sağlar. Bilgi teorisi temelli olan AIC, model verisi ile ilgili bilgileri temsil eder ve modelin uygunluğu ile karmaşıklığı arasındaki ilişkiyi anlamak için kullanılır. Ancak AIC, bir modelin mutlak kalitesini değil, diğer modellere göre olan kalitesini ölçer. Ele alınan istatistiksel bir modelin AIC'si, modelin olabilirlik fonksiyonunun maksimize değerinin, modelde tahmin edilen parametre sayısının bir fonksiyonu olarak hesaplanır. AIC, k bileşenli bir modelin, $L(\psi)$ olabilirlik fonksiyonu olarak tanımlanır ve $2k - 2\ln L(\psi)$ olarak hesaplanır (Akaike, 1974). Daha düşük bir AIC değeri, daha iyi bir model uygunluğunu ve daha düşük bir model karmaşıklığı anlamına gelir.

Bayesci bilgi kriteri (BIC) ise, n gözlem sayısı olmak üzere, $BIC = -\ln(n) k - 2\ln L(\psi)$ formülüyle hesaplanır (Schwarz, 1978). AIC kriterine benzer olarak, modeller arasında en küçük değerin en iyi uyuma sahip olduğu yorumu yapılır (Yalta, 2011). Özellikle büyük örneklemelerde en doğru modeli seçebilmesi sebebiyle yüksek bir geçerliliğe sahiptir (Schwarz, 1978).

Sonuç olarak; AIC, veri kümesine en uygun karma dağılım modelini seçmek için kullanılan bir kriterdir. Bu kriter, modelin karmaşıklığına ve uygunluğuna bağlı olarak bir ceza puanı uygular. BIC ise, olasılık dağılımının doğruluğunu ve basitliğini birleştirir. BIC, AIC'den daha güçlü bir ceza kullanılır ve bu, n olarak belirtilen örnek boyutuna bağlıdır (Schwarz, 1978). Bu iki kriter arasındaki farklılıklar ve avantajları, seçilen modellerin performansına bağlıdır. AIC, genellikle büyük veri setleri için daha iyi bir performans gösterirken, BIC küçük veri setleri için daha iyi sonuçlar verir. Her iki yöntemde de en küçük değer en uygun model olarak temsil edilir. BIC ayrıca, modeldeki parametre sayısını azaltarak aşırı uyuma (overfitting) karşı bir çözüm sunar. Aşırı uyum, modelin veri setine çok iyi uyması ama genelleme yapamaması durumudur. Bu durumda modelin öğrenme yeteneği yüksek olmasına rağmen, yeni veri için tahmin yaparken performansı düşük olabilir. BIC, daha basit bir modelin seçilmesini sağlayarak aşırı uyuma karşı koruma sağlar.



Şekil 3.26. Pearson Yengeç Verisi BIC Grafı

Şekil 3.26.'da Weldon'un topladığı gözlemlerin en iyi karma dağılım modelini ("E" eşit varyans veya "V" değişken varyans) ve optimum bileşen sayısını bulmak için BIC analizi yapılmıştır. Pearson bu çalışmayı yaptığı zaman AIC ve BIC yaklaşımı henüz ortaya koyulmamıştı. Pearson yaptığı çalışmada, tüm hesaplamaları elle yapmış ve nihayetinde iki bileşenli normal karma dağılım yaklaşımını benimsemişti. Yapılan BIC analizi sonucunda, en iyi karma dağılım modelinin iki bileşenli olduğu doğrulanmıştır. Böylece Pearson'ın bulduğu sonuçların doğruluğu ile BIC yönteminin geçerliliği sınanmıştır.

3.6.2. Karma Dağılımlarda Parametre Tahmini

Karma dağılım modellerinde bileşen sayısı belirlendikten sonra, her bir bileşene ait parametrelerin tahmini ikinci adımdır. Bu parametreler, her bileşenin karma oranlarını ve OYF parametrelerinin tahminlerini içerir. Bu bölümde karma dağılım parametre tahmini için kullanılan bu yöntemler ele alınmıştır.

En Çok Olabilirlik Yöntemi (Maximum Likelihood - ML)

Parametrelerin tahmininde temel olarak maksimum olabilirlik (ML) yöntemi kullanılır. Bu yöntemle ilgili ayrıntılı bilgi 3.4. Parametre Tahmini başlığında verilmiştir. Ancak, karma dağılım modelinin olabilirlik fonksiyonu karmaşık olduğundan, bu tahminlerin hesaplanması zor olabilir. Bu nedenle ML algoritması temelli daha gelişmiş algoritmaların kullanılması, parametre tahminin daha hızlı ve verimli hale getirir.

Beklenti Maksimizasyonu Yöntemi (Expectation Maximization - EM)

Beklenti maksimizasyon yöntemi (Expectation Maximization - EM) algoritması, Expectation (Beklenen değeri hesaplama adımı) ve Maximization (Maksimum yapma adımı) kelimelerinin baş harflerinden oluşturalara isimlendirilmiş bir algoritmadır. Bu algoritma ilk olarak 1977 yılında Dempster ve arkadaşları tarafından önerilmiştir (Dempster ve ark., 1977). Son 20 yılda, özellikle sonlu karma dağılım modellerinde parametre tahmini için sıkça kullanılmaktadır (Lakshmi ve Vaidyanathan, 2016). EM algoritması, genellikle verinin belirli bir olasılık dağılımının karışımından geldiği varsayımını benimser ve model yapısı ve parametreleri belirlemek için kullanılır. EM algoritmasının temel yöntemi, maksimum olabilirlik tahmin yöntemidir. EM algoritması, kapalı formda bir çözüm mevcut olmadığında uygun bir yöntem sağlar ve olasılık denklemi için bir çözüm elde eder. Bu nedenle, EM algoritması, karma dağılım parametrelerin tahmin edilmesi amacıyla kullanılır (Wengrzik, 2012). EM algoritması, karma dağılım modellerinde parametre tahmini için en uygun yöntemlerden biri olarak nitelendirilebilir. Bu sayede, karmaşık dağılım modellerindeki parametreleri tahmin etmek için etkili bir yöntem sunmaktadır.

Temelde, EM algoritması önerilen ve olasılığı maksimize eden veri setinin parametrelerini bulmaya çalışır. n gözlemin oluşturduğu k bileşenli bir karma dağılım modelinin tüm değerlerinde $L(\psi)$ 'yi maksimize eden $\theta(x_1, \dots, x_n)$ değerine θ 'nin en çok olabilirlik tahmini olarak adlandırılır.

E-adımı olarak dağılımlar için merkez çap değerleri gibi bir tahmini parametre vektörü ile başlanır. EM algoritmasının E-adımı olarak herhangi bir x değeri için, bu kümelere ait olma olasılığının hesaplanması gerekmektedir.

M-adımı olarak hesaplanan p_i değerleri göre parametreler p_{i+1} yeni adım değeri ile tekrardan hesaplanır ve işlemler iteratif olarak tekrar edilir. Algoritma olabilirlik fonksiyonunu maksimize etmeye çalışır.

EM algoritmasının birkaç önemli özelliği şöyle sıralanabilir:

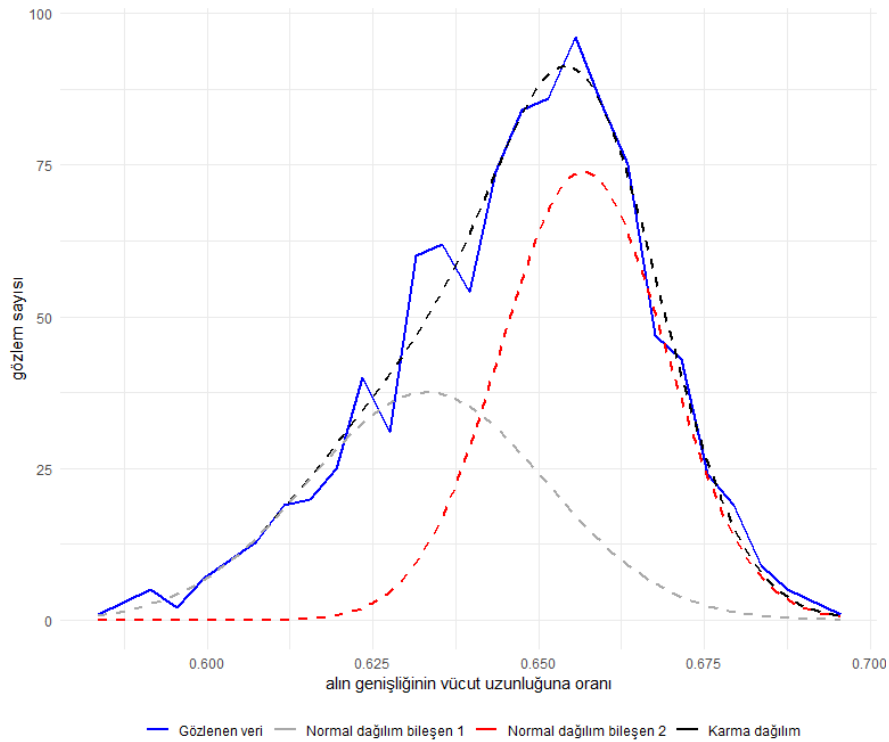
- Her EM iterasyonunda, olabilirlik fonksiyonu artar ve bu sayede EM algoritması, maksimum olabilirlik tahminlerine ulaşır.
- Genellikle düzgün koşullar altında, EM algoritması güvenilir bir şekilde yakınsar ve bu sayede model parametrelerinin doğru bir şekilde tahmin edilmesini sağlar.

- EM algoritması, hesaplamaları anlaşılır ve analitik olarak kolayca gerçekleştirilebilir. Adımları oldukça basittir ve uygulaması kolaydır.
- EM algoritması, kayıp veri tahminlerini de hesaplayabilir. Bu özellik, özellikle eksik veri durumlarında çok yararlıdır ve model parametrelerinin doğru bir şekilde tahmin edilmesine yardımcı olur.

EM algoritmasının bazı dezavantajları ise şöyle sıralanabilir:

- EM algoritması, parametre tahminlerinin kovaryans matrisini otomatik olarak hesaplayamaz. Bu nedenle, bazı durumlarda kovaryans matrisinin ayrıca tahmin edilmesi gerekebilir.
- Yakınsama hızı bazı durumlarda oldukça yavaş olabilir. Özellikle veri setleri büyük olduğunda veya modelin karmaşıklığı arttığında bu durum gözlemlenebilir.
- Bazı problemlerde E ve M adımları analitik olarak ifade edilemez ve bu nedenle sayısal çözümler kullanılması gerekebilir.
- EM algoritmasının sonuç tahminleri, parametrelerin başlangıç değerlerine bağlıdır. Bu nedenle, farklı başlangıç noktaları kullanarak aynı veri seti için farklı sonuçlar elde edilebilir.

Pearson ve Weldon verisine geri dönülecek olursa, Pearson tarafından elle yapılan hesaplama sonuçlarına dayanan parametre tahminlerini grafik olarak gösterilebilir. Şekil 3.27. Pearson yengeç verisi hesaplanmış parametrelerin OYF'sidir.



Şekil 3.27. Pearson Yengeç Verisi ve Karma Dağılım OYF'leri

Şekil 3.27.'de, mavi ile çizilen gerçek gözlenen değerlerin OYF'sidir. Weldon tarafından Napoli'de örneklenen 1000 yengeç için alın genişliğinin vücut uzunluğuna oranını içermektedir. Pearson bu veri setine uyması için iki bileşenli bir normal karma modeli kullanmıştır. Bu karma normal dağılımın karma oranları $\pi_1 = 0,441$, $\pi_2 = 0,559$; $\mu_1 = 0,634$, $\mu_2 = 0,657$; $\sigma_1^2 = 0,000339$, $\sigma_2^2 = 0,000158$ şeklinde bulunmuştur. Böylece, çarpık gözlemlerin karma dağılım modeli kullanılarak modellenmesi fikri ilk olarak Weldon'un verisini kullanan ve açtığı yolda ilerleyen Pearson tarafından önerilmiştir (Pearson, 1894).

3.6.3. Performans Kriteri

Bir verinin istatistiksel olarak parametre tahmini yapıldıktan sonra bu modelin geçerliliğinin test edilmesi önemli bir adımdır. Analiz edilen veri seti bilinen herhangi bir dağılıma uymuyorsa; geleneksel olarak ampirik dağılım oluşturmak ya da bu çalışmanın önemli noktalarından biri olan karma dağılım yaklaşımı benimsenebilir. Oluşturulan ampirik ve karma dağılımlardan elde edilen tahmin değerlerinin gerçek değerler ile arasındaki benzerliğinin ölçülmesi için Ortalama Mutlak Yüzde Hatası (MAPE) VE Mann-Whitney U (MWU) testi kullanılabilir.

Ortalama Mutlak Yüzde Hatası (MAPE)

Ortalama Mutlak Yüzde Hatası (MAPE), tahmin edilen değerler ile gerçek değerler arasındaki farkın yüzdesel olarak ifade edildiği bir ölçüttür. Bu ölçüt genellikle öngörülebilirlik performansının ölçülmesi için kullanılır yüzde (%) cinsinden ifade edilir.

MAPE, gerçek (y_i) ve tahmin edilen değerler ($\hat{y}_i$) arasındaki yüzde hatayı temsil eder ve $100/n \sum_{i=1}^n |y_i - \hat{y}_i|/y_i$ formülüyle hesaplanır. MAPE hesaplanırken, negatif değerlerin mutlak değerleri kullanılır. Bu nedenle, MAPE değeri herhangi bir negatif değerde olamaz (Makridakis ve ark., 1998).

MAPE değeri, %0 ile %100 arasında değerler alır. Düşük bir MAPE değeri, gerçek ve tahmin edilen değerler arasındaki benzerliğin daha yüksek olduğunu gösterir. %10'nun altında hesaplanan MAPE yüksek doğruluklu tahmin, %10-20 arası iyi bir tahmin, %20-50 kabul edilebilir, %50'den büyük olduğu durumlarda iyi kabul edilemez bir tahmin olarak tanımlanır.

Avantajları:

- Yüzdesel olarak ifade edildiği için, farklı büyüklükteki veri arasında karşılaştırma yapmak mümkündür.
- MAPE, bir öngörünün doğruluğunu ölçmek için basit ve kullanışlı bir ölçüttür.
- MAPE, genellikle işletme, finansal analiz ve pazarlama araştırması gibi alanlarda yaygın olarak kullanılır.

Dezavantajları:

- MAPE, bazı durumlarda yüksek hatalara neden olabilir. Özellikle, gözlemlerin yüzdesel değişkenlik gösterdiği durumlarda MAPE, yüksek hata oranlarına neden olabilir.
- Sıfıra yakın veya sıfıra eşit değerlerde MAPE hesaplanması zor olabilir.
- MAPE, öngörülen değerlerin doğruluğunu ölçmek için tek bir ölçüt olarak kullanılmamalı, farklı performans kriterleriyle birlikte kullanılmalıdır.

Bu çalışmada, performans kriteri olarak, MAPE değeriyle birlikte istatistiksel bir kıyaslama testi olan Mann-Whitney U testi de kullanılmıştır.

Mann-Whitney U testi

Mann-Whitney U (MWU) testi, parametrik olmayan bir hipotez testidir. Bu test, iki bağımsız örneklemin ortalamalarının farklılığına karar vermek için kullanılır. Bu test normal dağılım varsayımını gerektirmez. Bunun yerine, sadece örneklemelerin sıralı değerlerine dayanarak çalışır. Bu nedenle, normal dağılımı olmayan veri için idealdir. MWU testinin hesaplanması şu adımlarla yapılır:

- İki bağımsız örneklem grubu için veri toplanır ve birleştirilir.
- Birleştirilmiş veri setindeki tüm değerler sıralanır.
- İki örneklem grubunun sıralanmış verisi ayrı ayrı incelenir ve her bir gözlem için sırasıyla " U_1 " ve " U_2 " değerleri hesaplanır. Bu değerler, o gözlemin kaçınıcı sırada olduğuna ve hangi örneklem grubuna ait olduğuna bağlıdır.
- U_1 ve U_2 değerleri toplamı hesaplanır: $U = U_1 + U_2$.
- N_1 ve N_2 , sırasıyla her bir örneklem grubunun örneklem büyüklükleridir. Eğer her iki örneklem grubu da eşit büyüklükte ise, $N_1 = N_2 = N/2$ formülü kullanılabilir. Burada N , toplam örneklem büyüklüğüdür.
- U_0 değeri, birleştirilmiş veri setindeki medyanın kaçınıcı sırada olduğunu belirler. $U_0 = N_1 * N_2 / 2$ formülü kullanılarak hesaplanır.
- Test istatistiği $U = \min(U, N_1 * N_2 - U)$ formülüyle hesaplanır.
- U_0 değeri, test istatistiğinin ortasında mı yoksa dışında mı olduğuna bakarak p -değeri hesaplanır. Eğer test istatistiği U_0 'ın solunda kalıyorsa, p -değeri hesaplanırken U_0 'ın sağındaki alan kullanılır; eğer test istatistiği U_0 'ın sağında kalıyorsa, p -değeri hesaplanırken U_0 'ın solundaki alan kullanılır.

Test sonucunda elde edilen p -değeri, iki grup arasındaki farklılık skorlarının dağılımı altında bir ihtimal hesaplamasını ifade eder. Bu ihtimal, H_0 hipotezinin doğru olması durumunda, iki grup arasındaki farkın bu kadar büyük veya daha büyük olması olasılığıdır. MWU testi, birinci örneklemin ikinci örnekleme göre daha büyük olduğu veya daha küçük olduğu durumlar için bir alternatif hipotez oluşturur. Bu hipotezler iki yönlü olarak ya "iki grup birbirinden farklıdır" ya da "iki grup eşittir" şeklinde ifade edilir. H_0 hipotezi altında, iki popülasyonun dağılımları aynıdır. MWU testinin p -değeri anlamlılık düzeyinden küçük veya buna eşitse, H_0 hipotezi reddedilir. MWU testinin uygulanması, istatistiksel yazılımların çoğunda mevcuttur. İstatistiksel yazılımlar, test sonuçlarını sağlar ve çoğu zaman testin gücünü ve etkinliğini değerlendirmek için birçok istatistiksel ölçüt de sağlar (Mann ve Whitney, 1947). Bu testin avantajları ve dezavantajları şu şekildedir:

Avantajları:

- Parametrik olmayan bir test olduğu için normal dağılım varsayımına ihtiyaç duyulmaz. Bu nedenle veri normal dağılmadığında oldukça kullanışlıdır.
- Küçük örneklem boyutları ve büyük varyanslı veride de kullanılabilir.
- İki grup arasındaki medyan farkını karşılaştırmak için kullanışlıdır.

Dezavantajları:

- Test, verinin sadece medyanlarını karşılaştırır ve diğer özellikleri göz ardı eder. Bu nedenle, tamamını değerlendirmek için farklı bir test gerekir.
- Sadece iki grup arasındaki farkı belirleyebilir, daha fazla grup için uygun değildir.

Performans kriterleriyle uyumu ispatlamış girdi parametreleri belirlendikten sonra simülasyon modelinin kurulması ve çalıştırılması gerekir.

3.7. Yoğun Bakım Ünitesi Modelinin Oluşturulması

Bu bölümde YBÜ sistemi bir kuyruk modeli olarak tanımlanmış ve bu model kullanılarak teorik olarak doğrulanmış bir simülasyon modeli oluşturulmuştur. Kuyruk Teorisi ayrıntılı bir şekilde açıklanmış ve ardından kurulan simülasyon modelinin tüm aşamaları adım adım ayrıntılı bir şekilde işlenmiştir.

3.7.1. Kuyruk Teorisi

Kuyruk teorisi, olasılık teorisi ve istatistik bilim dallarını birleştirerek, kuyruğu oluşturan belirli bir olayın gelişimini, yoğunluğunu, dağılımını, ortalama bekleme süresini ve hizmet kapasitesini matematiksel olarak analiz eden bir disiplindir (Sundarapandian, 2009). Bir kuyruk, müşterilerin, ürünlerin veya hizmetlerin beklediği bir alan olarak tanımlanabilir. Kuyruk teorisi, özellikle hizmet sektörü, haberleşme, ulaşım, finans, tıp ve endüstri gibi yoğun kuyruk oluşan alanlarda kuyrukların davranışını analiz ederek, bir sistemin performansını ölçmek için birçok alanda uygulanır.

İlk olarak Danimarkalı bir mühendis olan Agner Krarup Erlang, telefon santraline gelen görüşmelerin sayısını Poisson süreci kullanarak modelledi (Erlang, 1909). Kuyruk teorisi, olasılık dağılımlarını kullanarak, bir sistemin performansını ölçmek için matematiksel modeller geliştirir. Bu modeller, kuyruğa giren varlıkların geliş oranları, kuyrukta bekleyenlerin sayısı, kuyrukta bekleyenlerin hizmet zamanları ve hizmet kapasitesi gibi parametrelerin belirlenmesinde kullanılır.

Kuyruk, sistemlerdeki gecikmeler, sıralamalar veya beklemler nedeniyle oluşabilir. Kuyruk teorisi, bu kuyrukların büyüklüğünü, uzunluğunu, beklemlerini, bekleme sürelerini gibi özelliklerini analiz ederek, bir sistemin performansını ölçer ve iyileştirmek için önerilerde bulunur (Kingman, 2009). Kuyruk teorisi, bir kuyruğun uzunluğunun ne kadar süreceğini tahmin ederek, bir sistemin verimliliğini ve kapasitesini arttırmak için tasarlanmıştır. Bu teori, kuyruğa giren varlıkların zamanlamasını ve işlem zamanlarını dikkate alarak, hizmet sunan bir sistemin kapasitesinin ne kadarının kullanıldığını belirler. Bekleme süresinin kısaltılması, hizmet kalitesinin iyileştirilmesi, hizmet kapasitesinin artırılması ve maliyetlerin azaltılması gibi hedefleri gerçekleştirmeye yardımcı olur. Kuyruk teorisi, aynı zamanda, bir sistemin kapasitesinin ne kadarının kullanıldığına dair anlamlı bir ölçü olarak da kullanılabilir. Örneğin, bir banka kuyruğundaki müşteri sayısını tahmin etmek, bir üretim tesisi kuyruğundaki ürün sayısını tahmin etmek, bir havaalanında kuyruktaki yolcu sayısını tahmin etmek, bir hastane kuyruğundaki hastaların bekleme sürelerini tahmin etmek, bir web sitesindeki kuyruktaki talepleri tahmin etmek gibi ölçüler kuyruk teorisi modelinin çıktıları olabilmektedir. Bu çıktıların analiziyle kuyruk teorisi, sistemlerin performansını analiz etmek ve iyileştirmek için kullanılan önemli bir araçtır. Bu nedenle, kuyruk teorisi Endüstri Mühendisliği ve İşletme Yönetimi gibi birçok alanda önemlidir.

Kuyruk teorisi, tek kuyruklu düğümleri genellikle $A/S/C$ şeklindeki Kendall'ın gösterimi ile tanımlar. A , sıraya girenler arasındaki süreyi, S işlemlerin boyutunu ve C düğümü sunucu sayısını belirtir (Kendall, 1953). Kendall gösterimine gerekirse bir terim daha eklenebilir ve bu sistemin kabul edeceği maksimum müşteri sayısını ifade eder. Kuyruk teorisindeki pek çok teorem, Markov zincirleri olarak adlandırılan matematiksel sistemlere indirgenerek ispatlanır. Bu yöntem, ilk olarak Andrey Markov tarafından 1906 yılında yayınlanan bir makalede açıklanmıştır (Markov, 1906). Kendall'ın gösteriminde, M Markov veya hafızasız anlamına gelir ve varışlar Poisson sürecine göre gerçekleşir. D deterministik anlamına gelir ve sıraya giren işlerin sabit bir servis

süresine sahip olduğunu ifade eder. k , kuyruğun hizmet veren sunucu sayısını belirtir ve bu sayı genellikle 1 veya daha fazla olabilir. Eğer sunucuların sayısı, düğümdeki iş sayısından az ise işler sıraya girecek ve hizmet için bekleneyecektir.

Kuyruk teorisi içinde farklı hizmet stratejileri ve davranışları uygulanabilir. İlk Giren İlk Çıkar ilkesi, müşterilere birer birer hizmet verildiğini ve en uzun süre bekleyen müşteriye öncelik verildiğini belirtir.

İlk Giren Son Çıkar ilkesi ise müşterilere birer birer hizmet verildiğini ve en kısa bekleme süresine sahip müşteriye önce servis edildiğini gösterir. İşlemci Paylaşımı, hizmet kapasitesinin müşteriler arasında eşit olarak paylaşıldığını ve kuyruk modelinin bu kurala göre ilerlediğini belirtir. Öncelik stratejisi, yüksek önceliği olan müşterilere ilk hizmet sunulması prensibine dayanır. Kısa İş İlk stratejisi, sunulacak bir sonraki işin en küçük boyuta sahip olanı seçilmesini ifade eder. Öncelikli En Kısa İş ilkesinde sunulacak bir sonraki iş, orijinal en küçük boyuta sahip olan seçilir (Harchol-Balter, 2012). En Kısa Kalan İşlem Süresi stratejisi, sunulacak bir sonraki işin kalan işleme gereksiniminin en kısa olanı olması gerektiğini ifade eder. Servis Tesisi ise müşterilerin hizmet alma şekillerini belirler. Tek sunucu stratejisi, müşterilerin sırayla tek sunucudan hizmet almasını, paralel sunucu stratejisi ise müşterilerin sırayla birçok sunucudan hizmet almasını sağlar. Tandem sırası ise birçok sayacın bulunduğu ve müşterilerin nereye sıraya gireceğine karar verebildiği bir stratejidir. Kuyruk teorisinde, bekleyen müşterilerin davranışları da farklı stratejilere dayandırılabilir. Kaçınma, müşterilerin çok uzun sıralar görürlerse sıraya katılmamayı tercih etmeleri anlamına gelirken, kandırma stratejisi, müşterilerin daha hızlı hizmet alacaklarını düşünerek sıralar arasında geçiş yapmalarını ifade eder.

Kuyruk teorisinde, Little Yasası (*Little's Law*) 1961 yılında John Little tarafından geliştirilmiş ve sistem performans değerinin arasında matematiksel ilişki kurulmasını sağlamıştır (Little, 1961). Eğer sistem durağan durum halinde, birim zamanda gelen müşteri sayısı (λ) ve servis süresi ($1/\mu$) sabitse; tüm sistemde beklenen müşteri sayısı L , kuyruktaki beklenen müşteri sayısı L_q ve müşterinin kuyruktaki beklenen bekleme süresi W_q , müşterinin sistemdeki beklenen bekleme süresi W olmak üzere, aşağıdaki denklemler yardımıyla hesaplanır.

$$L = \lambda W$$

$$L_q = \lambda W_q$$

$$W = W_q + \frac{1}{\mu}$$

M/M/1

$M/M/1$ kuyruğu, üstel dağılıma sahip hizmet gereksinimlerinin olduğu ve tek bir sunucunun gelen işlere hizmet verdiği basit bir kuyruk modelidir. Bu birimde geliş hızı $1/\lambda$ ve servis hızı $1/\mu$ ortalamalı üstel dağılıma uyar. En temel kuyruk modeli olan $M/M/1$ sistemine N sayıda müşteriye izin verildiği durumda gösterimi $M/M/1/N$ şeklinde güncellenir. Bu sistemin formülleri Çizelge 3.4.'te verilmiştir. Bir $M/M/1$ kuyruk sisteminin sunucu sayısı s sayıda olabilir ve $M/M/s$ kuyruk modeli şekilde gösterilir. $M/M/1$ sisteminde kullanılan tüm varsayımlar $M/M/s$ kuyruk sisteminde geçerlidir.

Çizelge 3.4. $M/M/1/N$ Kuyruk Sistemi için Formüller

Sembol	Açıklama	Formül
ρ	Sistemin Kullanım Oranı	$\frac{\lambda}{\mu}$
P_0	Sistemin Boş Kalması Olasılığı	$\frac{1 - \rho}{1 - \rho^{N+1}}$
P_n	Sistemde n Birim Olması Olasılığı	$\rho^n \frac{1 - \rho}{1 - \rho^{N+1}}$
P_N	Sistemin Dolu Olması Olasılığı	$\rho^N \frac{1 - \rho}{1 - \rho^{N+1}}$
P_r	Sistemde En Az k Birim Bulunması	$1 - \frac{1 - \rho^k}{1 - \rho^{N+1}}$
L_q	Kuyrukta Olması Beklenen Birim Sayısı	$\frac{\rho^2(1 + \rho^N(N - 1) - N\rho^{N-1})}{(1 - \rho^{N+1})(1 - \rho)}$
λ_{eff}	Bir Müşterinin Sistemde Harcadığı Toplam Ortalama Zaman	$\lambda[1 - P_N]$
L	Sistemde Olması Beklenen Birim Sayısı	$\frac{\rho}{1 - \rho} + (N + 1) \left[1 - \frac{1}{1 - \rho^{N+1}} \right]$
W_q	Kuyrukta Geçen Beklenen Süre	$\frac{\lambda(1 + \rho^N(N - 1) - N\rho^{N-1})}{\mu(\mu - \lambda)(1 - \rho^N)}$
W	Sistemde Geçen Beklenen Süre	$\frac{\mu(1 - \rho^N) - N\lambda\rho^{N-1}(1 - \rho)}{\mu(\mu - \lambda)(1 - \rho^N)}$

$M/G/1$

Servis hızı çoğu durumda üstel dağılım ile temsil edilemez. 1930 yılında servis süresi kısıtını ortadan kaldırmak için geliştirilen $M/G/1$, Markov gelişlerin hizmet sürelerinin genel dağılıma sahip olduğu ve tek bir sunucunun bulunduğu bir kuyruk modelidir (Pollaczek, 1930). Model, hizmet süreleri üstel dağılıma olan $M/M/1$ kuyruğunun bir uzantısıdır. $M/G/1$ modeli tarafından temsil edilen bir kuyruk, durum uzayı $\{0,1,2,3,\dots\}$ kümesi olan, değerin kuyruktaki

müşteri sayısına karşılık geldiği stokastik bir süreçtir. İki durum arasındaki geçişler, iki durum arasındaki geçiş sürelerinin parametresi λ olan üstel dağılıma sahiptir. Müşterilerin hizmet süresinin bitirilmesi ve ayrılması durumunda, i 'den $i - 1$ 'e yapılan geçişler, hizmet verilen bir müşteriyi temsil eder. Bir müşteriye hizmet verilmesi için gereken süre, genel dağılım fonksiyonuna sahip olmalıdır. Varışların arasındaki sürelerin ve hizmet sürelerinin uzunlukları, istatistiksel olarak bağımsız olduğu varsayılan rasgele değişkenlerdir.

$M/G/1$ modelinde ise, hizmet gereksinimleri rasgele bir olasılık dağılımı tarafından belirlenir ve bu model Pollaczek tarafından 1930'da çözülmüştür (Pollaczek, 1930). Aleksandr Khinchin tarafından, bu çözüm olasılık açısından tekrarlanabilir olarak adlandırılmış ve Pollaczek-Khinchine formülü olarak bilinir hale gelmiştir (Whittle, 2002). Pollaczek-Khinchine formülü, Laplace'ın bir $M/G/1$ kuyruğu için kuyruk uzunluğu ile hizmet süresi dağılımı arasındaki ilişkiyi belirtir. Terim aynı zamanda böyle bir modelde ortalama kuyruk uzunluğu ile ortalama bekleme/hizmet süresi arasındaki ilişkileri ifade etmek için de kullanılır (Asmussen, 2003). Ancak, $M/G/s$ kuyruk modeli için performans metrikleri gibi bazı sorunlar hala çözülmeyi bekleyen bir durumdur.

Çizelge 3.5. $M/G/1$ Kuyruk Sistemi için Formüller

Sembol	Açıklama	Formül
L_q	Kuyrukta Olması Beklenen Birim Sayısı	$\frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2 \sigma^2}{2\left(1 - \frac{\lambda}{\mu}\right)}$
L	Sistemde Olması Beklenen Birim Sayısı	$\frac{\lambda}{\mu} + \frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2 \sigma^2}{2\left(1 - \frac{\lambda}{\mu}\right)}$
W_q	Kuyrukta Geçen Beklenen Süre	$\frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2 \sigma^2}{2\left(1 - \frac{\lambda}{\mu}\right) \lambda}$
W	Sistemde Geçen Beklenen Süre	$\frac{\frac{\lambda}{\mu} + \frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2 \sigma^2}{2\left(1 - \frac{\lambda}{\mu}\right)}}{\lambda}$

$G/G/s$

1940'lardan sonra kuyruk teorisi matematikçiler için önemli bir araştırma alanı haline gelmiştir (Whittle, 2002). 1957'de Pollaczek, $G/G/1$ sırasındaki integral denklemi kullanarak çalışmalar yaptı. John Kingman $G/G/1$ kuyruk sırasındaki ortalama bekleme süresine ilişkin bir

formül verdi ve Kingman'ın formülü olarak bilinir (Kingman, 1961). Bu kuyruk modelinde gelişler arası süre ve servis süresi üstel dağılım dışında bir dağılıma sahiptir. $G/G/1$ kuyruk modeli için Kingman, Pollaczek-Khinchine formülleri yardımıyla Kingman Difüzyonu adı verdiği bir yaklaşık çözüm metodu oluşturmuştur.

C_a^2 gelişler arası sürenin, C_s^2 servis süresinin değişim katsayısı olmak üzere, kuyrukta geçen beklenen süre yaklaşık olarak $W_q(G/G/1) \cong \left(\frac{C_a^2 + C_s^2}{2} \right) W_q(M/M/1)$ formülüyle hesaplanır.

Yapılan bir çalışmada Kingman Difüzyonu formülüyle elde edilen bu yaklaşık çözümün gerçek bekleme süresinin üst süresi olduğu hesaplanmıştır (Kraemer ve Langenbach-belz, 1976).

Kuyruk teorisinde sistemi tanımlayabilmek için üç ana karakteristik özellik tanımlanır. Bunlar, müşterilerin gelişler arası süreleri, servis süresi ve servis veren sunucu sayısıdır. Gelişler arası süre ve servis süresinin istatistiksel olarak tanımlanarak uygun parametrelerin ve kuyruk modelinin belirlenmesi, sistemin matematiksel olarak çözülmesi sağlar.

Yöneylem araştırması tekniklerinde, bir YBÜ kuyruk modelleri kullanılarak analiz edilebileceği düşünülebilir. Ancak mevcut durum için, hastalara hemşire, doktor ve yatak tahsisine yönelik karmaşık politikalar nedeniyle kuyruk modellerinin uygulanması büyük olasılıkla yetersiz kalır. Ayrıca, bu sistemin durağan olmayan gelişler, değişen personel seviyeleri gibi dinamik yapısı nedeniyle mevcut analitik kuyruk modelleriyle analitik olarak çözülemez. Kuyruk modelleri, sistemin bölümlerini analiz etmek için kullanılabilir, ancak tüm sistemin dinamik davranışının toplam analizi, bu tür modellerin kapasitesinin ötesindedir. Bu nedenle kuyruk teorisi yerine bu çalışmada simülasyon modelleme kullanılmıştır.

Doğrulama (verification), modelin istenen çıktıyı karşılayıp karşılamadığını kontrol etmekle ilgilidir. Bir model doğrulanırken, sonuçları bilinen test senaryolarından bir örnek seçilir ve modelin test senaryolarında beklenen çıktıyı ürettiğini gösterir. Test senaryolarının tarafsız bir şekilde seçilmesi ve test senaryolarının, modeli kullanırken ilgili olacak alanın büyük bir bölümünü kapsamaması önemlidir. Test durumları, modelin karşılaması gereken standardı oluşturur ve bu nedenle başkalarını doğrulama sürecinin yüksek kalitede olduğuna ikna etmede kritik öneme sahiptir. Geçerlilik (validation) ise, modelin amaçlanan hedefi karşılayıp karşılamadığıdır. Bir modelin geçerliliği test edilirken, modelin amaçlandığı şekilde kullanılabileceğine dair güvenilirlik oluşturmaya çalışır. Bu çalışmada, kuyruk teorisi yaklaşımından, simülasyon modelinin doğrulaması ve geçerliliğinin yapılması noktasında faydalanılır. Girdi parametresi olarak karma dağılımların kullanıldığı ve oldukça stokastik bir yapıya sahip YBÜ modelinin, kuyruk teorisinde herhangi bir kesin analitik çözümü bulunmamaktadır. Oluşturulan simülasyon modeli bloklarına analitik çözüme sahip $M/M/S$ girdi parametreleri verilerek çıktılar kontrol edilmiş ve simülasyon modelinin kesin sonuç değerlerini ürettiği görülmüştür. Böylece kurulan simülasyon modeli doğrulanmış ve geçerliliği kabul edilmiştir.

3.7.2. Simülasyon Modeli

Simülasyon, karmaşık sistemlerin modellenmesi için önemli bir araçtır. Sistemin bir bilgisayar programı aracılığıyla temsili, mühendislik tasarımı değişikliklerinin modellenen sistemde kesintiye uğramadan test edilmesi sağlanır.

Simülasyon modelleme; sistem modelleme, bilgisayar programlama, olasılık ve istatistik ve mühendislik tasarımı unsurlarını içerir. Simülasyon modelleme, bu bireysel olarak zorlayıcı konuları içerdiğinden, simülasyon modellemenin öğretilmesi ve öğrenilmesi zor olabilir. Ek olarak, simülasyon modelleme ile ilgili karmaşıklık nedeniyle, özel bilgisayar dillerine ihtiyaç duyar. Genel olarak, simülasyonlar birçok biçimde olabilir. Bir simülasyon, gerçek şeyin yalnızca bir temsildir. Simülasyonlarınızı zamanın ve teknolojinin elverdiği kadar gerçekçi hale getirilebilir ancak tamamen gerçek sistem gibi davranamaz.

Bir acil servis bölümünü simülasyon akışı şu şekilde tanımlanabilir. Hastalar acil servise gelir. Hastaların acil servise gelişi rasgele gerçekleşir ve haftanın gününe hatta günün saatine göre değişkendir. Hastane, gelen hastanın durumunun izlendiği bir triyaj istasyonuna sahiptir. Hastanın durumu acil müdahale gerektiriyorsa, hasta bir doktor ve bir hemşirenin gözetiminde acil servis yatağına götürür. Her hasta bilgilerinin alındığı kabul sürecinden geçer. Hasta daha sonra bir oda, doktor ve hemşire tahsis edilmesini beklemek üzere bekleme odasına yönlendirilir. Acil servisteki doktorlar ve hemşireler, testler yaparak ve hastanın semptomlarını teşhis ederek hastaların sağlığını izler. Bu döngü periyodik olarak çalışma süresi boyunca gerçekleşir. Akışın sonunda, hasta ya tedavi gördükten sonra taburcu edilir ya da hayatını kaybederek sistemden ayrılır.

Hastane, mevcut kaynakları daha iyi kullanabilmek için belirli performans ölçütlerinin davranışını anlamak gerekmektedir (Rossetti, 2021). Bu performans ölçütleri, bekleyen ortalama hasta sayısı, hastaların ortalama bekleme süresi ve acil serviste ortalama toplam süreleri, gereken ortalama yatak sayısı, doktorların ve hemşirelerin ortalama kullanımı gibi çıktılarıdır. Sağlık sistemi ilgili bir çalışma yürütürken, ünitenin gelecekte nasıl çalışacağına ilişkin değişiklikler önerilmesi ve böylece, sistemde değişiklik yaparken sistemin gelecekteki davranışını ve performansını tahmin edilmesi gerekir. Bu durumda, hastaların yaşamlarını veya bakımlarını tehlikeye atmadan gerçek sistemle gerçekçi bir şekilde denemeler yapılamaz. Bu nedenle, sistemin modellenmesi ve değişikliklerin model üzerindeki etkisinin test edilmesi daha doğrudur. Modelin kabul edilebilir bir doğruluğu varsa, değişikliklerin gerçek sistemi nasıl etkileyeceği tespit edilebilir. Simülasyon tekniklerinin kullanılacağı nokta tam olarak burasıdır.

Simülasyon modelinin avantajları şöyle sıralanabilir (Rossetti, 2021).

- Sistemdeki karmaşık etkileşimlerin ve rasgeleliğin performansı nasıl etkilediğini tespit eder,
- Farklı tasarımın hangi koşullar altında performans hedeflerini karşıladığını belirler ve alternatifleri karşılaştırır,

- Farklı durumlarda tekrar kullanılabilir karar destek sistemleri oluşturur,
- Yüksek maliyetli değişimlerde yeni kararı değerlendirir,
- Sistem henüz mevcut olmadığında tasarımı test edilir.

Bir simülasyon modeli hem tanımlayıcı hem de tahmine dayalı bir modeldir. Tanımlayıcı modelleme, bir sistemin geçmişteki davranışını anlamak için ne olduğunu açıklamak için zaman içerisinde kaydedilmiş verisi kullanır. Tahmine dayalı modelleme ise, gelecekteki davranışları anlamaya yardımcı olan modeller geliştirmek için kullanır.

Simülasyon modeli, bileşenler arasındaki ilişkileri fiziksel ve mantıksal kurallara bağlı olarak oluşturulur. Saf istatistiksel öngörü modellerinden farklı olarak simülasyon, sadece ilişkileri keşfetmeye yerine, bu ilişkileri açıkça temsil etme avantajına sahiptir. Simülasyon modellemenin önemli bir avantajı, tüm sistemi ve onun karmaşık ilişkilerini modelleme yeteneğine sahip olmasıdır. Simülasyonun temsil gücü, karmaşık süreçleri yakalamak için gerekli olan esnek modellemeyi sağlar. Sonuç olarak, sistemin farklı bileşenleri arasındaki tüm önemli etkileşimler model içinde açıklanır. Bu etkileşimlerin modellenmesi simülasyon modellemesinin doğasında vardır, çünkü simülasyon gerçek sistemin davranışını taklit eder. Sistemin gelecekteki davranışının tahmini, simüle edilmiş zamanın bir fonksiyonu olarak farklı modelleme senaryolarının davranışı izlenerek elde edilir. Gerçek dünya sistemleri genellikle analitik modeller için çok karmaşıktır ve genellikle doğrudan deney yapmak için çok pahalıdır. Simülasyon modelleri, bu karmaşıklığın modellenmesine izin verir ve gerçek sistemin nasıl davranabileceği hakkında çıkarımlar yapmak için düşük maliyetli deneyler sağlar.

Bir simülasyon modelinin temel amacı, belirli bir sistem hakkındaki gözlemlerin zamanın bir fonksiyonu olarak toplanmasına izin vermektir. Bu bakış açısından, ayrık (discrete) ve sürekli (continious) olmak üzere iki farklı simülasyon türü bulunur. Ayrık sistemde belirli değişiklikler meydana geldiğinde zamanın seçilen noktalarında toplanır. Zaman içinde seçilen bu noktalara olaylar denir. Öte yandan, sürekli simülasyonda, gözlemlerin zamanın her noktasında sürekli olarak toplanması gerekir (Rossetti, 2021).

Bir YBÜ'ye gelen hastaların dinamikleri oldukça hızlı değişir. Bu nedenle, bu çalışmada temel araç olarak farklı talep, yatış süresi ve ünite kapasitesi kombinasyonları altında kabul ve programlama politikalarının performansını araştırmak için simülasyon modeli yaklaşımı kullanılmıştır. Simülasyon, sistemin özelliklerine özel bir varsayım yapmadan karmaşık stokastik sistemleri ayrıntılı olarak modelleyebilme esnekliği ve karmaşık "eğer... ise" senaryolarını araştırabilme özelliği ile birlikte çok yönlülük sağlamaktadır. İyi kurulan bir simülasyon modeli, aynı zamanda sistemin gelecekteki ve herhangi bir başlangıç durumunda sistem davranışlarını da anlaşılmasına olanak sağlar (Antmen ve Oğulata, 2013).

Gerçek bir YBÜ sisteminin detaylı olarak modellenmesi oldukça zordur. Sistemin pek çok kısıtı ve değişkeni söz konusudur. Sağlık sisteminin tepkisi vakadan vakaya farklılaşabilir. Bu

nedenle sistemin kesin bir şekilde modellenebilmesi imkansıza yakındır. Sadeleştirici varsayımlar ve eldeki gerçek veri ile bir takım teknikler kullanılarak model oluşturulabilir.

Bu çalışmada kullanılan simülasyon modelinin amacı, YBÜ'ye gelen tüm hastalara sıra beklemeden hizmet verebilmek için gerekli yatak sayısını belirlemektir. Simülasyon modeli oluşturulmadan önce hastalara uygulanacak olan prosedürün daha iyi anlaşılabilmesi için bir algoritma oluşturulmuştur.

Oluşturan YBÜ akışında, modeldeki varlıklar rasgele aralıklarla gelen hastalardır. Algoritma, hastanın YBÜ'ye gelmesiyle başlar. Gelen hastaların sisteme kabul edilip edilmeyeceği boş yatağın varlığına bağlıdır. t anında gelen hastaya tüm yatakların dolu olması durumunda hasta tedavi görmeden sistemden çıkar. Her hastanın geliş zamanı hastane veri tabanına kaydedilir. YBÜ'ye kabul eden hastalar ilk gelen ilk hizmet alır (FCFS) politikasıyla yatakları işgal etme sırasına karar verilir. Yatağını temin edebilen hastalar ise belirlenen servis süresi kadar yatakta kalır ve servis süresi sonunda sistemden ayrılır. Sistemden ayrılış gerçek sistemde taburcu, ölüm ya da farklı bir üniteye sevk şeklinde gerçekleşebilir ancak veri setinde bu bilgi mevcut olmadığı için modelde herhangi bir ayırım yapılmamıştır.

ARENA, bilgisayar simülasyon modellerinin geliştirilmesini ve yürütülmesini kolaylaştıran blok tabanlı yazılım programıdır. Sürükle ve bırak akış şeması metodolojisini kullanarak modellerin oluşturulmasına izin veren bir ortam aracılığıyla SIMAN adlı temel bir simülasyon diline erişim sağlar. Ortam, dil modelleme yapılarına erişim sağlayan panellere sahiptir. Ayrıca ortam, modelin canlandırılması, modelin çalıştırılması ve sonuçların görüntülenmesi gibi yaygın simülasyon etkinliklerine araç çubuğu ve menü erişimi sağlar.

YBÜ'de hizmet gören hastaların olduğu sistem temelinde bir kuyruk modelidir. Fakat bu kuyruk modelinin matematiksel olarak çözülmesi mümkün değildir. Bu nedenle simülasyon modeli oluşturmak daha uygulanabilir bir yaklaşımdır.

Simülasyon modeli oluşturulurken bir takım varsayımlar kabul edilmiştir. Bu varsayımlar şu şekilde sıralanır.

- Hastalar sisteme rasgele gelir.
- Sistemde uygun yatak varsa hasta ilk gelen ilk hizmet alır (FCFS) politikasıyla kabul edilir, herhangi bir hasta öncelik politikası izlenmemektedir.
- YBÜ'ler 7 gün ve 24 saat (7/24) çalışmaktadır.
- Her iki (Pediatri ve Reanimasyon) YBÜ'de doktor ve hemşire sayısı yeterlidir.
- Tüm yataklar tam donanımlıdır ve gelen hastalara eksiksiz hizmet verir.

Replikasyon Sayısının Belirlenmesi

Simülasyon çalışmalarında, simüle edilen sistemin performansını değerlendirmek için birden fazla kez çalıştırılması gerekir. Bunun nedeni, simüle edilen sistemin rassalığı nedeniyle

farklı sonuçlar verecek olmasıdır. Bu yüzden, simülasyon çalışmasının sonuçlarının güvenilirliğini arttırmak için birçok replikasyon yapılır. Replikasyon sayısı, simülasyon çalışmasının planlamasının önemli bir parçasıdır. Replikasyon sayısı, simülasyon çalışmasının hedeflerine ve simüle edilen sistemin özelliklerine göre belirlenir. Daha karmaşık ve değişken bir sistemin simüle edildiği durumlarda, daha fazla replikasyon sayısı gerekebilir.

Bu çalışmada replikasyon sayısı belirlenirken ardışık yöntem metodu uygulanmıştır. Bu yöntemde, istenen güven aralığında (e) gereken replikasyon sayısı (n), $n = t_{\alpha/2, n-1} \frac{\sigma}{e}$ formülüyle hesaplanır.

Isınma Periyodunun Belirlenmesi

Kesikli olay simülasyonlarında, sistemin duraklama durumuna ulaşip ulaşmadığını belirlemek önemli bir sorundur. Doğru performans ölçütlerinin elde edilmesi, bu ölçümlerin tarafsız tahmin ediciler olmasına bağlıdır. Performans ölçütlerinin tarafsız olması için başlangıç koşullarından bağımsız olmaları gerekir. Başka bir deyişle, başlangıç koşulları farklı olsa bile aynı sonuçları verebilen bir simülasyon oluşturulmalıdır (Uncu, 2017).

Sistemin başlangıç zamanı $t = 0$ anında, hizmet noktası uzun dönem yoğunluğuna henüz başlanmamıştır ve bu koşullar gerçek sistemi temsil edemez. Sistemin gerçek durumu ifade edebileceği noktadan başlangıç süresine olan periyot, ısınma periyodu olarak adlandırılır. Bu ısınma periyodunun belirlenebilmesi için grafik yöntemleri veya yaklaşık süreler gibi bir takım metotlar geliştirilmiştir.

Isınma periyodunun olması gerekenden kısa tanımlandığı durumlarda simülasyon sonuçları yanlış olur, uzun tanımlandığı durumlarda ise, veri kaybı yaşanacağı için simülasyon sonucunda elde edilecek olan durağan durum parametrelerinde güvenilirlik düşer.

Bu çalışmada, literatür ışığında, ısınma periyodu toplam simülasyon süresinin %10'una ayarlamaya yönelik olan yöntem benimsenmiştir (Banks ve ark. 2005).

4. BULGULAR VE TARTIŞMA

Bu bölümde, Pediatri ve Reanimasyon Yoğun Bakım Ünitelerine (YBÜ) hasta gelişler arası (GAS) ve bu ünitelerde servis süreleri (SS) analiz edildi, simülasyon modeli oluşturuldu, karma dağılım model bileşen sayısı ve parametreleri tahmin edildi. Oluşturulan doğrulanmış simülasyon modeliyle, tüm hastalara bekleme süresi olmadan hizmet verecek yatak kapasitesi belirlenmesi dahil olmak üzere kurulan farklı senaryolar üzerinde sistem performansı izlendi.

4.1. Veri Analizi

Dağılım uydurma, istatistiksel veri analizinde önemli bir süreç olarak kabul edilir ve veri kümesine en iyi uyan istatistiksel dağılımı seçmek için kullanılır. Bir veri setinin istatistiksel olarak hangi dağılım kullanılarak temsil edilebileceğinin belirlenmesi, modelin gerçeğe yakın olmasını sağlar. Yanlış dağılım seçiminde, süreç doğru şekilde yansıtılamaz.

Pediatri YBÜ ve Reanimasyon YBÜ'den alınan GAS ve SS'nin tanımlayıcı istatistikleri, örneklem büyüklüğü (n), ortalama (μ), standart sapma (σ), medyan, çarpıklık ve basıklık olmak üzere Çizelge 4.1.'de özetlenmiştir.

Çizelge 4.1. Yoğun Bakım Ünitelerinin Tanımlayıcı İstatistikleri

YBÜ	Tür	n	μ	σ	Medyan	Çarpıklık	Basıklık
Pediatri	GAS	758	10,59	10,68	6,77	1,99	6,77
	SS	759	115,63	162,27	46,03	2,73	8,21
Reanimasyon	GAS	254	32,93	39,19	23,71	2,8	10,35
	SS	255	251,03	276,62	126,15	1,48	1,68

Ortalama ve medyan, merkezi eğilimi ölçerken, standart sapma, basıklık ve çarpıklık; verinin değişkenliği ölçer. Basıklık değeri +1'den büyükse dağılım tepe noktasına ulaşır. Normal bir tek değişkenli dağılımı göstermek için -1 ile +1 arasındaki bir çarpıklık kabul edilebilir olarak yorumlanır; aksi halde oldukça çarpık ve asimetric bir dağılıma işaret edilir.

Çizelge 4.1. incelendiğinde, Pediatri YBÜ'de GAS için 758 örnekleme yapılmıştır. İncelemeye alınan örneklemelerin ortalama değeri 10,59'dur. Standart sapma 10,68 olarak hesaplanmıştır, bu da örneklem verisinin dağılımının yayılımının geniş olduğunu gösterir. Medyan değeri 6,77'dir, yani örneklem verisinin yarısından fazlası bu değer altında yer almaktadır. Medyan ve ortalama değerinin birbirinden uzak olması verinin simetrik dağılım göstermediğine bir işaret olarak kabul edilebilir. Çarpıklık değeri 1,99'dur, bu pozitif bir çarpıklığı gösterir ve dağılımın sağa çekilmiş olduğunu gösterir. Basıklık değeri ise 6,77'dir ve dağılımın daha düz ve normal dağılıma göre daha sivri olduğunu gösterir.

Pediatric YBÜ'de SS için 115,63 olan 759 örneklem kullanılmış ve standart sapma 162,27 olarak bulunmuştur. Ortalama ve standart sapma değerinin birbirine yakın olması gözlemlerin geniş bir aralıkta dağıldığını gösterir. Medyan değeri 46,03'tür, yani örneklem verisinin yarısından fazlası bu değerin altında yer almaktadır. Medyan değerinin ortalama değerde olan uzaklığı oldukça fazladır, Pediatric YBÜ GAS'ye benzer olarak verinin asimetrik bir şekilde dağılım göstereceği bu değerler üzerinden anlaşılabilir. Çarpıklık değeri 2,73'tür, yani dağılım sağa çarpıktır. Basıklık değeri 8,21'dir, bu değer verinin normal dağılıma göre daha sivri bir dağılıma sahip olduğunu gösterir.

Reanimasyon YBÜ'de GAS için 254 örnekleme yapılmıştır. İncelemeye alınan örneklemelerin ortalama değeri 32,93'tür. Standart sapma 39,19 olarak hesaplanmıştır. Medyan değeri 23,71'dir. Bu değerlere göre verinin, geniş bir yaygınlık ve çarpıklık gösterdiği yorumu yapılır. Çarpıklık değeri 2,8'dir, bu pozitif bir çarpıklığı gösterir. Basıklık değeri ise 10,35'tir ve veri sivri bir dağılıma sahiptir.

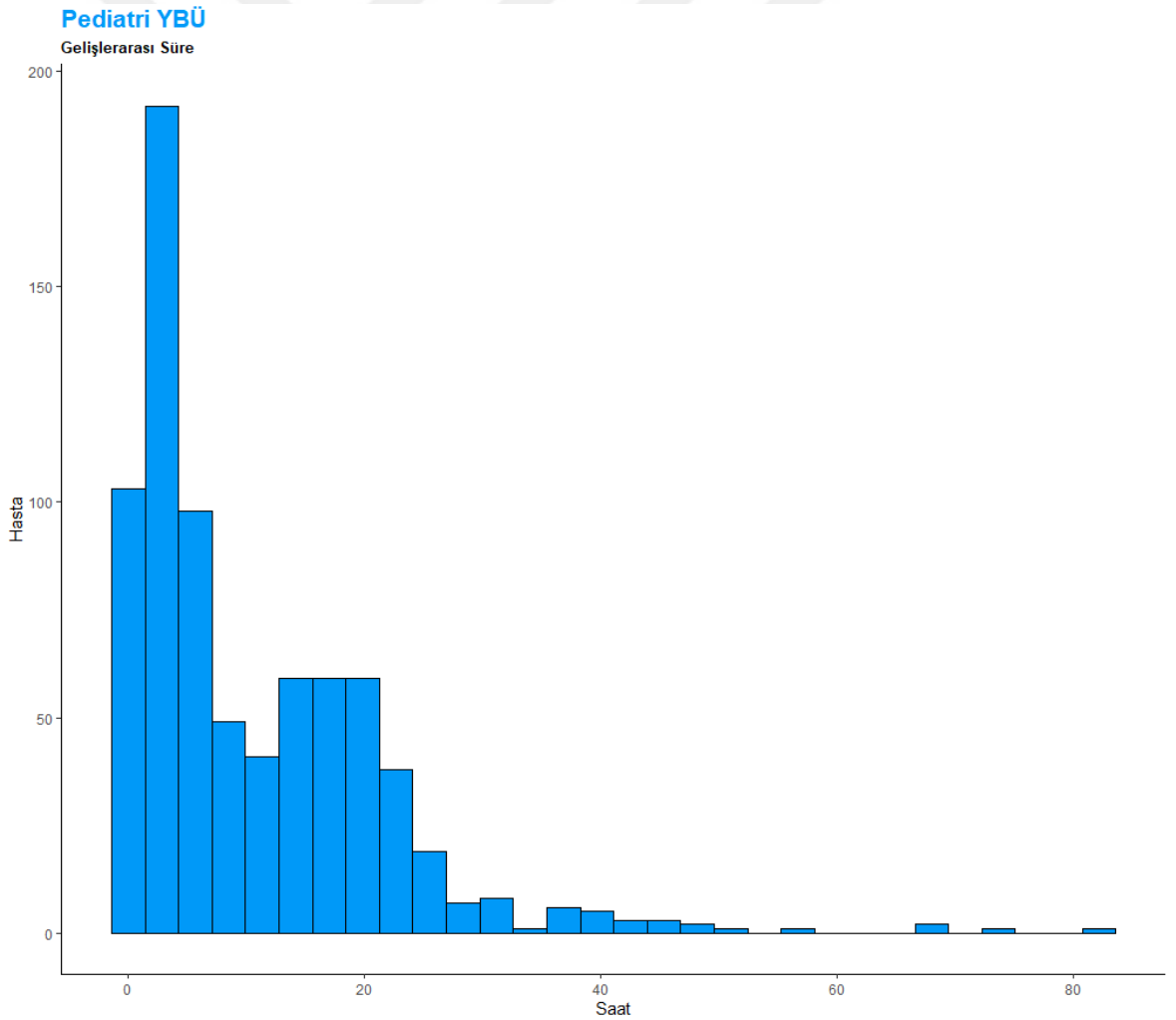
Reanimasyon YBÜ'de SS ortalama değeri 251,03 olan 255 örneklem için standart sapma 276,62'dir, bu verinin yayılımının daha geniş olduğunu gösterir. Medyan değeri 126,15'tir, yani örneklem kümesinin tepe noktası simetri çizgisinden uzaktır. Çarpıklık değeri 1,48'dir, bu dağılımın sağa çekik olduğunu gösterir. Basıklık değeri ise 1,68'dir ve dağılımın sivri bir tepe noktası olduğunu ifade eder.

Çizelge 4.1.'de hesaplanan tanımlayıcı istatistikler, her bir YBÜ için GAS ve SS gözlemlerinin dağılımı ve karakteristikleri hakkında bilgi sağlar. Örneğin, her iki birimde çarpıklık değerlerinin pozitifdir, yani dağılımların sağa çekik olduğu dikkat çeker. Basıklık değerleri ise birimler arasında farklılık göstermektedir, ancak her iki birimin de normal dağılıma göre daha sivri dağılımlara sahip olduğu yorumu yapılabilir. Her iki YBÜ içinde GAS ve SS verisi asimetrik dağılmaktadır.

Çizelge 4.1.'de gösterildiği gibi, tüm veri asimetrik, basıklık ve çarpıklık değerlerine göre keskin tepe noktalarına sahiptir. Veri simetrik ise, ortalama ve medyan benzer olmalıdır. Ortalama ve medyan değerlerinden görüldüğü üzere; Pediatric ve Reanimasyon YBÜ'lerinin GAS ve SS'leri simetrik bir dağılım göstermez. Bu yorumlar, gözlemlerin istatistiksel analizini ve birimler arasındaki farkları anlamak için kullanılabilir. Ancak, daha kapsamlı bir analiz yapmak için diğer faktörleri de dikkate almak ve daha fazla istatistiksel yöntem uygulamak gerekebilir. Bu yöntemlerin başında grafiksel yöntemler gelmektedir.

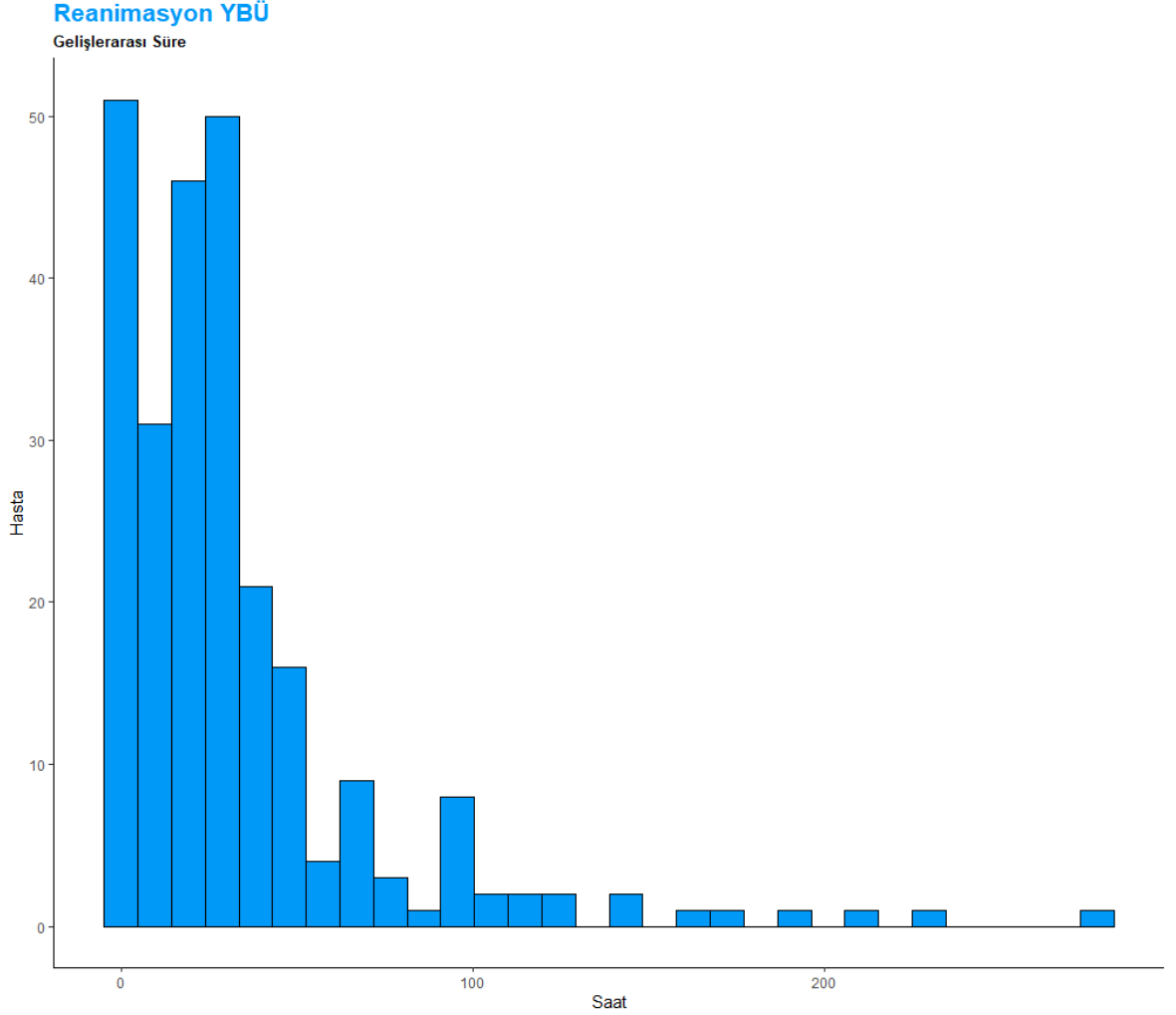
Histogram grafiği, verinin dağılımını görselleştirmek ve gözlemlerin nasıl dağıldığını anlamak için kullanılan bir grafiksel yöntemdir. Grafiğin x-ekseni, veri aralığını temsil ederken, y-ekseni ise her bir veri aralığındaki gözlem sayısını temsil eder. Histogramdaki sütunların yükseklikleri, her bir veri aralığındaki gözlem sayısını temsil eder. Yüksek sütunlar, o veri aralığında daha fazla veri noktasının bulunduğunu gösterirken, düşük sütunlar daha az veri noktasının olduğunu gösterir. Histogramdaki sütunların genişlikleri, veri aralıklarının büyüklüğünü

temsil eder. Geniş sütunlar, daha geniş bir veri aralığına sahip olduğunu gösterirken, dar sütunlar daha dar bir veri aralığına sahip olduğunu gösterir. Histogramdaki zirve noktaları, verinin yoğunlaştığı yerleri temsil eder. Sütunların en yüksek olduğu nokta, gözlemlerin en yoğun olduğu değeri gösterir. Aynı şekilde, dip noktaları, verinin en seyrek olduğu noktaları temsil eder. Histogram, aykırı değerleri tespit etmek için de kullanılabilir. Aykırı değerler, diğer veri noktalarından belirgin bir şekilde farklı olan nadir gözlemlerdir. Histogramda, aykırı değerler genellikle diğer sütunlardan uzakta tek bir veya birkaç sütun olarak görülebilir. Histogram, veri dağılımının simetri veya çarpıklık durumunu gözlemlenmesine yardımcı olur. Eğer sütunlar simetrik bir şekilde dağılmışsa, veri dağılımı simetrik olarak kabul edilir. Eğer sütunlar sağa veya sola yığılmış bir şekilde dağılıyorsa, veri dağılımı çarpık olarak tanımlanabilir. Histogram grafiği, veri dağılımını hızlı ve görsel bir şekilde yorumlanmasını sağlar. Yükseklik, genişlik, simetri, çarpıklık, zirve noktaları ve aykırı değerler gibi özelliklere odaklanarak, veri hakkında hızlı bilgi sahibi olunur.



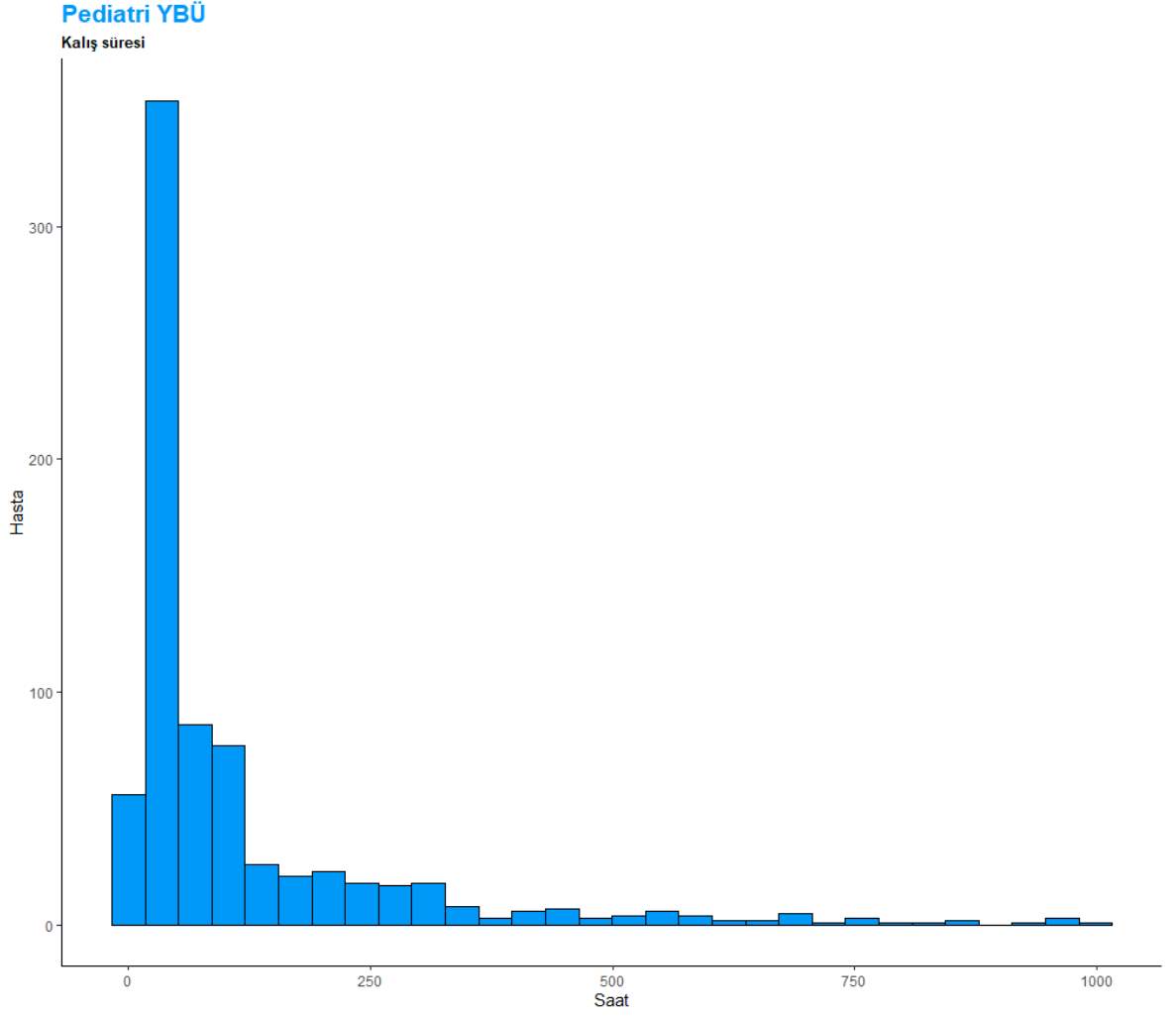
Şekil 4.1. Pediatri YBÜ GAS Histogramı

Şekil 4.1.'de Pediatri YBÜ'ye hastaların GAS verisi histogramı çizilmiştir. Buna göre veri, asimetrik ve sağa çarpık bir dağılıma sahiptir. Histogramdaki zirve noktası, sisteme giren ardışık hastaların GAS'nin kısa periyotlarda olduğunu gösterir.



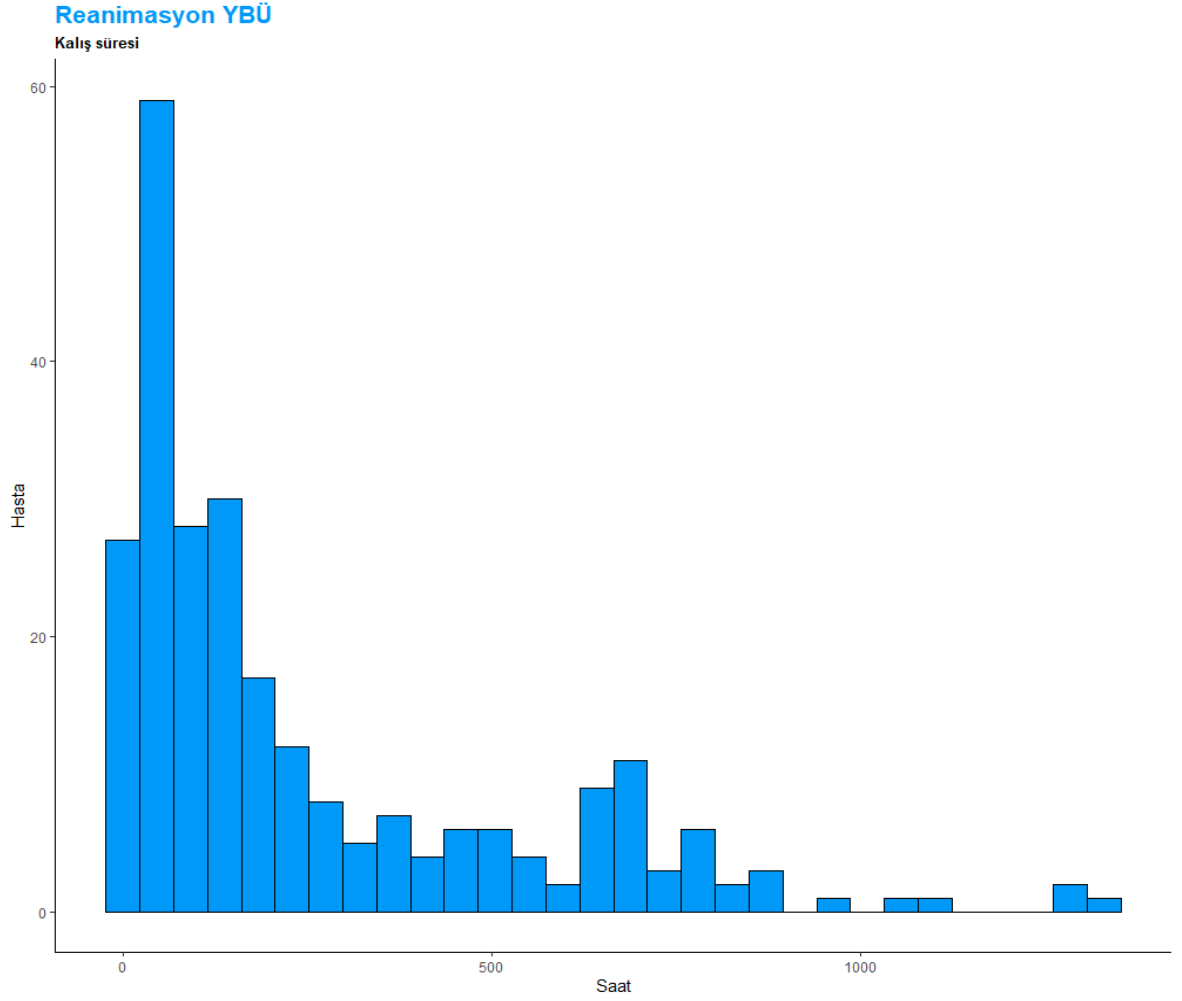
Şekil 4.2. Reanimasyon YBÜ GAS Histogramı

Şekil 4.2.'de Reanimasyon YBÜ'ye hastaların GAS verisi histogramı çizilmiştir. Pediatri YBÜ'ye gelen hastaların GAS histogramı grafiğine benzer olarak, gözlemlerin asimetrik ve sağa çarpık bir dağılıma sahiptir. Histogramdaki zirve noktası, sisteme giren ardışık hastaların GAS verisinin Pediatri YBÜ'ye gelen hastaların GAS verisine göre daha uzun periyotlarda olduğunu göstermektedir.



Şekil 4.3. Pediatri YBÜ SS Histogramı

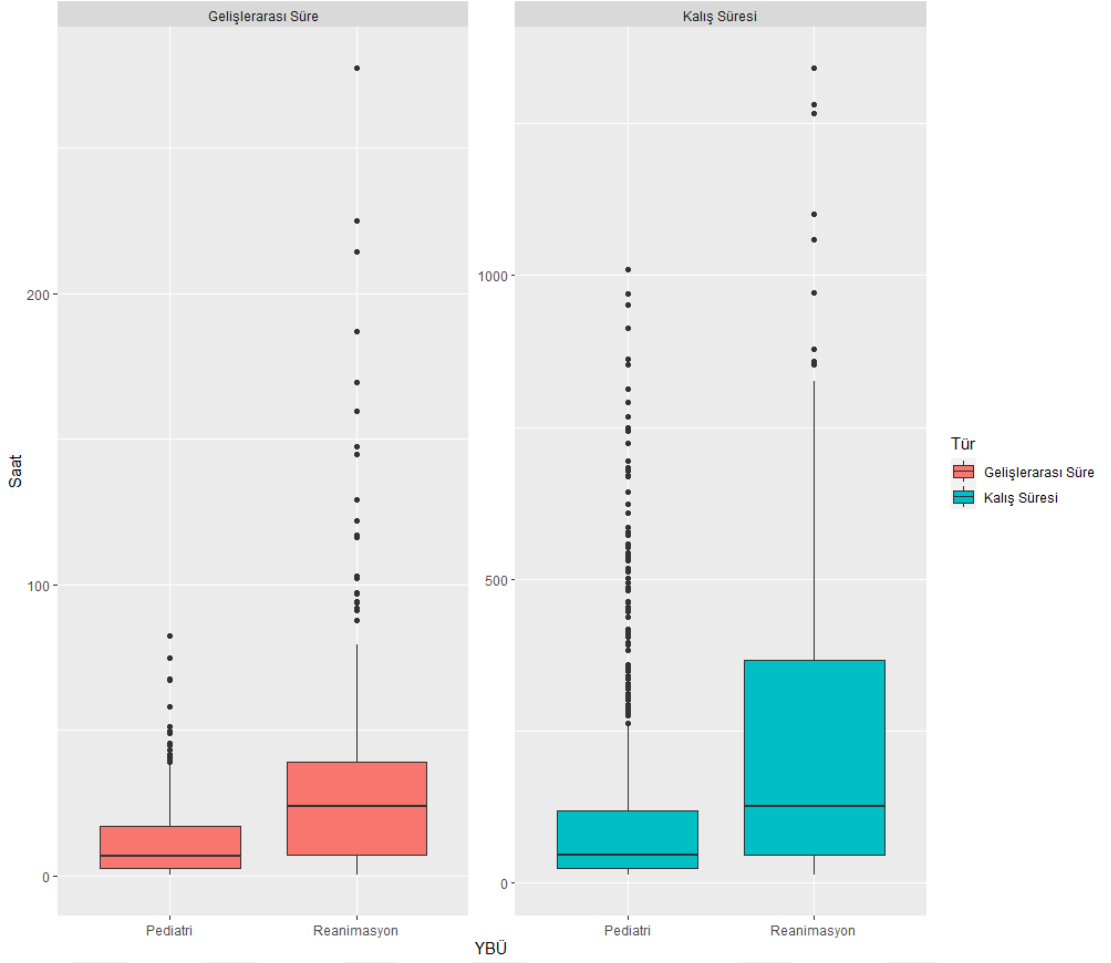
Şekil 4.3'te Pediatri YBÜ'de kalan hastaların SS histogramı çizilmiştir. Buna göre verinin, asimetrik ve sağa çarpık bir dağılıma sahiptir. Histogramdaki zirve noktası, hastaların ortalama kalış süresine yakın bir değer aldığı görülmüştür. SS gözlemlerin 250 saatten fazla olduğu durumlar aykırı durumlar olarak nitelendirilebilir, bu değerden daha büyük değeri alan gözlem sayıları oldukça azdır.



Şekil 4.4. Reanimasyon YBÜ SS Histogramı

Şekil 4.4'te ise, Reanimasyon YBÜ'de hastaların SS verisi histogramı paylaşılmıştır. Buna göre Pediatri YBÜ'ye benzer olarak Reanimasyon YBÜ'de de, SS verisi asimetric ve sağa çarpıktır. Histogramdaki zirve noktası, gözlemlerin yoğunlaştığı ve hastaların ortalama kalış süresine yakın değerlerdir.

Pediatri ve Reanimasyon YBÜ'lere gelen hastaların GAS ve bu YBÜ'lerde gözlenen SS gözlem histogramları grafikleri incelendiğinde, gözlemlerin sağa çarpık ve asimetric olduğu yorumu yapılır. Ancak histogramı grafikleri bir veri setini tanımlamak için tek başına yeterli değildir. Bu nedenler farklı grafikler de incelenmeli ve veri en uyumu şekilde istatistiksel olarak analiz edilmelidir.



Şekil 4.5. Pediatri ve Reanimasyon YBÜ'lerin GAS ve SS Kutu Grafikleri

Kutu grafiği, veri setinin minimum, birinci çeyrek, medyan, üçüncü çeyrek, maksimum ve aykırı değerlerin görsel olarak sunulduğu bir grafik türüdür. Kutu grafiğinin ortasında yer alan çizgi, veri setinin ortanca değeri olan medyanyı temsil eder. Eğer çizginin tam ortasında bir kutu bulunuyorsa, veri seti simetrik olarak kabul edilir. Aksi halde, medyanın konumu veri setinin sağa veya sola çarpık olduğunu gösterir. Kutu grafiğindeki kutu, verinin çeyreklerini temsil eder. Kutunun yüksekliği, gözlemlerin yayılımını gösterir. Eğer kutu yüksekse, veri setinin yayılımı geniş; düşükse, veri setinin yayılımı dar olur. Kutu grafiğindeki "whisker" (bıyık) adı verilen çizgiler, veri kümesinin yayılımını gösterir. Eğer çizgilerden öteye uzanan noktalar yani aykırı değerler bulunuyorsa, bu noktalar veri setinden belirgin bir şekilde farklı olan nadir gözlemleri temsil eder. Kutu grafiği, veri dağılımının simetri veya çarpıklık durumunun gözlemlenmesine de yardımcı olur. Eğer kutu grafiği simetrik olmak yerine, çizgilerden biri diğerinden daha uzunsa, veri seti sağa veya sola çarpık olduğu yorumu yapılmaktadır.

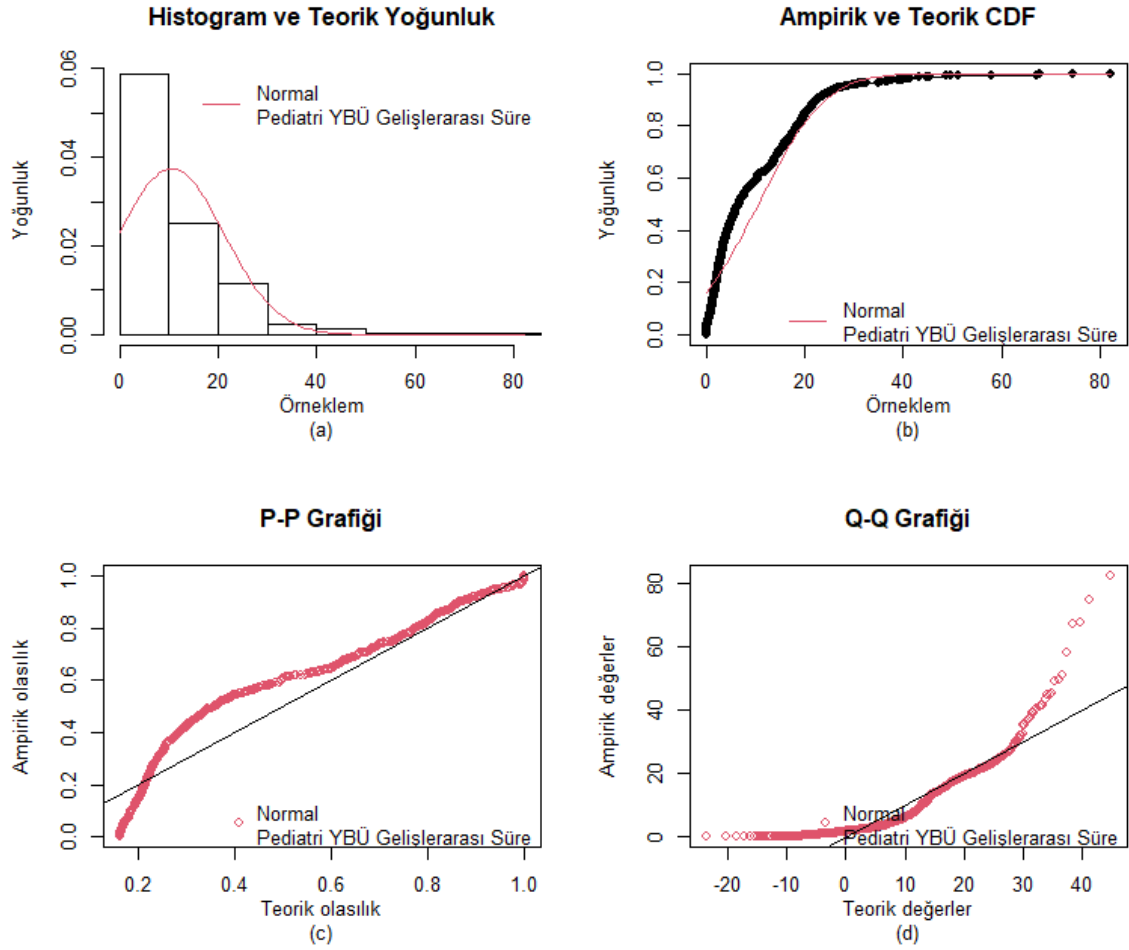
Pediatri ve Reanimasyon YBÜ'lerin GAS ve SS veri analizi için kutu grafikleri çizilmiş ve Şekil 4.5.'te paylaşılmıştır. Şekil 4.5. incelendiğinde, Pediatri YBÜ GAS için, medyan değeri kutunun alt çizgisine yakın bulunmaktadır. Bu veri setinin asimetric ve sağa çarpık olduğunu göstermektedir. Benzer şekilde, Reanimasyon YBÜ GAS gözlemlerinde de asimetric söz

konusudur. Alt ve üst değerleri incelendiğinde; Reanimasyon YBÜ'nün GAS verisi, Pediatri YBÜ GAS verisine göre göre daha yaygın bir değer aralığına sahiptir. Pediatri YBÜ'ye hasta gelişleri, Reanimasyon YBÜ'ye hasta gelişlerinden daha sık şekilde gerçekleşir. Her iki YBÜ'de de GAS verisinde aykırı değerler bulunmaktadır. Bu veri setinin OYF'sinde uzun kuyruklara sebep olur. SS gözlemlerinin medyan çizgisi ise, her iki YBÜ için de ortalamadan küçüktür ve sağa çarpıklık ve aykırı değerlerin varlığı yorumu yapılır.

Şekil 4.5.'de paylaşılan kutu grafikleri genel olarak yorumlandığında, Pediatri ve Reanimasyon YBÜ'lerin GAS ve SS verisi asimetrik, sağa çarpık ve aykırı değerlere sahiptir. Verilerin bu genel yapısı, verisi bilinen herhangi bir saf istatistik dağılıma uymama ihtimali giderek kuvvetlenmektedir. Ancak bu yöntemler kesin yargı için yeterli değildir, Materyal ve Metot bölümünde ayrıntılı şekilde bahsedildiği gibi, seçilen her bir aday dağılım için OYF, P-P, Q-Q grafikleri de incelenmeli ve seçilecek aday dağılımların uyum iyiliği testleri sonuçları yorumlanmalıdır.

Verilerin genel yapısını ve eğilimini izlemenin başka bir görsel yolu, bir P-P grafiği oluşturmaktır. P-P grafiği, ampirik KDF değerlerini teorik KDF değerlerine karşı çizer. P-P grafiği düz bir çizgiye yakınsa, belirtilen dağılım veriye uyar. Bir veri kümesinin potansiyel olarak bazı teorik dağılımlardan gelip gelmediğini değerlendirmek için ise Q-Q grafiği kullanılır. Veri seçilen dağılıma uygun dağılıyorsa, bir Q-Q çizimindeki noktalar düz bir çapraz çizgi üzerinde yer alır. Tersine, Q-Q grafiğindeki noktalar düz çapraz bir çizgiden ne kadar fazla saparsa, veri setinin seçilen dağılımı izlemesi o kadar az olasıdır.

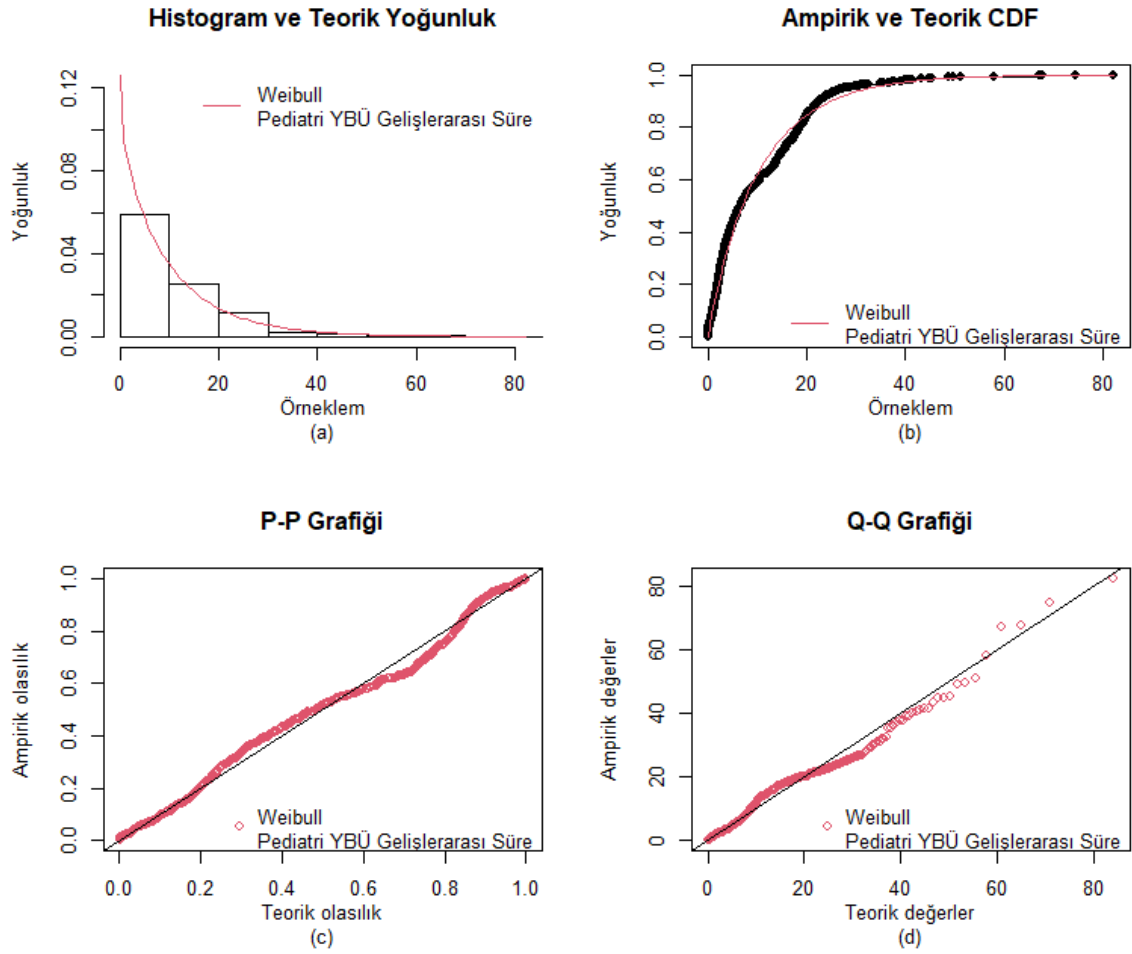
Pediatri ve Reanimasyon YBÜ'lere gelen hastaların GAS ve YBÜ'lerde SS veri kümeleri için aday dağılım olarak normal, Weibull, gamma, lognormal ve üstel dağılımlar seçilmiştir. Bu aday dağılımların ilk istatistiksel testleri olarak OYF'ler, P-P ve Q-Q grafiklerinden faydalanılmıştır.



Şekil 4.6. Pediatri YBÜ'ye GAS Verisinin Normal Dağılım Uydurma Grafikleri

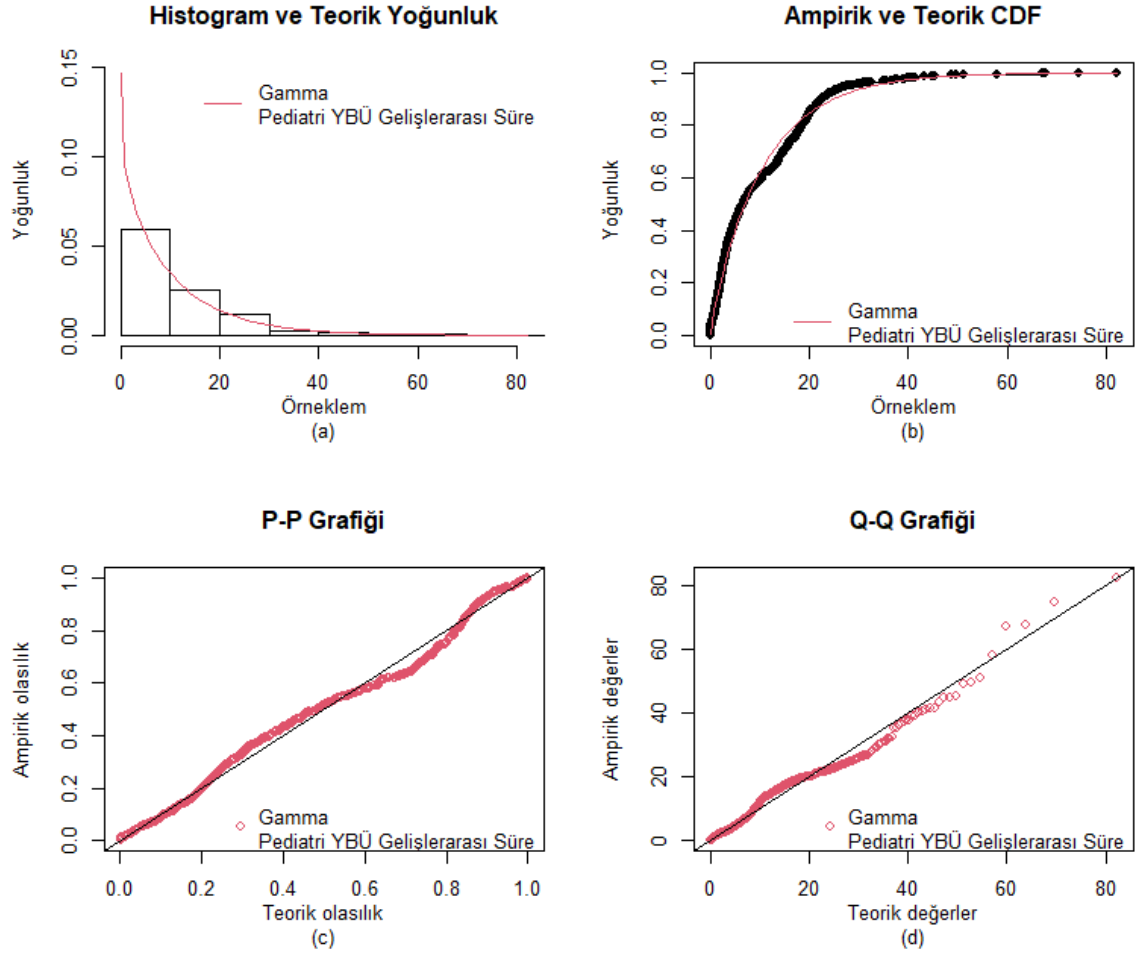
Şekil 4.6.'da, Pediatri YBÜ'ye GAS verisi normal dağılıma uydurulmaya çalışılarak elde edilen grafikler çizilmiştir. Şekil 4.6. (a), normal dağılım teorik OYF'si ile gözlenen Pediatri YBÜ'ye GAS gözlemlerinin histogramıdır. Gözlenen değerlerin normal dağılıma uymayan bir yayılım gösterdiği ilk bakışta söylenebilir. Şekil 4.6. (b), Pediatri YBÜ'ye GAS verisinin KDF'si ile normal dağılım KDF'nin karşılaştırıldığı grafiğine iyi bir uyum sağlamadığı ve sapmalar olduğu görülmektedir. Şekil 4.6. (c), teorik normal dağılım olasılık değerinin Pediatri YBÜ'ye GAS olasılık değeriyle kıyaslandığı P-P grafiğidir. Gözlenen Pediatri YBÜ'ye GAS verisinin seçilen normal dağılıma uyması durumunda, veri P-P grafiğindeki lineer çizgi üzerine konumlanacaktır. Görüldüğü üzere, verisi temsil eden noktalar çizgi üzerinden sapmıştır. Böylece gözlenen Pediatri YBÜ'ye GAS verisinin normal dağılıma uymadığı söylenebilir. Şekil 4.6. (d) ise, Pediatri YBÜ'ye GAS değerlerinin, seçilen aday dağılım olan normal dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. P-P grafiğine benzer şekilde, veri seçilen dağılıma uyarsa, Q-Q grafiğinde noktalar düz bir çapraz çizgi üzerinde sıralanması beklenmektedir. Şekil 4.6. (d) incelendiğinde, gözlemlerin Q-Q lineer çizgisi etrafında toplanmadığı ve seçilen normal dağılımı takip etmediği

yorumu kolaylıkla yapılabilir. Genel olarak Şekil 4.6. değerlendirildiğinde, analiz edilen verinin aday dağılım olarak seçilen normal dağılıma uygun olmadığı söylenebilir.



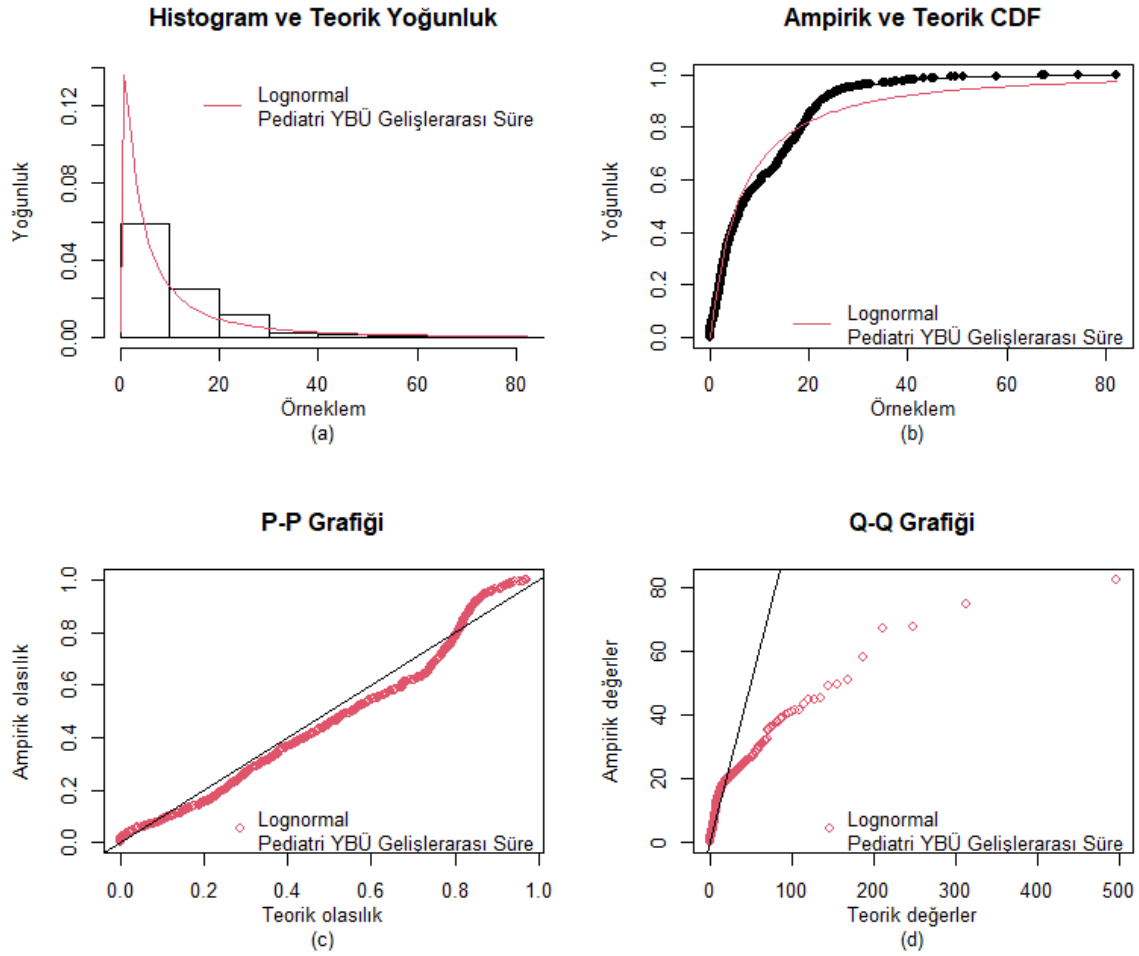
Şekil 4.7. Pediatri YBÜ’ye GAS Verisinin Weibull Dağılım Uydurma Grafikleri

Şekil 4.7.’de, Pediatri YBÜ’ye GAS Weibull dağılıma olan uyumu izlenmek için grafikler çizilmiştir. Şekil 4.7. (a), Weibull dağılım teorik OYF’si ile gözlenen Pediatri YBÜ’ye gelişler arası süre değerlerin histogramıdır. Gözlenen değerlerin Weibull dağılıma uymadığı görülmektedir. Şekil 4.7. (b)’de, Pediatri YBÜ’ye GAS verisinin KDF’si ile Weibull dağılım KDF’si karşılaştırılmış ve seçilen aday Weibull dağılımın veriye iyi bir uyum sağlamadığı tespit edilmiştir. Şekil 4.7. (c), teorik Weibull dağılım olasılık değerinin Pediatri YBÜ’ye GAS olasılık değeriyle kıyaslandığı P-P grafiğinde çizgi üzerinde veri toplanmamış ve salınım yapmıştır, böylece gözlemlerin Weibull dağılımına uymadığı söylenebilir. Şekil 4.7. (d) ise, Pediatri YBÜ’ye GAS değerlerinin, seçilen Weibull dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. Bu grafik incelendiğinde, gözlemlerin seçilen Weibull dağılımı takip etmediği yorumu kolaylıkla yapılır. Genel olarak Şekil 4.7. değerlendirildiğinde, analiz edilen verinin aday dağılım olarak seçilen Weibull dağılıma uymadığı yorumu yapılır.



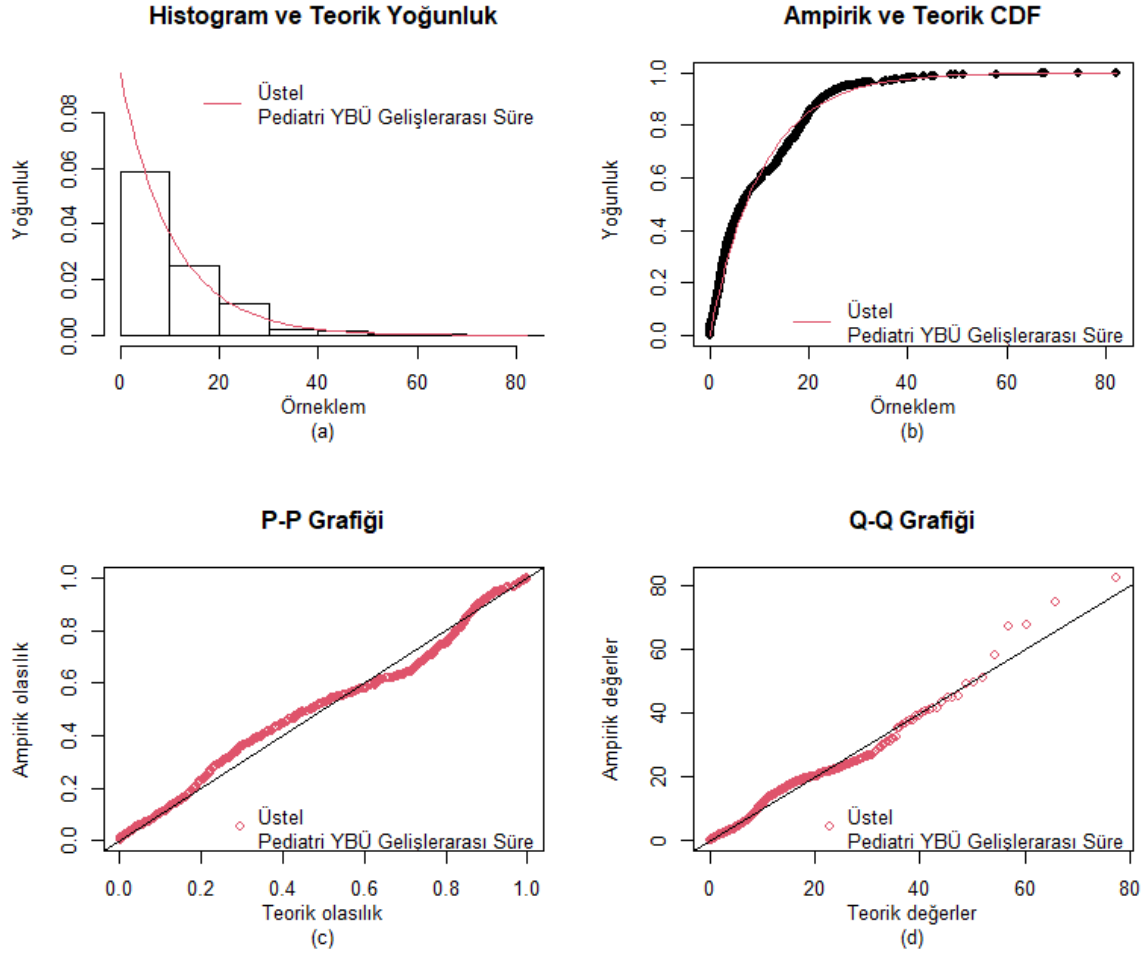
Şekil 4.8. Pediatri YBÜ’ye GAS Verisinin Gamma Dağılım Uydurma Grafikleri

Şekil 4.8.’de, Pediatri YBÜ’ye GAS verisinin gamma dağılıma olan uyumu izlenmiştir. Şekil 4.8. (a)’da gamma dağılım teorik OYF’si ile gözlenen Pediatri YBÜ’ye GAS değerleri histogramının uyumsuz olduğu görülmektedir. Şekil 4.8. (b)’de gözlemlerin KDF’si seçilen aday dağılım olan gamma dağılımın KDF’sine tam olarak uyum sağlamadığı görülmektedir. Şekil 4.8. (c)’deki P-P grafiğinde, gözlemler çizgi etrafında salınım yapmaktadır ve aday gamma dağılımına verinin uymadığı yorumunu yapılabilir. Şekil 4.8. (d)’deki Q-Q grafiğinde de veri çizgi etrafında toplanmamıştır ve gözlemler gamma dağılımının uygun değildir. Genel olarak Şekil 4.8. değerlendirildiğinde, ele alınan Pediatri YBÜ’ye GAS verisinin, aday dağılım olarak seçilen gamma dağılımına bir uyum göstermediği yorumu yapılır.



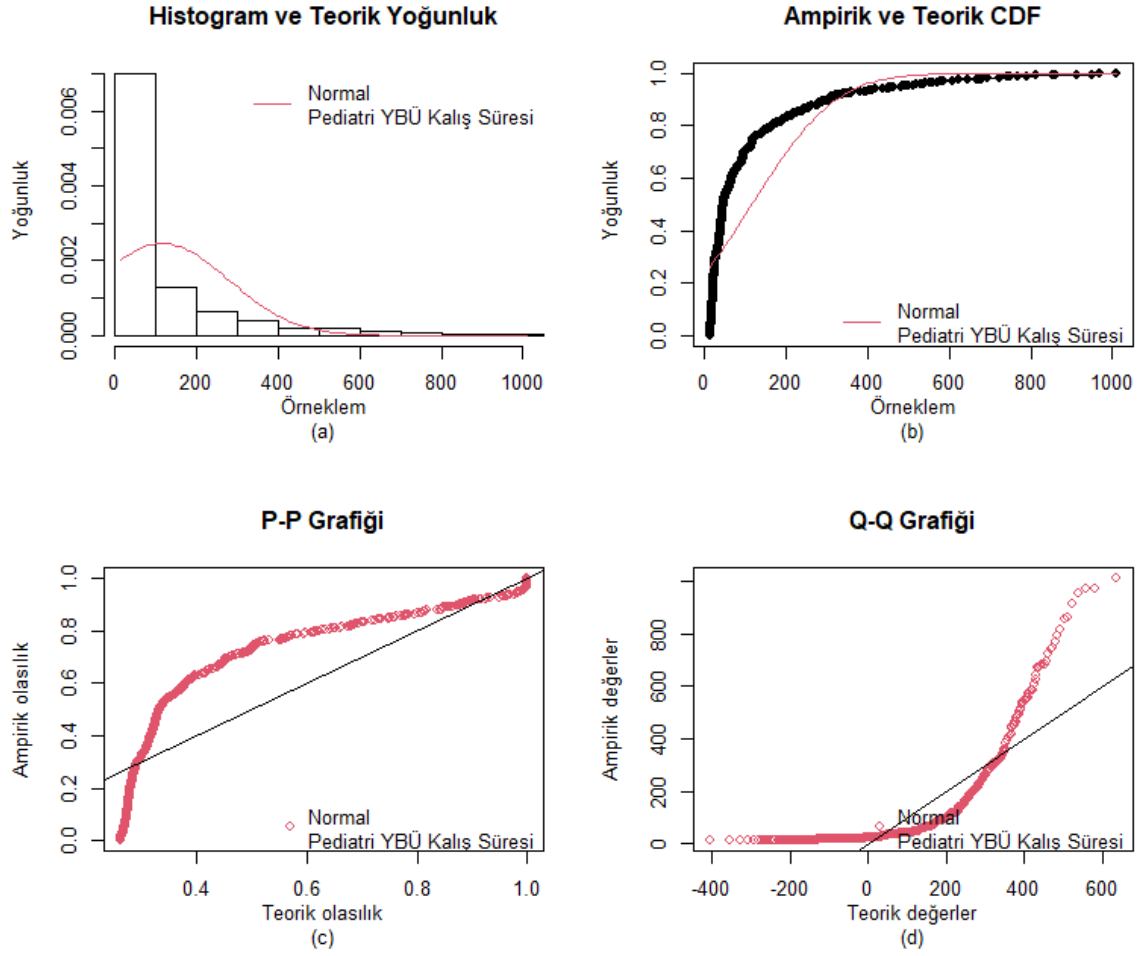
Şekil 4.9. Pediatri YBÜ’ye GAS Verisinin Lognormal Dağılım Uydurma Grafikleri

Şekil 4.9.’da, Pediatri YBÜ’ye GAS lognormal dağılıma olan uyumu test edilmek için grafikler çizilmiştir. Şekil 4.9. (a), lognormal dağılım teorik OYF’si ile gözlenen Pediatri YBÜ’ye GAS değerlerinin histogramıdır ve uyumsuz olduğu görülmektedir. Şekil 4.9. (b)’de, Pediatri YBÜ’ye GAS verisinin KDF’siyle ile aday lognormal dağılım KDF’si de birbirine uyum göstermemektedir. Şekil 4.9. (c)’de, P-P incelendiğinde gözlemler lognormal dağılımı destekleyen çizgiden sapmıştır ve aday dağılıma uyum sağlamamıştır. Şekil 4.9. (d) ise gözlemlerin, seçilen lognormal dağılımdan gelen değerler ile karşılaştırıldığı Q-Q grafiğidir ve veri, bu grafikte oldukça çarpık bir salınım göstermesi nedeniyle seçilen lognormal dağılıma kesinlikle bir uyum sağlamadığı ifade edilebilir. Genel olarak Şekil 4.9. değerlendirildiğinde, analiz edilen veri, aday dağılım olarak seçilen lognormal dağılıma istatistiksel olarak uyumsuzdur.



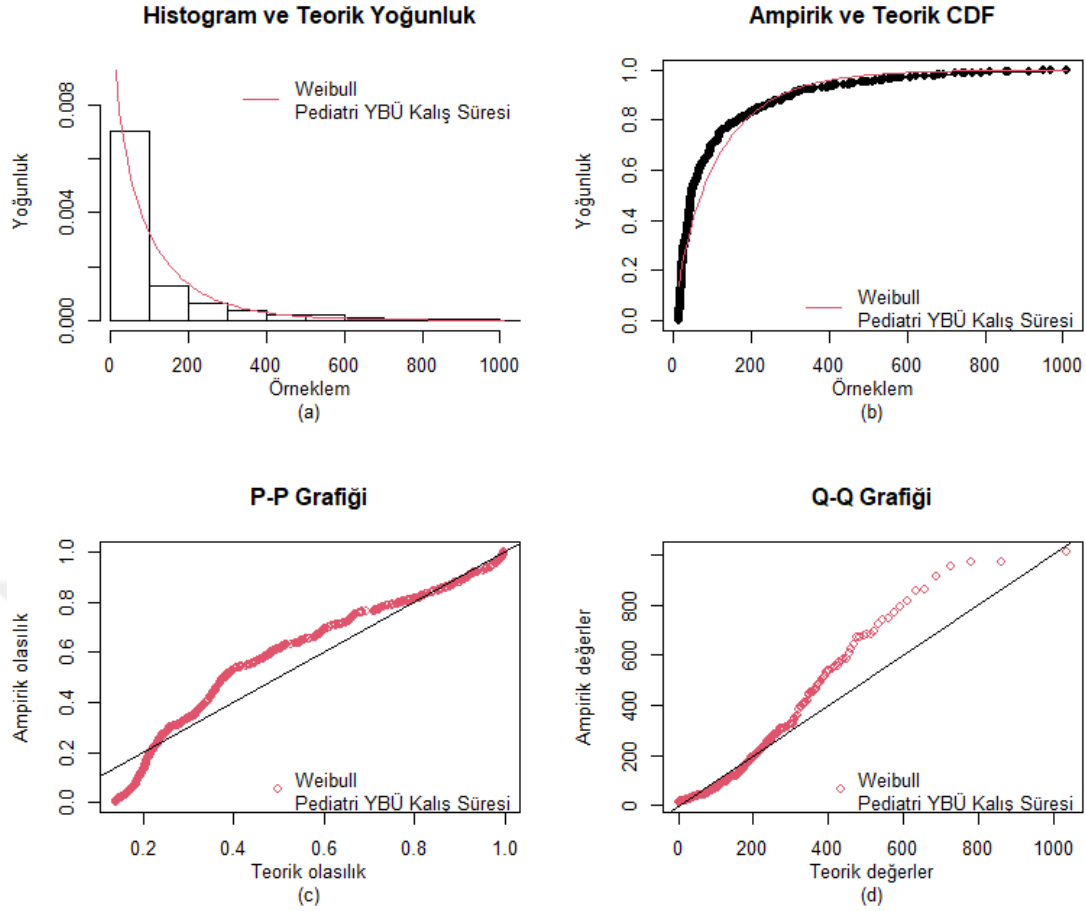
Şekil 4.10. Pediatri YBÜ'ye GAS Verisinin Üstel Dağılım Uydurma Grafikleri

Şekil 4.10.'da, Pediatri YBÜ'ye GAS üstel dağılıma uyup uymadığı test edilmek istenmiştir. Şekil 4.10. (a), üstel dağılım teorik OYF'si ile gözlenen değerlerin histogramının uyum göstermediği söylenebilir. Şekil 4.10. (b)'de, verilerin KDF'si seçilen aday dağılım olan üstel dağılımın KDF'sine tam olarak uyum sağlayamamıştır. Şekil 4.10. (c)'deki P-P grafiğine göre, veri çizgi etrafında salınım yapmaktadır. Böylece gözlene verinin, üstel dağılımına uymadığını söylenebilir. Şekil 4.10. (d)'deki Q-Q grafiğini incelendiğinde, seçilen aday üstel dağılıma, Pediatri YBÜ'ye GAS verisinin uygun olmadığını yorumu yapılabilir. Genel olarak Şekil 4.10. değerlendirildiğinde, ele alınan Pediatri YBÜ'ye GAS verisinin, aday dağılım olarak seçilen üstel dağılıma uyum göstermediği yorumu yapılabilir.



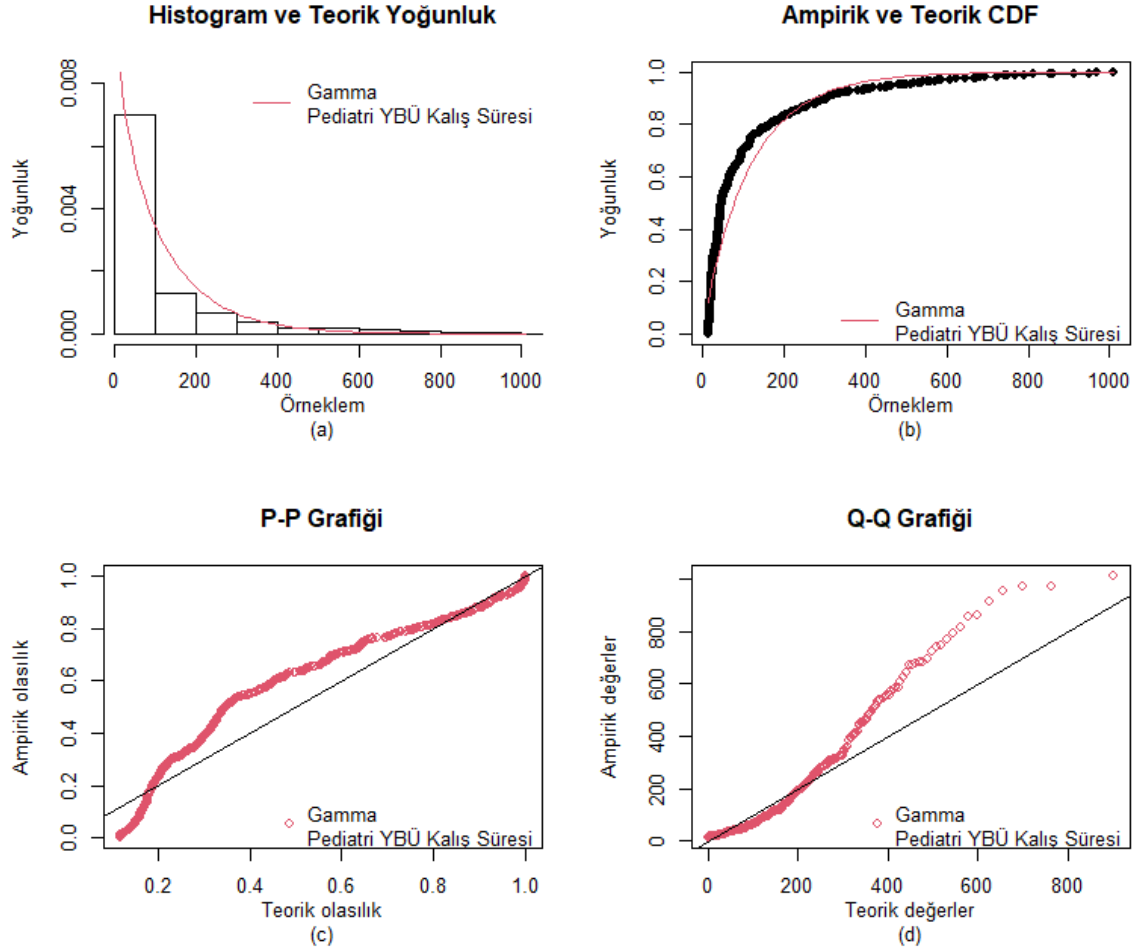
Şekil 4.11. Pediatri YBÜ’deki SS Verisinin Normal Dağılım Uydurma Grafikleri

Şekil 4.11.’de, Pediatri YBÜ’deki SS’in normal dağılıma uyup uymadığını test etmek için çizilen grafikler çizilmiştir. Şekil 4.11. (a), normal dağılım teorik OYF’si ile gözlenen Pediatri YBÜ’deki SS değerlerinin histogramıdır. Gözlenen verinin normal dağılıma uymayan bir yayılım gösterdiği ilk bakışta söylenebilir. Şekil 4.11. (b)’de gözlenen verinin KDF’sinin seçilen normal dağılım KDF’sine hiçbir şekilde uyum sağlamadığı açıkça görülmektedir. Şekil 4.11. (c)’de çizilen P-P grafiğinde, gözlenen verinin seçilen normal dağılıma uyması durumunda, lineer çizgi üzerine konumlanması beklenir. Görüldüğü üzere, verisi temsil eden noktalar çizgi üzerinden yüksek oranda sapmıştır. Böylece gözlemlerin normal dağılıma uymadığı rahatlıkla söylenebilir. Şekil 4.11. (d)’de Q-Q grafiği incelendiğinde, lineer çizgisi etrafında toplanmadığı ve bu sebeple seçilen normal dağılımı uymadığı söylenebilir. Genel olarak Şekil 4.11. değerlendirildiğinde, analiz edilen verinin, aday dağılım olarak seçilen normal dağılıma uygun olmadığı söylenebilir.



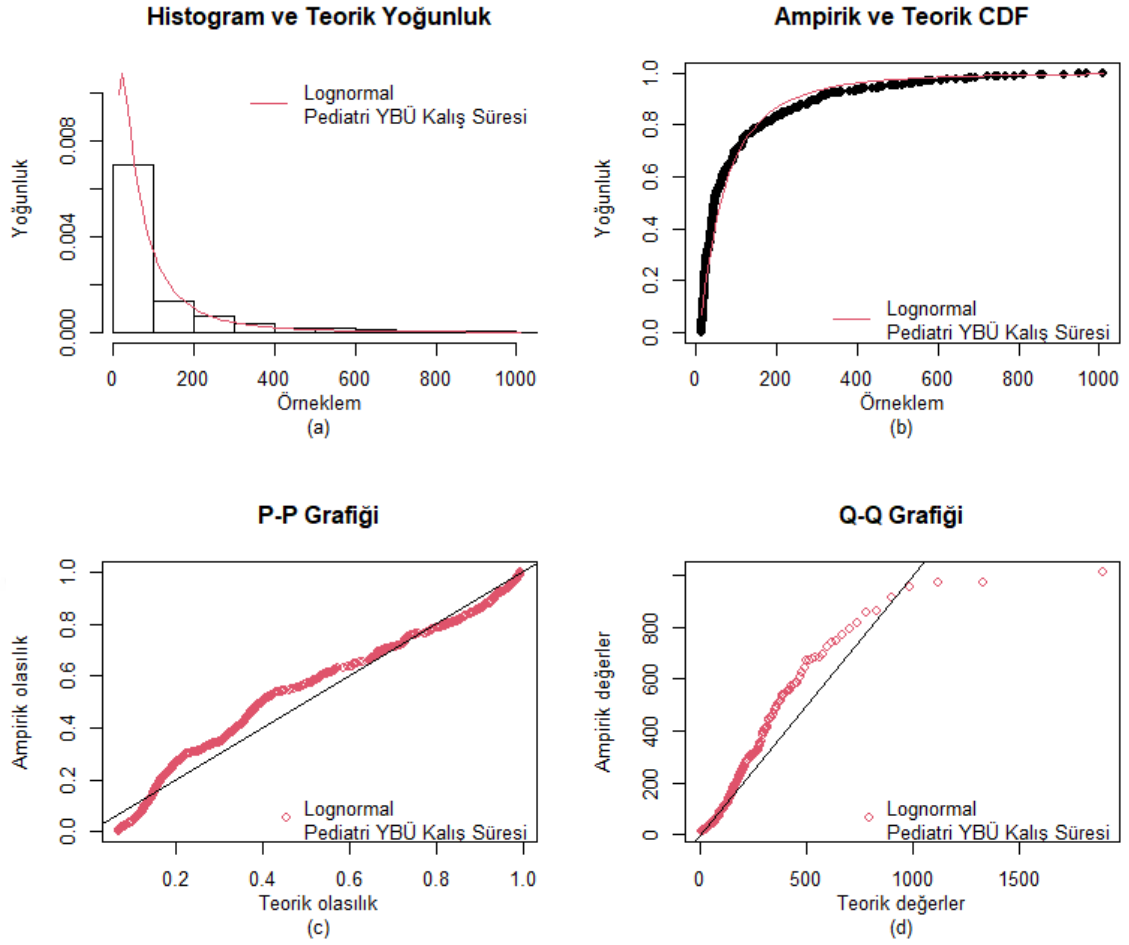
Şekil 4.12. Pediatri YBÜ’deki SS Verisinin Weibull Dağılım Uydurma Grafikleri

Şekil 4.12.’de, Pediatri YBÜ’deki SS’nin Weibull dağılıma olan uyumu izlenmek için grafikler çizilmiştir. Şekil 4.12. (a)’da, Weibull dağılım teorik OYF’si ile gözlenen Pediatri YBÜ’de SS verisi histogramının uyumlu olmadığı görülür. Pediatri YBÜ’deki SS gözlemlerin KDF’si ile Weibull dağılım KDF’nin karşılaştırıldığı Şekil 4.12. (b)’de, seçilen aday Weibull dağılımın veriye iyi bir uyum sağlamadığı tespit edilir. Teorik Weibull dağılım olasılık değerinin Pediatri YBÜ’deki SS olasılık değeriyle kıyaslandığı ve Şekil 4.12. (c)’de paylaşılan P-P grafiğinde, veri çizgi üzerinde toplanmamış ve salınım yapmıştır. Böylece gözlem değerlerinin Weibull dağılımına uymadığı yorumu yapılabilir. Şekil 4.12. (d) ise, gözlem değerlerinin, seçilen Weibull dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir ve seçilen Weibull dağılımı takip etmediği yorumu kolaylıkla yapılabilir. Genel olarak Şekil 4.12. değerlendirildiğinde, analiz edilen verinin aday dağılım olarak seçilen Weibull dağılıma uymadığı kabul edilir.



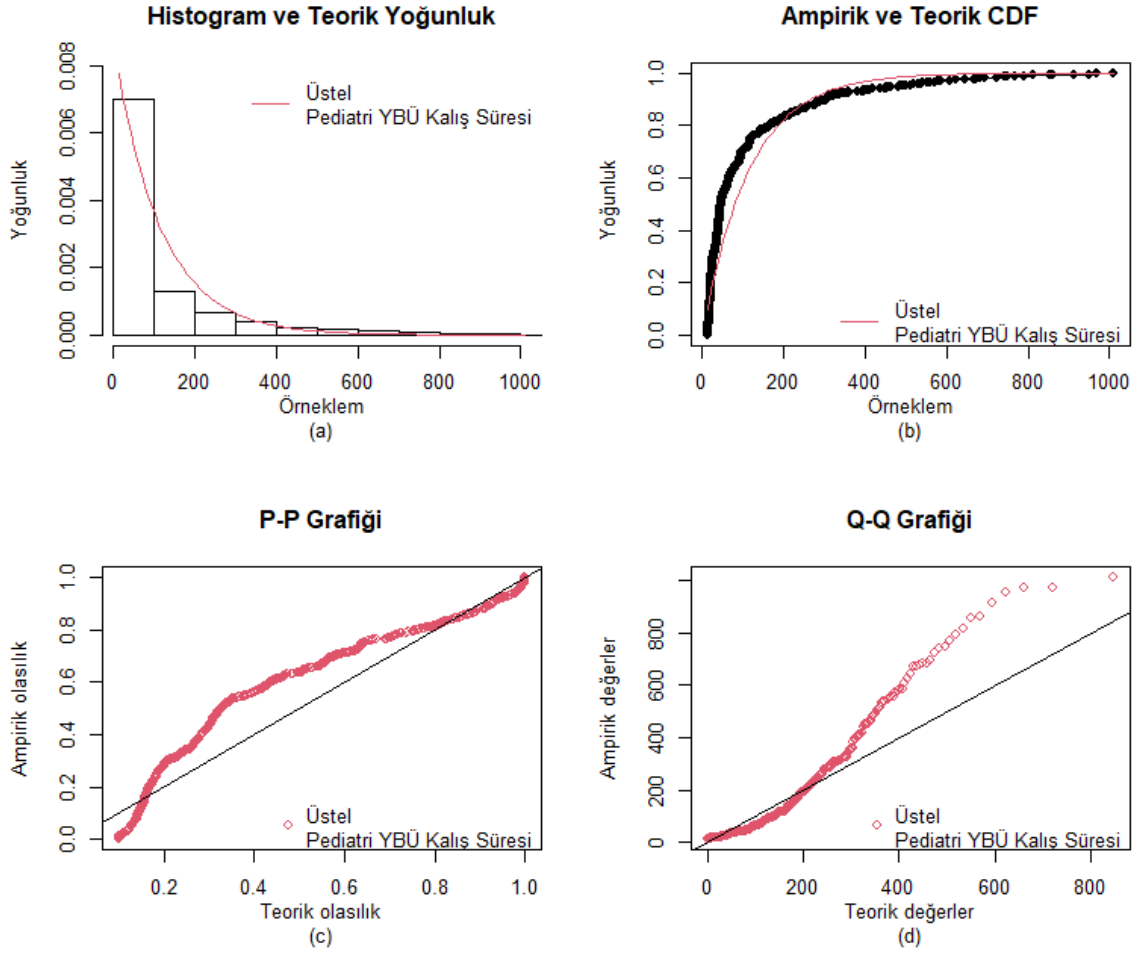
Şekil 4.13. Pediatri YBÜ’deki SS Verisinin Gamma Dağılım Uydurma Grafikleri

Şekil 4.13.’te, Pediatri YBÜ’deki SS gözlemlerinin gamma dağılıma olan uyumu izlenmiştir. Şekil 4.13. (a)’daki, aday gamma dağılımın OYF’sinin ile gözlenen verilere uymadığı görülür. Şekil 4.13. (b)’de, gözlemlerin KDF’si ile aday gamma dağılımın KDF’si kıyaslanmış ve uyumlu olmadığı tespit edilmiştir. Şekil 4.13. (c), teorik gamma dağılım olasılık değerinin Pediatri YBÜ’deki SS verisinin olasılık değeriyle kıyaslandığı P-P grafiğidir. Bu grafik incelendiğinde, gözlemler çizgi etrafında salınım yapmaktadır. Böylece verinin, seçilen gamma dağılımına uymadığı yorumu yapılabilir. Şekil 4.13. (d)’deki QQ grafiğinde ise gözlemlerin grafikte yaptığı salınımlar sebebiyle, seçilen gamma dağılımının veri temsiline uygun olmadığını söylenebilir. Genel olarak Şekil 4.13. değerlendirildiğinde, ele alınan Pediatri YBÜ’deki SS gözlemlerinin, aday dağılım olarak seçilen gamma dağılıma uyum göstermediği söylenebilir.



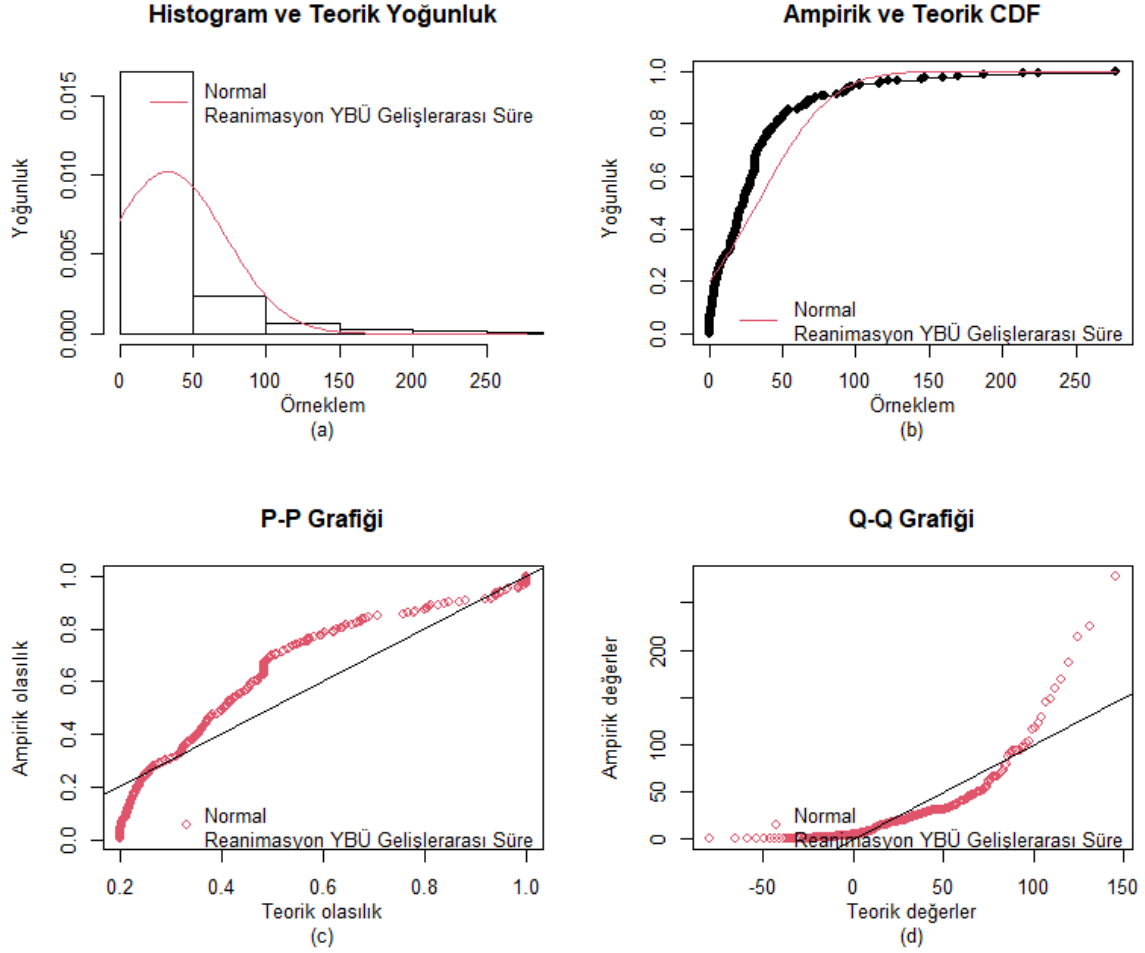
Şekil 4.14. Pediatri YBÜ’deki SS Verisinin Lognormal Dağılım Uydurma Grafikleri

Şekil 4.14.’te, Pediatri YBÜ’deki SS verisinin lognormal dağılıma olan uyumu test etmek için grafikler çizilmiştir. Şekil 4.14. (a), lognormal dağılım teorik OYF’si ile gözlemlerin histogramları kıyaslanmıştır ve birbirine uyum göstermediği yorumu yapılmıştır. Şekil 4.14. (b)’de çizilen, gözlemlerin KDF’si ile aday seçilen lognormal dağılımın KDF’si de sapmalara sahip olması nedeniyle birbirine uymadığı söylenebilir. Şekil 4.14. (c)’deki P-P grafiğinde, teorik lognormal dağılım olasılık değerinin gözlemlerin olasılık değeriyle kıyaslanır. Gözlemlerin, lognormal dağılımı destekleyen çizgiden sapmıştır ve bu dağılıma bir uyum göstermemektedir. Şekil 4.14. (d) ise, gözlemler, seçilen lognormal dağılımdan gelen değerler ile kıyaslanmış ve Q-Q grafiği çizilmiştir. Bu grafiğe göre veri, oldukça salınım göstermektedir ve seçilen lognormal dağılıma kesinlikle bir uyum sağlamadığı yorumu rahatlıkla yapılabilir. Genel olarak Şekil 4.14. değerlendirildiğinde, analiz edilen Pediatri YBÜ’deki SS verisinin, aday dağılım olarak seçilen lognormal dağılıma istatistiksel olarak uyum göstermediği söylenebilir.



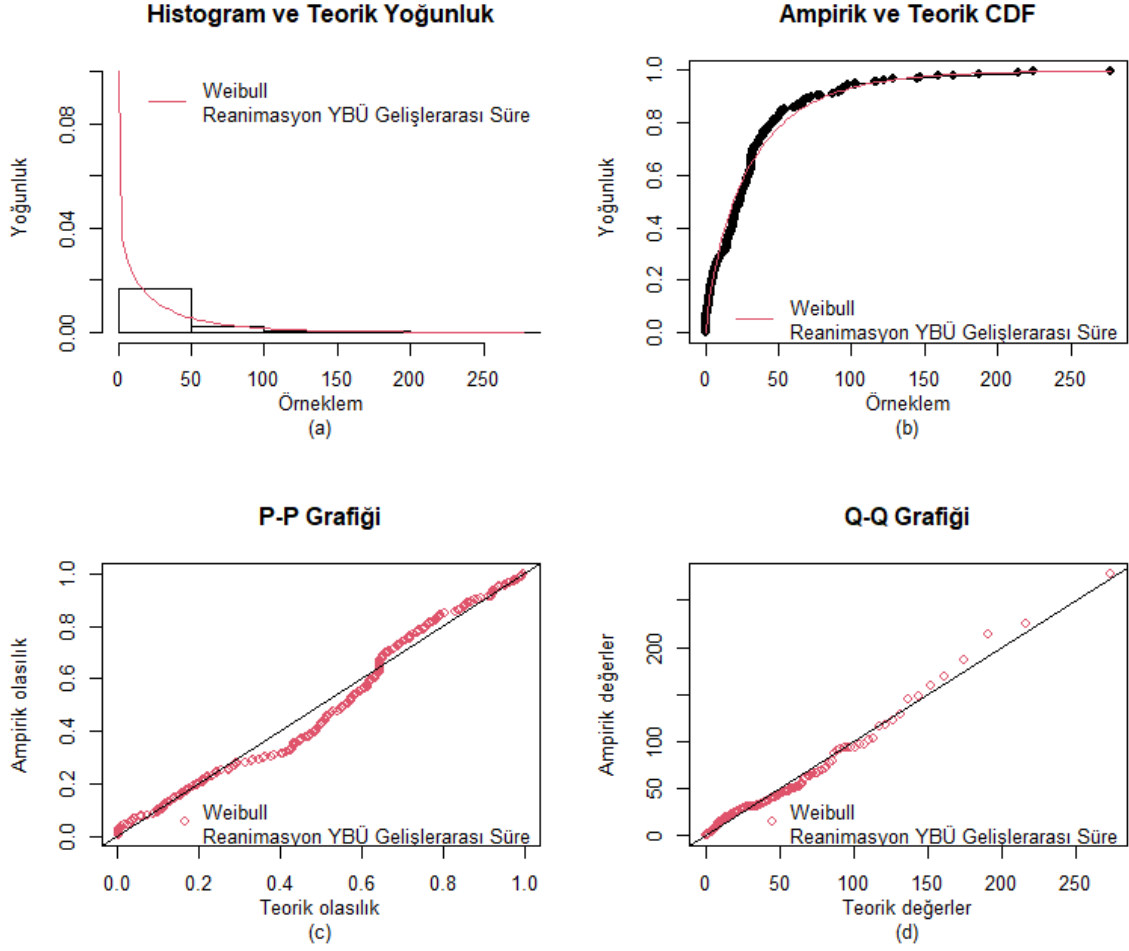
Şekil 4.15. Pediatri YBÜ’deki SS Verisinin Üstel Dağılım Uydurma Grafikleri

Şekil 4.15.’te, Pediatri YBÜ’deki SS gözlemlerinin üstel dağılıma olan uyumu izlenmiştir. Şekil 4.15. (a) incelendiğinde, üstel dağılım teorik OYF’si ile gözlenen değerlerin histogramına uyum göstermemektedir. Şekil 4.15. (b)’de, Pediatri YBÜ’deki SS gözlemlerinin KDF’si ile üstel dağılımın KDF’si karşılaştırılmıştır ve dağılım olan üstel dağılımın gözlemlere uyum sağlayamamıştır. Şekil 4.15. (c)’de paylaşılan P-P grafiğinden de yola çıkarak, verilerin aday üstel dağılıma uymadığı yorumu yapılabilir. Benzer şekilde Şekil 4.15. (d)’de çizilen Q-Q grafiğinde, gözlem değerleri, aday dağılımın beklenen değerlerinden sapmıştır. Genel olarak Şekil 4.15. değerlendirildiğinde, Pediatri YBÜ’deki SS gözlemlerinin, aday dağılım olarak seçilen üstel dağılıma uyum göstermediği söylenebilir.



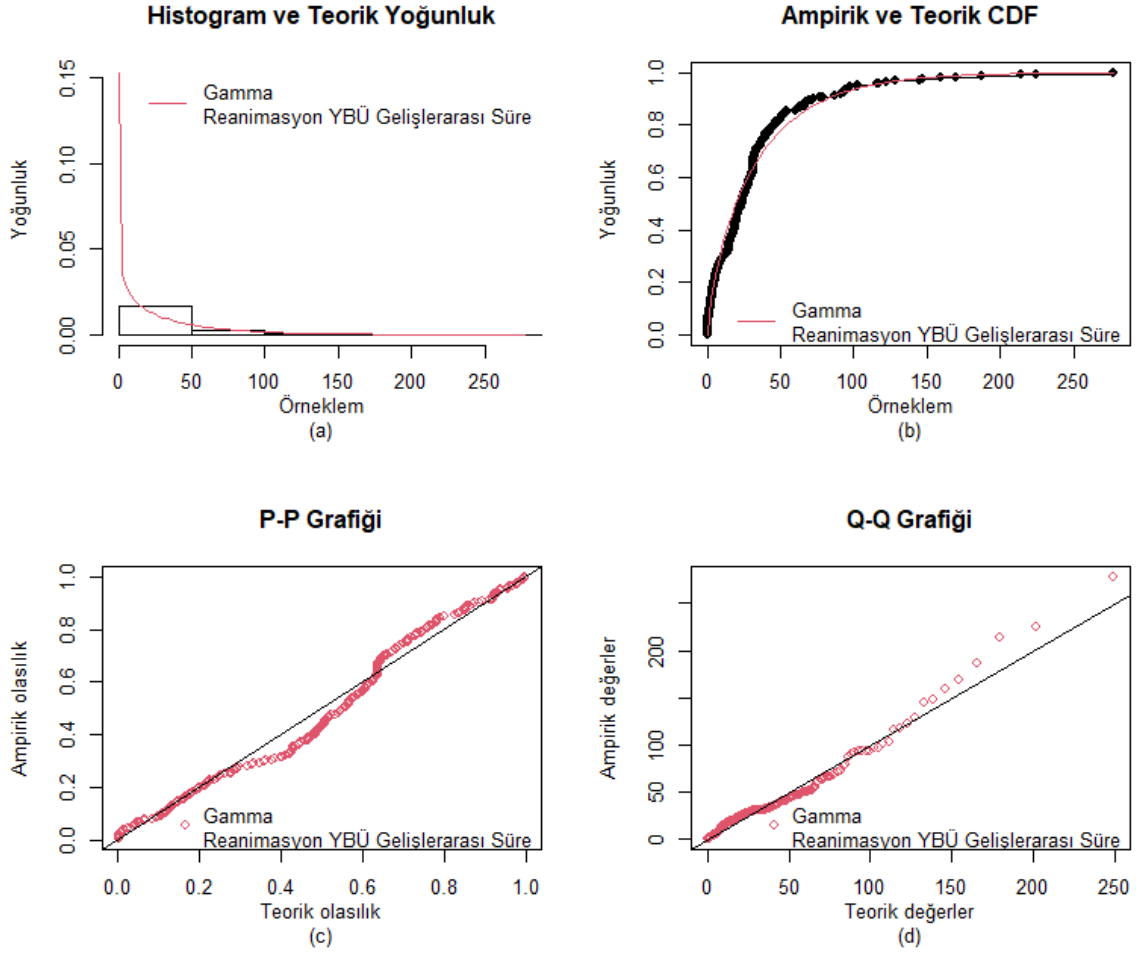
Şekil 4.16. Reanimasyon YBÜ’ye GAS Verisinin Normal Dağılım Uydurma Grafikleri

Şekil 4.16.’da, Reanimasyon YBÜ’ye GAS gözlemlerinin normal dağılıma uyup uymadığını test etmek için çizilen grafikler verilmiştir. Şekil 4.16. (a)’da, gözlemlerin normal dağılıma uymayan bir yayılım gösterdiği ilk bakışta söylenebilir. Şekil 4.16. (b)’de gözlemlerin KDF’si ile aday dağılımın KDF’si kıyaslanmış ve uyum göstermeyen bir grafik ortaya çıkmıştır. Şekil 4.16. (c)’de çizilen P-P grafiğinde de gözlemler aday dağılımı takip etmemektedir ve lineer çizgi üzerine konumlanmamıştır. Şekil 4.16. (d) ise, Reanimasyon YBÜ’ye GAS verisinin, seçilen normal dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiğidir. Grafik incelendiğinde, Q-Q lineer çizgisi etrafında toplanmadığı ve seçilen normal dağılımı uymadığı görülmektedir. Genel olarak Şekil 4.16. değerlendirildiğinde, analiz edilen verinin, aday dağılım olarak seçilen normal dağılıma uygun olmadığı yorumu yapılabilir.



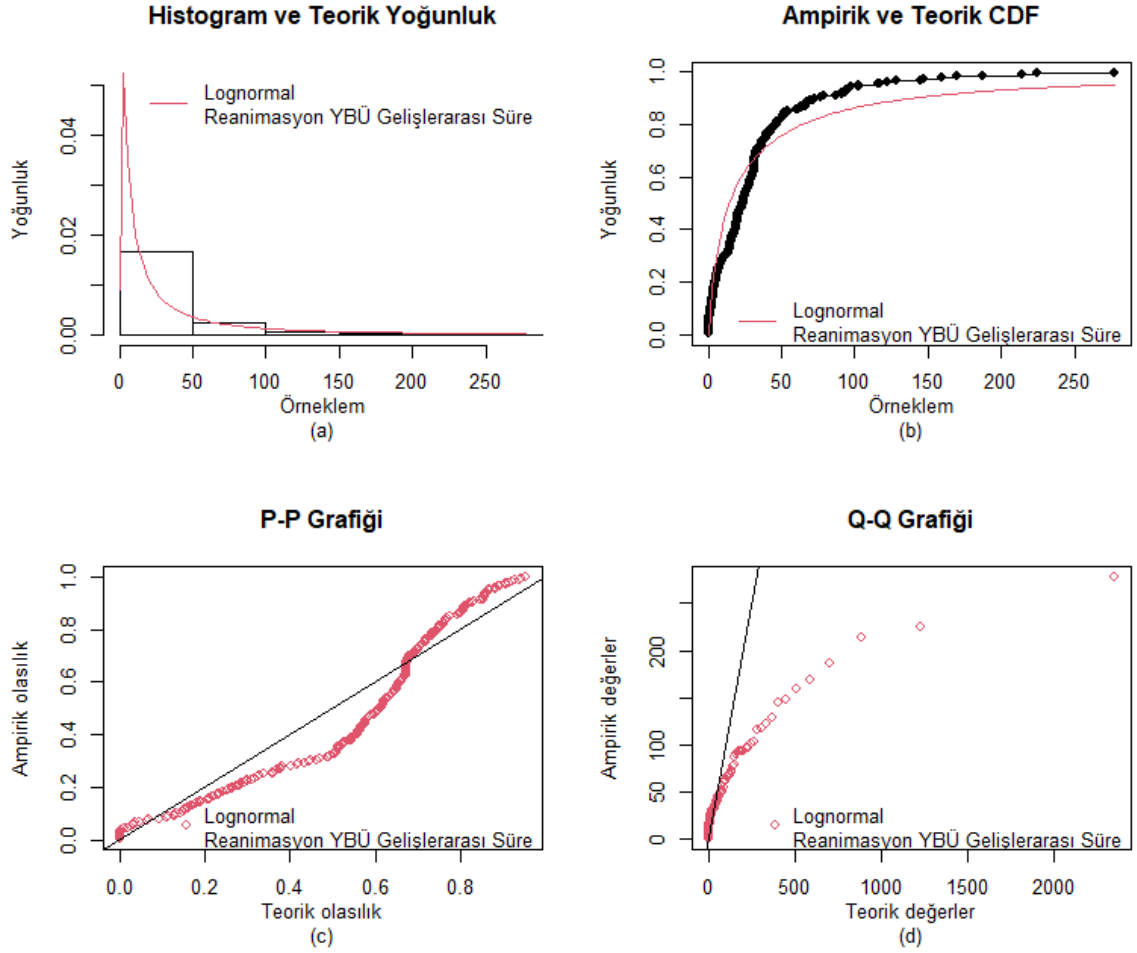
Şekil 4.17. Reanimasyon YBÜ’ye GAS Verisinin Weibull Dağılım Uydurma Grafikleri

Şekil 4.17.’de, Reanimasyon YBÜ’ye gelişler arası süresinin Weibull dağılıma olan uyumu izlenmek için grafikler çizilmiştir. Şekil 4.17. (a)’da çizilen Weibull dağılımının OYF’si ile gözlemlerin histogramına bakıldığında, değerlerin Weibull dağılıma uymadığı görülmektedir. Gözlemlerin KDF’si ile aday dağılım olan Weibull’un beklenen KDF’si kıyaslanmış Şekil 4.17. (b)’de çizilmiştir. Bu grafiğe göre, seçilen aday Weibull dağılımın veriye iyi bir uyum sağlamadığı yorumu yapılabilir. Teorik Weibull dağılım olasılık değerinin, gözlemlerin olasılık değeriyle kıyaslandığı ve Şekil 4.17. (c)’de paylaşılan, P-P grafiğinde çizgi üzerinde veri toplanmamış ve salınım yapmıştır. Böylece gözlemlerin Weibull dağılımına uymadığı söylenebilir. Reanimasyon YBÜ’ye GAS değerlerinin, seçilen Weibull dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiği Şekil 4.17. (d)’de çizilmiş ve P-P grafiğine benzer şekilde, gözlemlerin Q-Q lineer çizgisi etrafında toplanmadığı ve seçilen Weibull dağılımı takip etmediği görülmüştür. Genel olarak Şekil 4.17. değerlendirildiğinde, analiz edilen Reanimasyon YBÜ’ye GAS verisinin aday dağılım olarak seçilen Weibull dağılıma uymaz.



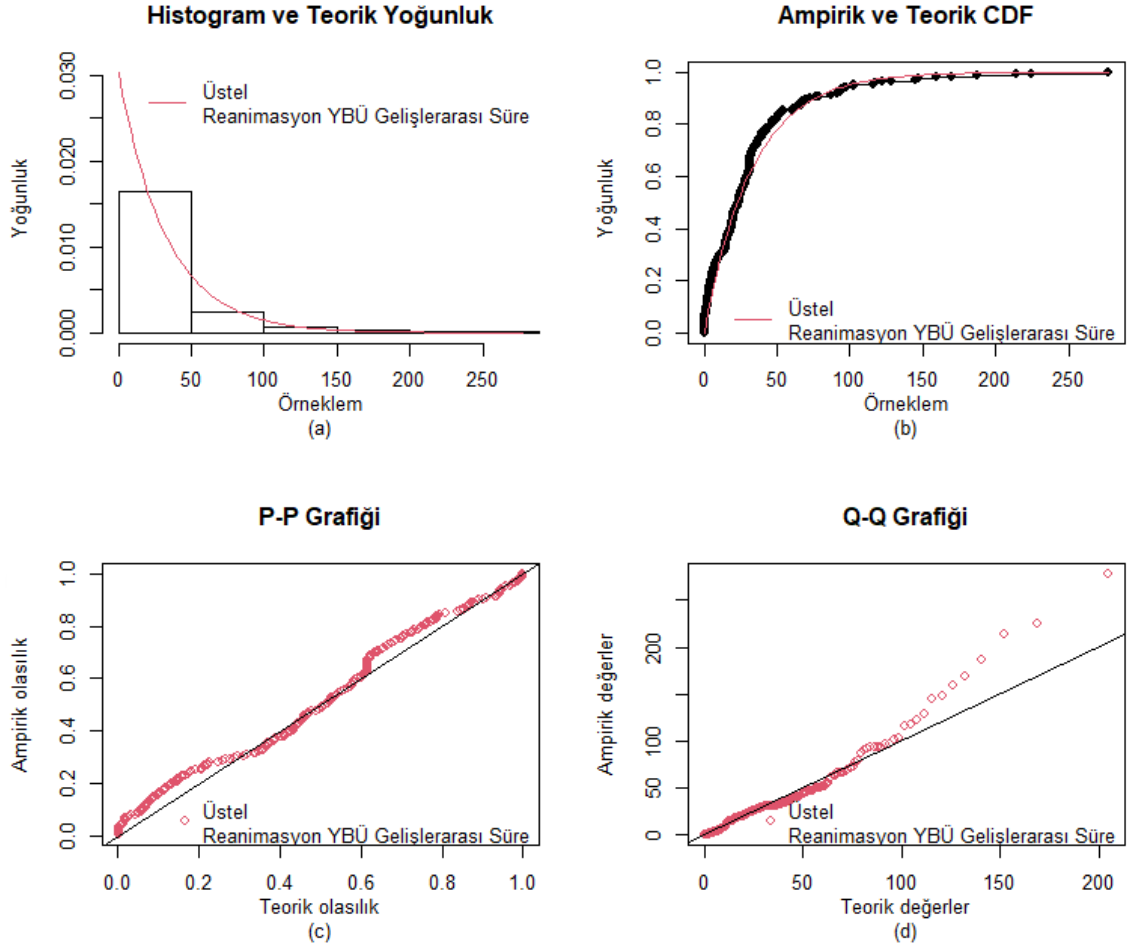
Şekil 4.18. Reanimasyon YBÜ’ye GAS Verisinin Gamma Dağılım Uydurma Grafikleri

Şekil 4.18.’de, Reanimasyon YBÜ’ye GAS verisinin gamma dağılımına olan uyumu izlenmiştir. Şekil 4.18. (a)’da, gamma dağılım teorik OYF’si ile gözlenen verinin histogramının uyumsuz olduğu görülmektedir. Şekil 4.18. (b) grafiğinde, gözlemlerin KDF’si ile gamma dağılım KDF’si karşılaştırılmıştır ve verinin seçilen aday dağılım olan gamma dağılıma tam olarak uyum sağlamadığı yorumu yapılabilir. Şekil 4.18. (c), teorik gamma dağılım olasılık değerinin gözlemlerin olasılık değeriyle kıyaslandığı P-P grafiğidir. Bu grafik incelendiğinde, veri çizgi etrafında salınım yapmaktadır. Böylece verinin, seçilen gamma dağılımına uymadığı yorumunu yapılabilir. Şekil 4.18. (d) ise, gözlemleri, seçilen gamma dağılımdan gelen değerler ile karşılaştırılan Q-Q grafiğidir. Q-Q grafiğinde, P-P grafiğinde olduğu gibi, verilere çizgi etrafında toplanmamıştır. Bu durum, gözlemlerin seçilen gamma dağılımının uygun olmadığını gösterir. Genel olarak Şekil 4.18. değerlendirildiğinde, ele alınan Reanimasyon YBÜ’ye GAS verisinin, aday dağılım olarak seçilen gamma dağılımına uymadığı göstermediği söylenebilir.



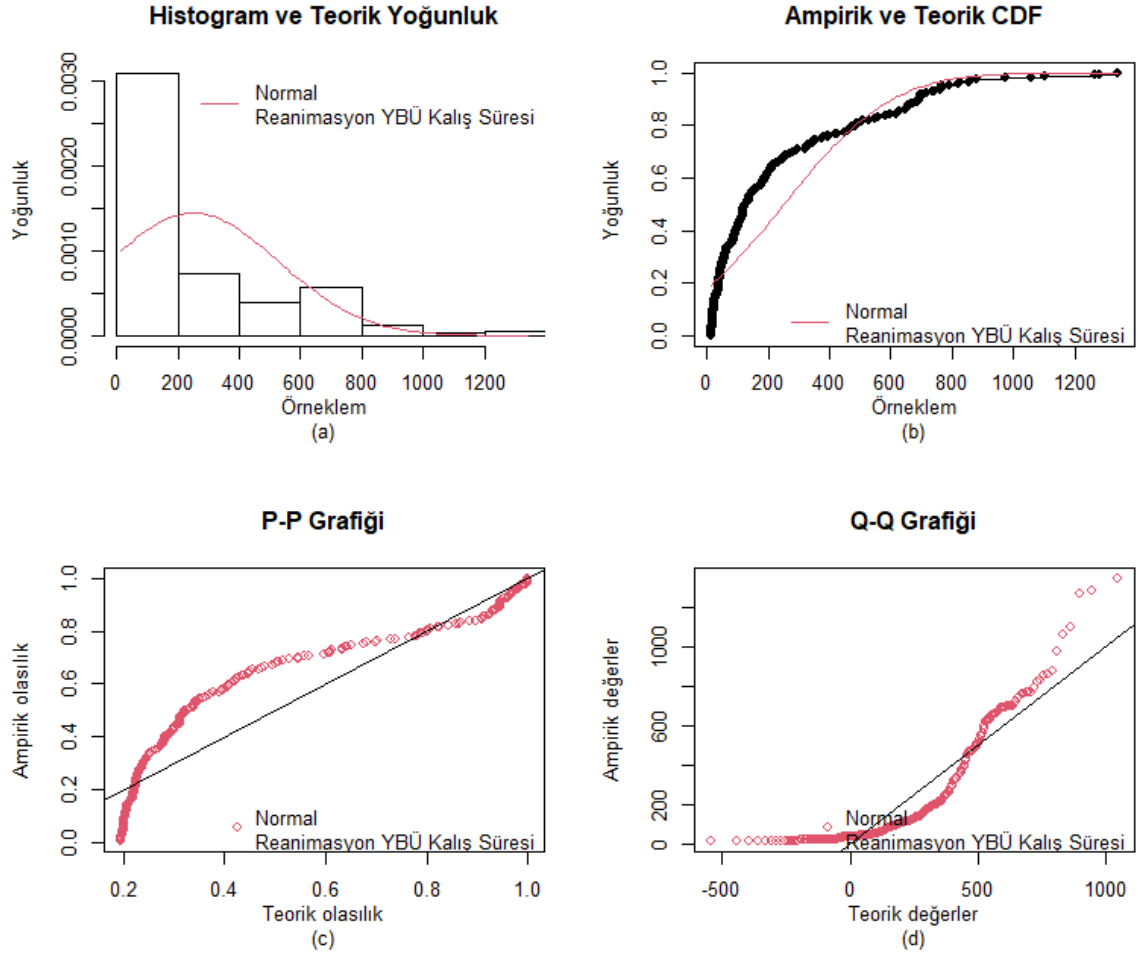
Şekil 4.19. Reanimasyon YBÜ'ye GAS Verisinin Lognormal Dağılım Uydurma Grafikleri

Şekil 4.19.'da, Reanimasyon YBÜ'ye GAS verisinin lognormal dağılıma olan uyumu test edilmek için grafikler çizilmiştir. Şekil 4.19. (a), lognormal dağılım teorik OYF'si ile gözlenen değerlerin histogramıdır ve gözlenen değerlerinin lognormal dağılıma uymadığı görülür. Şekil 4.19. (b), gözlenen verinin KDF'si ile lognormal dağılım KDF'sidir ve buna göre verinin, seçilen aday lognormal dağılımını takip etmediği söylenebilir. Şekil 4.19. (c), teorik lognormal dağılım olasılık değerinin, gözlemlerin olasılık değeriyle kıyaslandığı P-P grafiğidir. Gözlemler, lognormal dağılımı destekleyen çizgiden sapmıştır ve uyum göstermemektedir. Şekil 4.19. (d) ise, Reanimasyon YBÜ'ye GAS değerlerinin, seçilen lognormal dağılımdan gelen değerler ile karşılaştırıldığı Q-Q grafiğidir. Veri, bu grafikte oldukça çarpık bir salınım göstermektedir ve seçilen dağılıma kesinlikle bir uyum sağlamaz. Genel olarak Şekil 4.19. değerlendirildiğinde, analiz edilen Reanimasyon YBÜ'ye GAS verisinin, aday dağılım olarak seçilen lognormal dağılıma istatistiksel olarak uyum göstermediği söylenebilir.



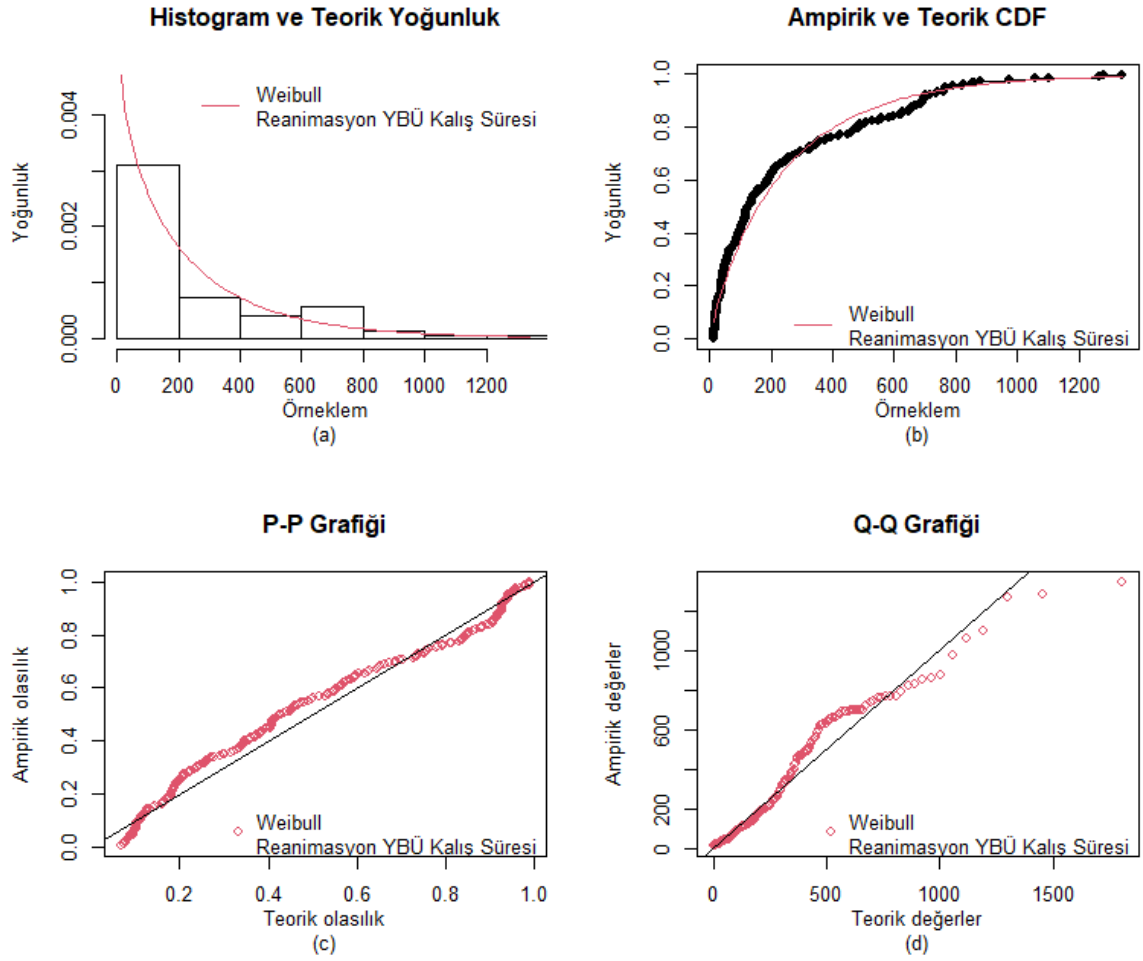
Şekil 4.20. Reanimasyon YBÜ'ye GAS Verisinin Üstel Dağılım Uydurma Grafikleri

Şekil 4.20.'de, Reanimasyon YBÜ'ye GAS verisinin üstel dağılıma olan uyumu izlenmiştir. Şekil 4.20. (a), üstel dağılım teorik OYF'si ile gözlenen değerlerinin histogramının uyum göstermediği söylenebilir. Şekil 4.20. (b), Reanimasyon YBÜ'ye GAS verisinin KDF'si ile üstel dağılımın KDF'si karşılaştırılmıştır ve gözlemlerin adaya dağılıma tam olarak uyum sağlamadığı görülmüştür. Üstel dağılım olasılık değerinin, gözlemlerin olasılık değeriyle karşılaştırıldığı ve Şekil 4.20. (c)'de çizilen P-P grafiğine göre, gözlemler çizgi etrafında salınım yapmaktadır. Böylece gözlemlerin seçilen üstel dağılımına yorumu yapılır. Şekil 4.20. (d) ise, Reanimasyon YBÜ'ye GAS değerlerinin, seçilen üstel dağılımdan gelen değerler ile karşılaştırılan Q-Q grafiğidir. Q-Q grafiğini incelendiğinde, seçilen aday üstel dağılıma, Reanimasyon YBÜ'ye GAS verisinin uygun olmadığını görülmektedir. Genel olarak Şekil 4.20. değerlendirildiğinde, ele alınan Reanimasyon YBÜ'ye GAS verisinin, aday dağılım olarak seçilen üstel dağılıma uyum göstermediği ifade edilebilir.



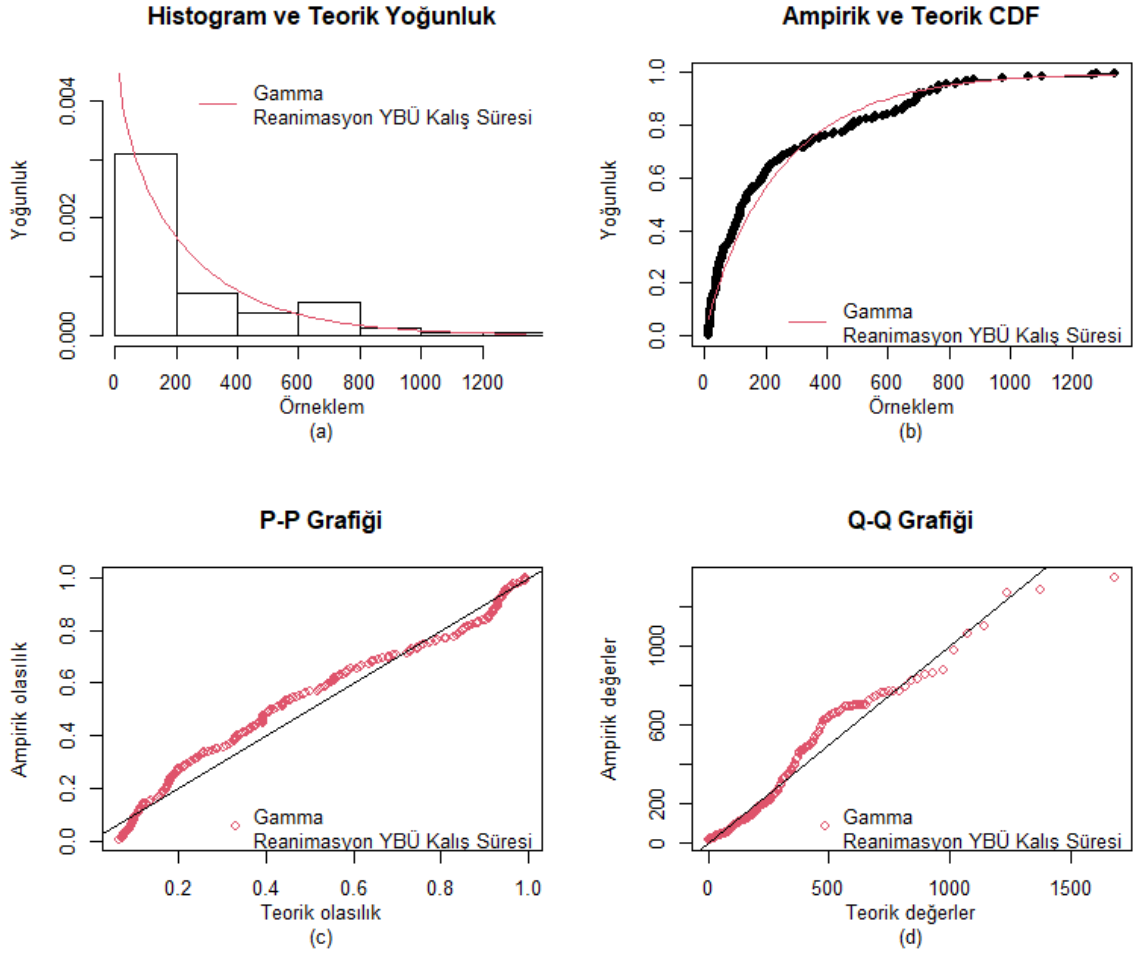
Şekil 4.21. Reanimasyon YBÜ’deki SS Verisinin Normal Dağılım Uydurma Grafikleri

Şekil 4.21.’de, Reanimasyon YBÜ’deki SS verisinin normal dağılıma uyup uymadığını test etmek için çizilen grafikler verilmiştir. Normal dağılım teorik OYF’si ile gözlenen değerlerin histogramının çizildiği Şekil 4.21. (a)’ya göre, gözlemler normal dağılıma uymaz. Şekil 4.21. (b), Reanimasyon YBÜ’deki SS verisinin KDF’si ile normal dağılım KDF’si karşılaştırılmış ve uyumsuz olduğu görülmüştür. Teorik normal dağılım olasılık değerinin, gözlemlerin olasılık değeriyle kıyaslandığı P-P grafiği Şekil 4.21. (c)’de paylaşılmış ve Reanimasyon YBÜ’deki SS verisini temsil eden noktalar çizgi üzerinden yüksek oranda sapmış olması nedeniyle normal dağılıma uymadığı yorumu yapılmıştır. Şekil 4.21. (d) ise, Reanimasyon YBÜ’deki SS verisinin, seçilen normal dağılımdan gelen değerler ile kıyaslandığı Q-Q grafiği çizilmiş ve buna göre gözlemlerin lineer çizgisi etrafında toplanmadığı ve seçilen normal dağılımı uymadığı görülmüştür. Genel olarak Şekil 4.21. değerlendirildiğinde, analiz edilen Reanimasyon YBÜ’deki SS verisinin, aday dağılım olarak seçilen normal dağılıma uygun olmadığı yorumu rahatlıkla yapılabilir.



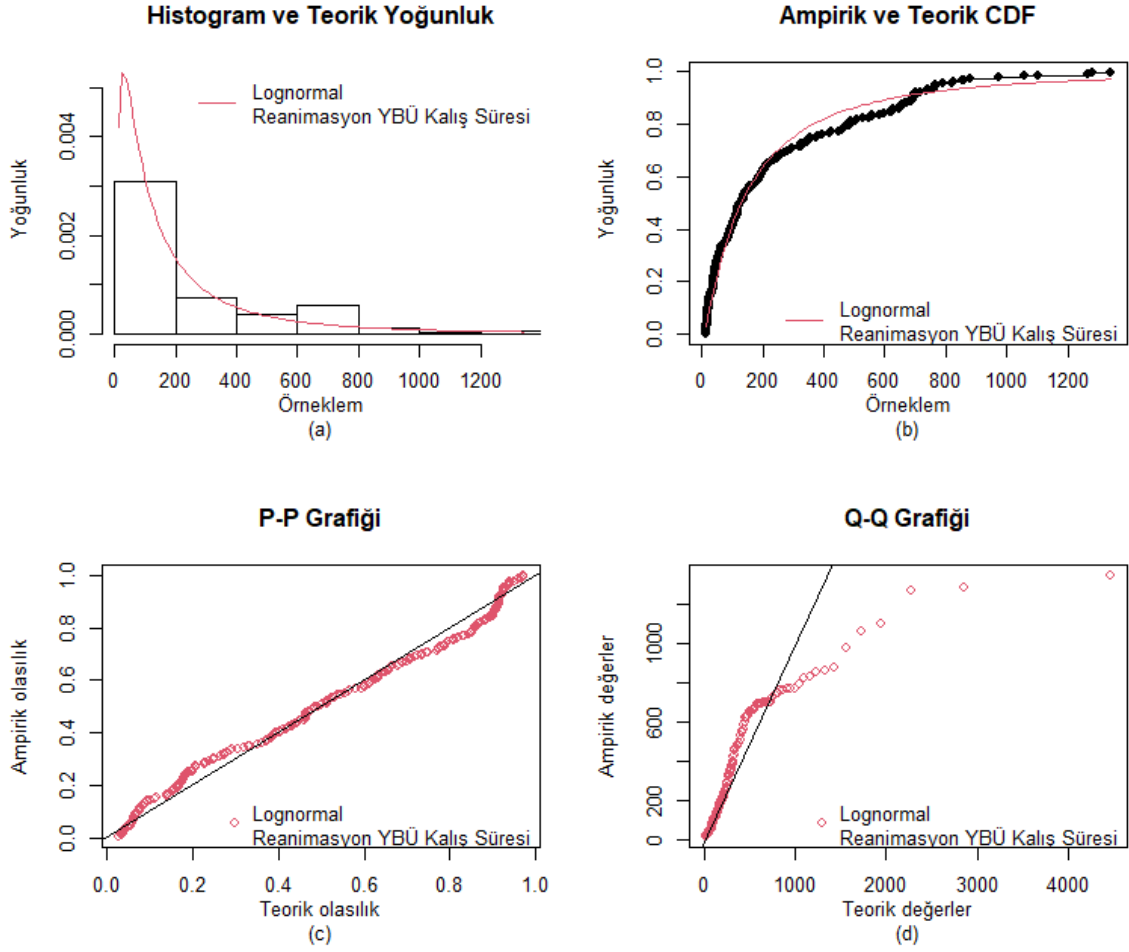
Şekil 4.22. Reanimasyon YBÜ’deki SS Verisinin Weibull Dağılım Uydurma Grafikleri

Şekil 4.22.’de, Reanimasyon YBÜ’deki SS verisinin Weibull dağılıma olan uyumu izlemek için grafikler çizilmiştir. Şekil 4.22. (a), Weibull dağılım teorik OYF’si ile gözlenen Reanimasyon YBÜ’deki SS verisinin histogramıdır ve Weibull dağılıma uymadığı görülmektedir. Şekil 4.22. (b)’de paylaşılan ve Reanimasyon YBÜ’deki SS verisinin KDF’si ile Weibull dağılım KDF’nin karşılaştırıldığı grafikte, gözlemlerin seçilen aday Weibull dağılımın iyi bir uyum sağlamadığı yorumu yapılabilir. Teorik Weibull dağılım olasılık değerinin Reanimasyon YBÜ’deki SS verisinin olasılık değeriyle kıyaslandığı ve Şekil 4.22. (c)’de çizilen P-P grafiğinde, veri çizgi üzerinde toplanmamış ve salınım yapmıştır. Böylece gözlemlerin Weibull dağılımına uymadığı söylenebilir. Reanimasyon YBÜ’deki SS değerlerinin, seçilen Weibull dağılımdan gelen değerler ile kıyaslandığı ve Şekil 4.22. (d)’de paylaşılan Q-Q grafiğinde, değerlerin seçilen Weibull dağılımı takip etmediği yorumu kolaylıkla yapılabilir. Genel olarak Şekil 4.22. değerlendirildiğinde, analiz edilen Reanimasyon YBÜ’deki SS verisinin aday dağılım olarak seçilen Weibull dağılıma uymadığı yorumu yapılabilir.



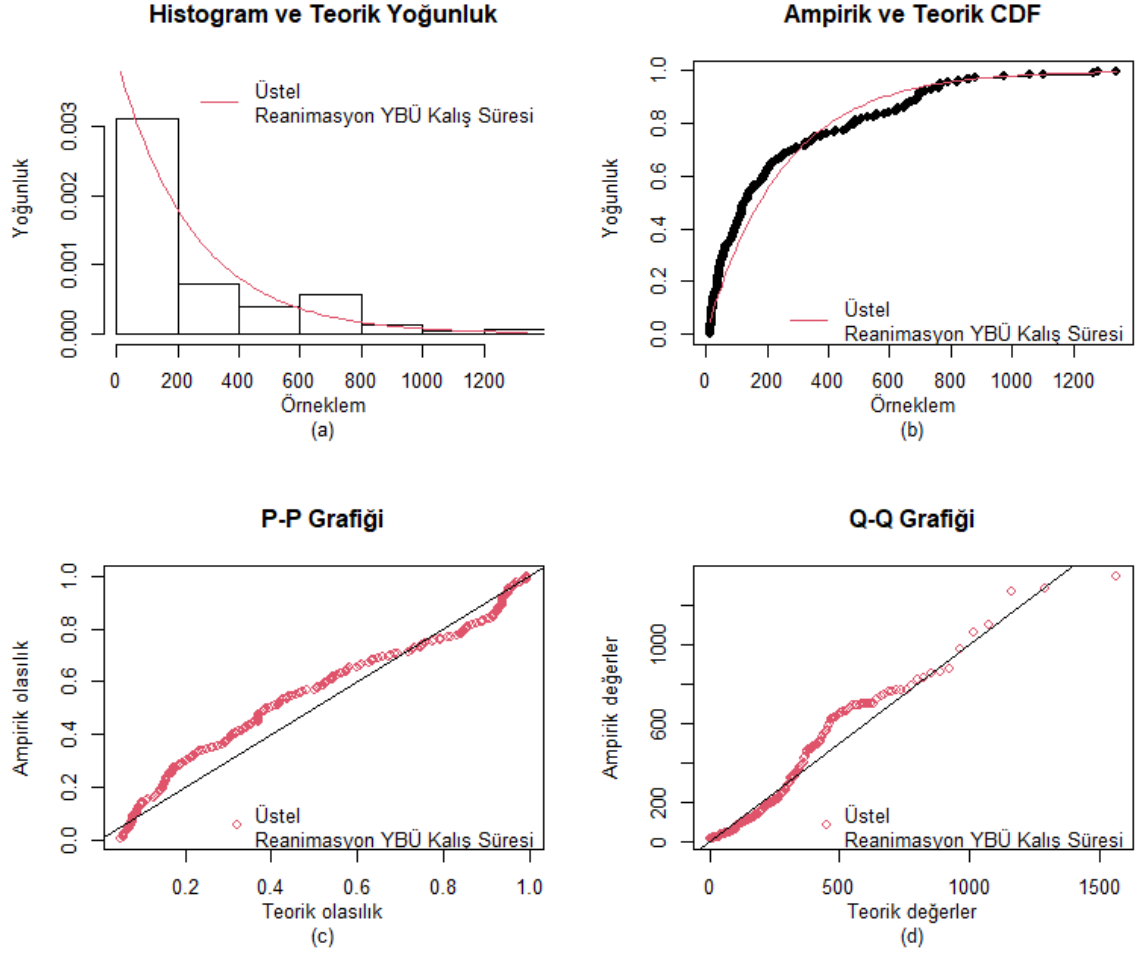
Şekil 4.23. Reanimasyon YBÜ’deki SS Verisinin Gamma Dağılım Uydurma Grafikleri

Şekil 4.23.’te, Reanimasyon YBÜ’deki SS verisinin gamma dağılımına olan uyumu test edilmiştir. Şekil 4.23. (a) grafiğinde, gamma dağılım teorik OYF’si ile gözlemlerin oluşturduğu histogramının uyumsuz olduğu görülmektedir. Şekil 4.23. (b) grafiğinde, Reanimasyon YBÜ’deki SS verisinin KDF’si ile gamma dağılım KDF’si karşılaştırılmıştır ve gözlemlerin seçilen aday dağılım olan gamma dağılımına uyum sağlayamadığı tespit edilmiştir. Şekil 4.23. (c), teorik gamma dağılım olasılık değerinin Reanimasyon YBÜ’deki SS verisinin olasılık değeriyle kıyaslandığı P-P grafiğidir. Bu grafik incelendiğinde, gözlemler çizgi etrafında salınım yapmaktadır. Böylece verinin, seçilen gamma dağılımına uymadığı yorumu yapılabilir. Şekil 4.23. (d) ise, gözlemlerin, seçilen gamma dağılımdan gelen değerler ile karşılaştırıldığı Q-Q grafiğidir. Bu grafikte, verinin gamma dağılımının uygun olmadığını yorumu yapılabilir. Genel olarak Şekil 4.23. değerlendirildiğinde, ele alınan Reanimasyon YBÜ’deki SS verisinin, aday dağılım olarak seçilen gamma dağılımla temsil edilemeyeceği söylenebilir.



Şekil 4.24. Reanimasyon YBÜ’deki SS Verisinin Lognormal Dağılım Uydurma Grafikleri

Şekil 4.24.’te, Reanimasyon YBÜ’deki SS verisinin lognormal dağılıma olan uyumu test edilmek için grafikler çizilmiştir. Şekil 4.24. (a), lognormal dağılım teorik OYF’si ile gözlenen Reanimasyon YBÜ’deki SS değerlerinin histogramı ve bu grafikte gözlemlerin lognormal dağılıma uymadığı görülmektedir. Şekil 4.24. (b), Reanimasyon YBÜ’deki SS verisinin KDF’si ile lognormal dağılım KDF’sidir ve buna göre gözlemler, seçilen aday lognormal dağılımıyla uyumlu değildir. Şekil 4.24. (c), teorik lognormal dağılım olasılık değerinin, gözlemlerin olasılık değeriyle kıyaslandığı P-P grafiğidir ve bu grafikteki sapmalar verinin lognormal dağılıma uyum göstermediğini destekler. Şekil 4.24. (d) ise, Reanimasyon YBÜ’deki SS değerlerinin, seçilen lognormal dağılımdan gelen değerler ile karşılaştırıldığı Q-Q grafiğidir. Veri, bu grafikte oldukça çarpık bir salınım göstermektedir ve seçilen aday dağılım olan lognormal dağılıma kesinlikle bir uyum sağlamaz. Genel olarak Şekil 4.24. değerlendirildiğinde, analiz edilen Reanimasyon YBÜ’deki SS verisinin, aday dağılım olarak seçilen lognormal dağılıma istatistiksel olarak uyum göstermediği ifade edilebilir.



Şekil 4.25. Reanimasyon YBÜ’deki SS Verisinin Üstel Dağılım Uydurma Grafikleri

Şekil 4.25.’te, Reanimasyon YBÜ’deki SS verisinin üstel dağılıma olan uyumu izlenmiştir. Şekil 4.25. (a)’da çizilen, üstel dağılım teorik OYF’si ile Reanimasyon YBÜ’deki SS değerleri histogramının uyum göstermediği söylenebilir. Şekil 4.25. (b)’de, gözlemlerin KDF’siyle üstel dağılımın KDF’si karşılaştırılmıştır ve uyum gözlenememiştir. Üstel dağılım olasılık değerinin Reanimasyon YBÜ’deki SS olasılık değeriyle karşılaştırıldığı P-P grafiği, Şekil 4.25. (c)’de çizilmiş ve verinin çizgi etrafında salınım yaptığı görülmüştür. Böylece gözlemlerin seçilen üstel dağılımına uymadığını söylenebilir. Şekil 4.25. (d) ise, Reanimasyon YBÜ’deki SS değerlerinin, seçilen üstel dağılımdan gelen değerler ile karşılaştırılan Q-Q grafiğidir. Bu grafik incelendiğinde, seçilen aday üstel dağılıma, gözlemlerin uygun olmadığını yorumu yapılabilir. Genel olarak Şekil 4.25. değerlendirildiğinde, ele alınan Reanimasyon YBÜ’deki SS verisinin, aday dağılım olarak seçilen üstel dağılıma uyum göstermediği yorumu yapılabilir.

Olasılık dağılım ailesi seçiminde faydalanılan grafiksel yöntemler ayrıntılı bir şekilde uygulanmış ve aday seçilen normal, Weibull, gamma, lognormal ve üstel dağılıma toplanan verilerin hiçbir uyum sağlamadığı tespit edilmiştir. Seçilen bu aday dağılımlar için aynı zamanda uyum iyiliği testleri de uygulanmış ve elde edilen sonuçlar desteklenmiştir.

Pediatric ve Reanimasyon YBÜ'lerin GAS ve SS değerlerine dağılım uydurma için Ki-kare (χ^2), Kolmogorov-Smirnov (K-S) ve Anderson-Darling (A-D) uyum iyiliği testleri uygulanmıştır. Bu adımda uyum iyiliği testleri p -değerleri RStudio programı kullanılarak hesaplanmıştır. H_0 hipotezi, seçilen aday dağılımın veriye uyumlu olduğu şeklinde kurulmuştur. Uyum iyiliği testi için p -değeri anlamlılık düzeyinden küçükse, H_0 hipotezi reddedilir ve verinin seçilen aday dağılıma uymadığı yorumu yapılır. p -değeri, bir test istatistiğinin gözlemlenen veriye ne kadar uyduğunu veya beklenen bir dağılımın ne kadar iyi olduğunu değerlendirmek için kullanılır.

Çizelge 4.2'de bir uyum iyiliği testi olan ki-kare (χ^2) testi her bir YBÜ için farklı dağılımlar (Normal, Weibull, Gamma, Lognormal, Üstel) için yapılan χ^2 uyum iyiliği testinin p -değerleri sunulmuştur.

Çizelge 4.2. Ki-kare (χ^2) Uyum İyiliği Testi p -değerleri

YBÜ	Tür	Normal	Weibull	Gamma	Lognormal	Üstel
Pediatric	GAS	0,00	0,00	0,00	0,00	0,00
	SS	0,00	0,00	0,00	0,00	0,00
Reanimasyon	GAS	0,00	0,00	0,00	0,00	0,04
	SS	0,00	0,00	0,00	0,00	0,00

Pediatric YBÜ GAS için, tüm test edilen aday dağılımlar (Normal, Weibull, Gamma, Lognormal, Üstel) için p -değeri 0,00'dır. Bu, gözlemlerin her bir dağılıma mükemmel bir şekilde uymadığını ve test edilen dağılımların veriye uygun olmadığını gösterir.

Pediatric YBÜ SS için, Pediatric YBÜ GAS'ye benzer şekilde, tüm test edilen dağılımlar için p -değeri 0,00'dır. Böylece aday seçilen dağılımların hiçbirine istatistiksel olarak uyum sağlamamaktadır.

Reanimasyon YBÜ'de hem GAS hem de SS için tüm dağılımlarda p -değeri H_0 hipotezi reddedilir ve gözlemlerin seçilen beş aday dağılımdan (normal, Weibull, gamma, lognormal ve üstel) hiçbirine uymadığını yorumu yapılır.

χ^2 testine ek olarak Kolmogorov-Smirnov (K-S) ve Anderson-Darling (A-D) uyum iyiliği testleri uygulanmış ve gözlemlerin seçilen beş aday dağılıma (normal, Weibull, gamma, lognormal ve üstel) olan uyumu ifade edilmek istenmiştir.

Çizelge 4.3.'te K-S uyum iyiliği testi p -değerleri paylaşılmıştır. χ^2 uyum iyiliği testi sonuçlarına benzer şekilde aday tüm dağılımlar için H_0 hipotezi reddedilmiştir.

Çizelge 4.3. Kolmogorov-Smirnov (K-S) Uyum İyiliği Testi p -değerleri

YBÜ	Tür	Normal	Weibull	Gamma	Lognormal	Üstel
Pediatri	GAS	0.00	0.00	0.00	0.00	0.00
	ST	0.00	0.00	0.00	0.00	0.00
Reanimasyon	GAS	0.00	0.01	0.02	0.00	0.13
	ST	0.00	0.13	0.04	0.14	0.00

Uygulanan son uyum iyiliği testi ise, Anderson-Darling uyum iyiliği testidir. Bu testten elde edilen p -değerleri Çizelge 4.4.'te paylaşılmıştır. χ^2 ve K-S uyum testine benzer şekilde, A-D uyum iyiliği testinde de beş aday dağılım için hesaplanan tüm p -değerleri H_0 hipotezini reddeder ve aday dağılımların hiçbiri Pediatri ve Reanimasyon YBÜ'lerin GAS ve SS verisinin istatistiksel olarak temsil edemediği ortaya konur.

Çizelge 4.4. Anderson-Darling (A-D) Uyum İyiliği Testi p -değerleri

YBÜ	Tür	Normal	Weibull	Gamma	Lognormal	Üstel
Pediatri	GAS	0.00	0.00	0.01	0.01	0.00
	SS	0.00	0.00	0.00	0.00	0.00
Reanimasyon	GAS	0.00	0.05	0.06	0.00	0.01
	SS	0.00	0.02	0.01	0.06	0.00

Hem grafiksel hem de istatistiksel uyum iyiliği testleriyle ayrıntılı bir şekilde analiz edilen Pediatri ve Reanimasyon YBÜ'lerden toplanan GAS ve SS verisinin aday seçilen beş dağılıma (normal, Weibull, gamma, lognormal ve üstel) uymadığı görülmüştür. Ancak bu gözlemlerin simülasyon modelinde girdi olarak kullanılabilmesi için istatistiksel olarak temsil edilmesi gerekmektedir. Bu noktada, klasik ancak özellikle çarpıklığı yüksek olan veriye temsil yeteneği düşük ampirik dağılımlar ya da çalışmamızın temel farklılığı oluşturan karma dağılımlar kullanılmalıdır. Çalışmamızın devamında hastaların Pediatri ve Reanimasyon YBÜ'lere GAS ve bu ünitelerdeki SS karma dağılım ile analiz edildi, bileşen sayısı belirlenerek parametre tahmini yapıldı. Karma dağılımların gerçek verisi temsil yeteneği seçilen performans kriterleri yardımıyla ampirik dağılım ile karşılaştırıldı. Performans kriterlerinin değerlendirilmesi sonucunda, etkinliği yüksek olan yaklaşım seçildi ve seçilen yaklaşımın parametreleri simülasyon modelinde girdi parametresi olarak kullanıldı.

4.2. Karma Dağılım Analizi

Materyal ve Metot ana başlığında bulunan 3.6. Karma Dağılımlar alt başlığında ayrıntılı olarak ele alındığı gibi, bir veri setinin karma dağılım analizine ilk olarak bileşen sayısı (k) belirlemeyle başlanır. Bileşen sayısı (k) belirlerken grafiksel yöntemler, likelihood oran istatistiği (LTR) ve AIC, BIC gibi bilgi kriterleri bütünsel olarak kullanılır.

Hastaların Pediatri ve Reanimasyon YBÜ'lere GAS ve bu ünitelerdeki SS'i temsil edecek en iyi karma dağılımı bulabilmek için bileşen sayısının 2, 3, 4, 5 olduğu ve normal, Weibull,

gamma ve lognormal karma dağılımları olacak şekilde 16 farklı durum test edildi ve logaritmik LTR, AIC ve BIC değerleri hesaplandı. AIC ve BIC'nin en düşük olduğu, logaritmik LTR değerinde ise, azaltmanın artmaya başladığı değer en iyi model uyumunu gösterecek şekilde model seçimi yapıldı. AIC ve BIC değerleri farklı modellerde en küçük değeri aldığı durumlarda, literatür ışığında BIC'nin AIC'den daha doğru bir tahmin sağlaması nedeniyle BIC değeri dikkate alındı (Keribin, 2000). Her iki YBÜ'nün GAS ve SS verisi karma dağılım analizinde kullanılan log LTR, AIC ve BIC değerleri RStudio programında hesaplandı.

Çizelge 4.5. Pediatri YBÜ GAS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri

Karma Dağılım	Bileşen Sayısı (<i>k</i>)	Model İstatistiği		
		log LTR	AIC	BIC
Normal	2	-2635,86	5281,71	5304,866
	3	-2558,12	5132,24	5169,285
	4	-2531,66	5085,32	5136,254
	5	-2526,91	5081,83	5146,658
Weibull	2	-2536,79	5083,57	5106,725
	3	-2505,96	5027,92	5064,963
	4	-2505,42	5032,85	5083,784
	5	-2502,57	5033,15	5088,449
Gamma	2	-2534,65	5079,30	5102,454
	3	-2505,53	5027,06	5064,107
	4	-2505,52	5033,04	5083,975
	5	-2503,28	5034,55	5099,381
Lognormal	2	-2547,94	5105,87	5129,026
	3	-2546,63	5109,36	5146,305
	4	-2497,59	5017,19	5068,127
	5	-2492,70	5013,40	5081,148

Çizelge 4.5.'te Pediatri YBÜ'ye hastaların GAS karma dağılım yaklaşımı için bileşen sayısı belirlemede kullanılan model istatistiklerini gösterilmektedir. Çizelge 4.5.'te, her karma dağılım modeli için log olabilirlik oranı (log LTR), AIC ve BIC değerleri verilmiştir. En iyi modeli seçmek için düşük AIC ve BIC değerlerine sahip olan model tercih edilir. Düşük AIC ve BIC değeri, daha iyi uyum sağlayan ve daha az karmaşık olan bir modeli temsil eder. Çizelge 4.5. incelendiğinde, en düşük AIC ve BIC değerine sahip karma dağılım; AIC değeri 5027,06 ve BIC değeri 5064,107 olan üç bileşenli gamma dağılımıdır. Böylece Pediatri YBÜ GAS verisinin üç bileşenli bir gamma dağılımı ile temsil edilebileceği söylenebilir.

Çizelge 4.6.'da, Pediatri YBÜ'deki SS'ye en iyi uyum gösteren karma dağılım modellenmesi için hesaplanan model istatistikleri paylaşılmıştır. Çizelge 4.6. dikkatli incelendiğinde, üç bileşenli lognormal karma dağılıma ait BIC değeri, tüm modeller arasında en küçük değeri almışken, AIC değeri için aynısı söylenemez. Ancak belirtildiği gibi, BIC değeri AIC'den daha güçlü bir bilgi kriteridir ve bu değerlerin ortak bir sonuç vermemesi durumunda BIC değeri dikkate alınır.

Çizelge 4.6. Pediatri YBÜ SS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri

Karma Dağılım	Bileşen Sayısı (k)	Model İstatistiği		
		log LTR	AIC	BIC
Normal	2	-4357,58	8725,16	8748,32
	3	-4235,44	8486,87	8523,93
	4	-4199,10	8420,21	8471,16
	5	-4196,07	8420,14	8484,99
Weibull	2	-4214,25	8438,49	8461,65
	3	-4187,41	8390,83	8427,88
	4	-4181,10	8384,20	8435,15
	5	-4178,63	8385,26	8450,11
Gamma	2	-4192,27	8394,55	8417,71
	3	-4144,19	8304,39	8341,44
	4	-4114,54	8251,09	8302,04
	5	-4124,94	8277,88	8342,73
Lognormal	2	-4165,93	8341,85	8365,01
	3	-4123,30	8262,59	8299,65
	4	-4114,31	8250,62	8301,58
	5	-4111,34	8250,67	8315,52

Pediatri YBÜ’de hastaların kalış süresi için karma dağılım modeli ve bileşen sayısı belirlenmesi için yapılan hesaplamalar sonucunda en iyi modelin (AIC = 8262,59 ve BIC = 8299,65) üç bileşenli lognormal karma dağılım olduğu tespit edilmiştir.

Çizelge 4.7.’de, Reanimasyon YBÜ’ye GAS’nin karma dağılım yaklaşımı için bileşen sayısı belirlemede kullanılan model istatistiklerini gösterilmektedir.

Çizelge 4.7. Reanimasyon YBÜ GAS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri

Karma Dağılım	Bileşen Sayısı (k)	Model İstatistiği		
		log LTR	AIC	BIC
Normal	2	-1186,13	2382,25	2399,94
	3	-1178,36	2372,71	2401,01
	4	-1137,59	2297,17	2336,09
	5	-1132,98	2293,96	2343,48
Weibull	2	-1118,42	2246,84	2264,53
	3	-1118,30	2250,80	2265,57
	4	-1115,90	2253,80	2292,71
	5	-1111,68	2251,37	2300,89
Gamma	2	-1116,04	2242,08	2259,76
	3	-1115,95	2247,89	2276,19
	4	-1115,45	2252,90	2291,81
	5	-1115,12	2258,23	2307,76
Lognormal	2	-1128,64	2267,28	2284,97
	3	-1124,49	2264,98	2293,28
	4	-1124,07	2270,14	2309,05
	5	-1118,10	2264,20	2313,72

Çizelge 4.7. incelendiğinde, Reanimasyon YBÜ GAS için karma dağılım modeli seçiminde, modeller arasında en küçük AIC ve BIC değerini sağlayan model iki bileşenli gamma dağılımı olduğu tespit edilmiştir.

Son olarak Reanimasyon YBÜ’de kalan hastaların SS verisinin karma dağılım bileşen analizi yapılmış ve model istatistikleri Çizelge 4.8.’de gösterildiği gibi hesaplanmıştır. Reanimasyon YBÜ’de kalan hastaların kalış süresine en iyi uyum gösteren model iki bileşenli lognormal karma dağılım olarak bulunmuştur.

Çizelge 4.8. Reanimasyon YBÜ SS Verisi Karma Dağılım Bileşen Sayısı Model İstatistikleri

Karma Dağılım	Bileşen Sayısı (<i>k</i>)	Model İstatistiği		
		log LTR	AIC	BIC
Normal	2	-1679,14	3368,28	3385,98
	3	-1641,01	3297,99	3326,32
	4	-1634,99	3291,99	3330,94
	5	-1631,48	3290,95	3340,53
Weibull	2	-1643,22	3296,44	3314,15
	3	-1637,87	3291,73	3320,06
	4	-1627,63	3277,25	3316,20
	5	-1625,41	3278,83	3328,41
Gamma	2	-1641,63	3293,26	3310,97
	3	-1628,58	3273,15	3301,48
	4	-1627,75	3277,49	3316,45
	5	-1623,04	3274,08	3323,66
Lognormal	2	-1635,01	3280,01	3297,71
	3	-1626,90	3269,80	3298,13
	4	-1622,85	3267,69	3306,65
	5	-1619,75	3267,49	3317,07

Çizelge 4.8. dikkatle incelendiğinde, Pediatri YBÜ’deki SS verisinin karma dağılım analizine benzer olarak, Reanimasyon YBÜ’de de BIC değerinin en düşük olduğu model ile AIC değerinin en düşük olduğu modelle örtüşmemektedir. Ancak BIC değerinin daha güçlü bir bilgi kriteri olması nedeniyle Reanimasyon YBÜ’de kalan hastaların SS verisini temsil etmek için iki bileşenli lognormal karma dağılım seçilmiştir.

Bir veri seti karma dağılım olarak modellenirken bileşen sayısı (*k*) belirlendikten sonra gelen aşama, karma dağılımı oluşturan dağılımların parametre tahminlerini yapmaktır.

Karma dağılım modeli; karma oranı $(\pi_1, \pi_2, \dots, \pi_k)$, bileşen sayısı (*k*) ve dağılım parametreleri $f_1(x; \theta_1), f_2(x; \theta_2), \dots, f_k(x; \theta_k)$ olmak üzere üç komponent ile tanımlanır. Parametre tahmini aşamasında karma oranı $(\pi_1, \pi_2, \dots, \pi_k)$ ve dağılım parametreleri $f_1(x; \theta_1), f_2(x; \theta_2), \dots, f_k(x; \theta_k)$ tahmin edilir. Bu çalışmada tahmin edici yöntem olarak Expectation-Maximization (EM – Beklenti Maksimizasyonu Algoritması) kullanılmıştır. Bu algoritmanın tüm ayrıntıları 3.6.2. Karma Dağılımlarda Parametre Tahmini bölümü altında açıklanmıştır. Kısaca Beklenti Maksimizasyonu (Expectation-maximization) algoritması, istatistik

alanında, istatistiksel modellerdeki parametrelerin en büyük olabilirlik tahminlerini bulmak için kullanılan bir yinelemeli arama yöntemidir. Beklenti Maksimizasyonu, beklenti (E) adımı ve maksimizasyon (M) adımı olmak üzere iki adımdan oluşmaktadır. E-adımı, parametrelerin güncel tahminlerini kullanarak bir log-olabilirlik beklentisi fonksiyonu oluşturur. M-adımı ise parametre değerlerini, log-olabilirlik beklentisini maksimize edecek şekilde günceller. Bu iki adım birbirini besleyerek çalışır, yani her bir adım diğerinin girdisini hesaplayarak ilerler. Beklenti Maksimizasyonu adımları, tahmindeki hata miktarı belirli bir eşik değerin altına düşene kadar tekrarlanır.

Karma dağılım parametre tahmini için uygulanan EM algoritması RStudio kullanılarak çalıştırılmıştır. RStudio model tabanlı kümeleme ve yoğunluk tahmini gibi denetimsiz öğrenmede yaygın olarak kullanılan sonlu karma modellerini oluşturmak ve analiz etmek için zengin bir paket koleksiyonu sağlamaktadır.

EM algoritması sonucunda elde edilen tahmin parametreleri Çizelge 4.9.'da gösterilmiştir. Her bir YBÜ, farklı karma dağılım modelleri ve bileşen sayıları ile temsil edilmektedir. Parametre tahminleri, her bir bileşenin ağırlığı ve ilgili dağılımın parametreleri olmak üzere Çizelge 4.9.'da ayrıntılı bir şekilde sunulmuştur. Bu bilgiler, veri setinin karma dağılım modelleri altında nasıl dağıldığını ve model parametrelerinin tahminlerini göstermektedir.

Çizelge 4.9. EM Algoritması Kullanılarak Tahmin Edilen Karma Dağılım Parametreleri

YBÜ	Tür	Karma Dağılım Modeli	Parametreler		
			k	π_k	θ_k
Pediatri	GAS	Gamma	3	0,68 0,25 0,07	Şekil (α)
					Oran (θ)
					1,07 21,33 5,96
					0,21 1,13 0,16
	SS	Lognormal	3	0,29 0,15 0,56	log-mu
					log-sd
					2,99 3,72 4,78
					0,22 0,11 0,93
Reanimasyon	GAS	Gamma	2	0,72 0,28	Şekil (α)
					Oran (θ)
					0,55 10,63
					0,016 0,37
	SS	Lognormal	2	0,89 0,11	log-mu
					log-sd
					4,65 6,54
					1,13 0,11

Karma oranı $(\pi_1, \pi_2, \dots, \pi_k)$, bileşen sayısı (k) ve dağılım parametreleri $f_1(x; \theta_1), f_2(x; \theta_2), \dots, f_k(x; \theta_k)$ olmak üzere üç komponent ile tanımlanır.

$$f(x) = \pi_1 f_1(x; \theta_1) + \pi_2 f_2(x; \theta_2) + \dots + \pi_k f_k(x; \theta_k)$$

$$0 \leq \pi_i \leq 1 \quad i = 1, 2, \dots, k$$

$$\sum_{i=1}^k \pi_i = 1$$

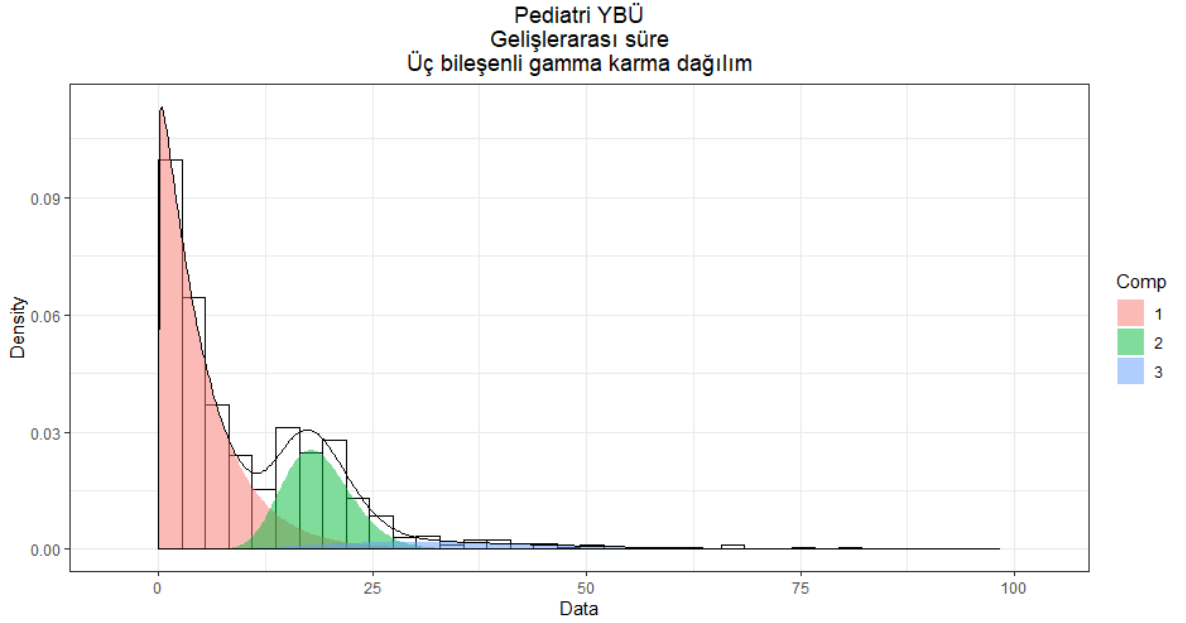
Karma dağılımların genel bir gösteriminde parametreler için ψ sembolü kullanılabilir ve $f(x, \psi) = \sum_{i=1}^k \pi_i f(x; \theta_i)$ şeklinde gösterilebilir. Buna göre, karma dağılım model parametreleri Pediatri ve Reanimasyon YBÜ'lerde toplanan GAS ve SS verileri için Çizelge 4.10.'da gösterildiği gibi genelleştirilebilir.

Çizelge 4.10. Tahmin Edilen Karma Dağılım Modellerinin Genel Gösterimi

YBÜ	Tür	Karma Dağılım Modeli
Pediatri	GAS	$\sum_{i=1}^k \pi_i f(x; \alpha_i, 1/\beta_i) = 0.68(1.07, 4.76) + 0.25(21.33, 0.88) + 0.07(5.96, 6.25)$
	SS	$\sum_{i=1}^k \pi_i f(x; \log \mu_i, \log \sigma_i) = 0.29(2.99, 0.22) + 0.15(3.72, 0.11) + 0.56(4.78, 0.93)$
Reanimasyon	GAS	$\sum_{i=1}^k \pi_i f(x; \alpha_i, \beta_i) = 0.72(0.55, 0.02) + 0.28(10.63, 0.37)$
	SS	$\sum_{i=1}^k \pi_i f(x; \log \mu_i, \log \sigma_i) = 0.89(4.65, 1.13) + 0.11(6.54, 0.11)$

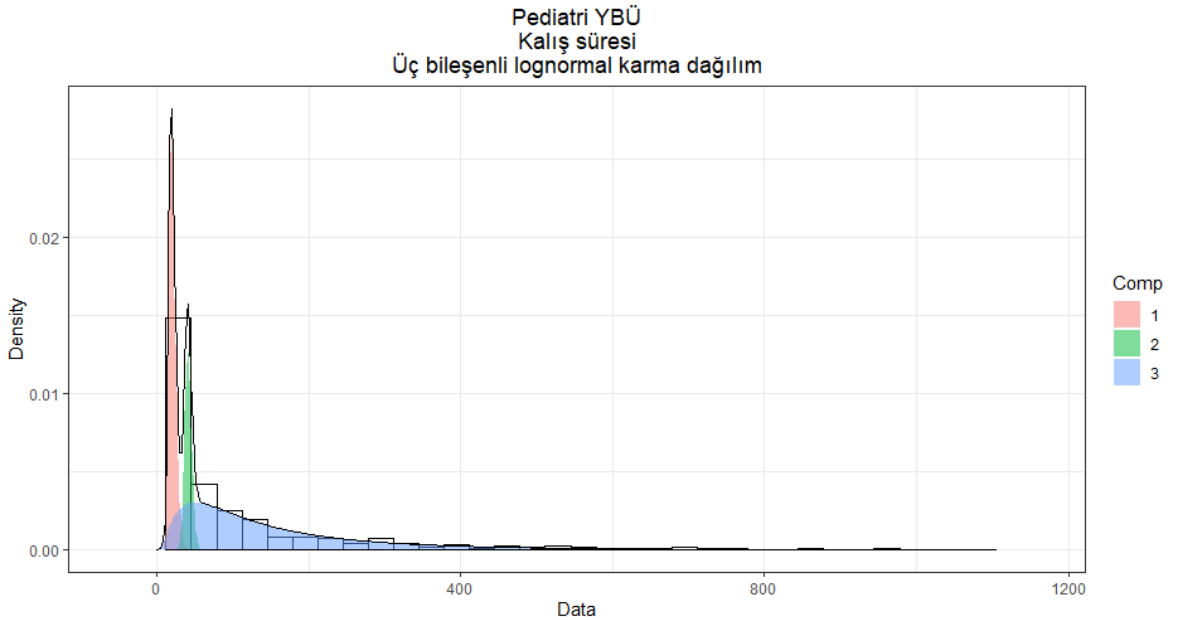
Tahmin edilen karma dağılım parametreleri kullanılarak çizilmiş OYF, gerçek verinin histogramı ile karşılaştırılmış ve uyumu grafiksel olarak izlenmiştir.

Şekil 4.26. Pediatri YBÜ'ye GAS için tahmin edilen üç bileşenli gamma dağılımın OYF'si ve gerçek gözlemlerin histogramını göstermektedir. Birinci bileşeni temsil eden kümenin ortalama değeri 5,09 ($\alpha\beta$) ve karma oranı %68 olarak hesaplanmıştır. Yani Pediatri YBÜ'ye hastaların geliş aralıkları %68 oranında 5,09 saat başı olmaktadır. İkinci bileşenin karma oranı %25 ve ortalama değeri ise 18,8 saattir. Üçüncü bileşeni temsil eden Pediatri YBÜ'ye seyrek varışların %7'lık karma oranı ve 37,25 saat değerine sahip ortalama değeri vardır.



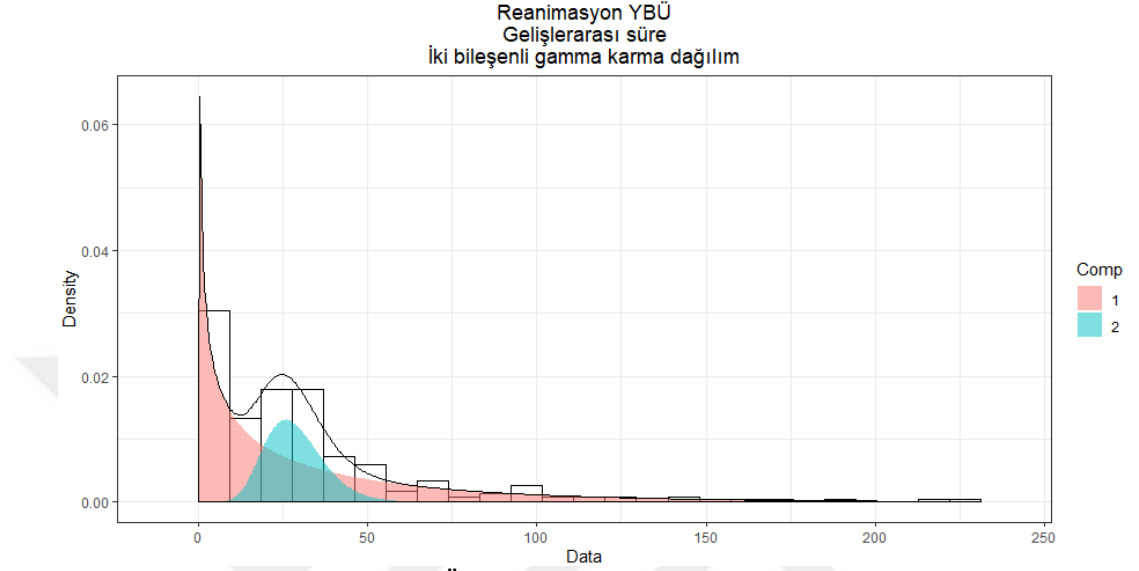
Şekil 4.26. Pediatri YBÜ’ye GAS Verisinin Karma Dağılım Yoğunluk Grafiği

Şekil 4.27.’de ise Pediatri YBÜ’deki SS için tahmin edilen üç bileşenli lognormal dağılımı OYF’si ve gerçek verinin histogramı aynı şekilde izlenmektedir. Birinci bileşen kısa süreli kalışları temsil eder. Bu kalışlar karma dağılımın %29’unu oluşturur ve ortalama değeri ($\exp(\log \mu) = \mu$) 19,88 saattir. İkinci bileşeni %15’lik karma oranı ve 41,26 saatlik ortalama değeri ile orta süreli kalışlar oluşturmaktadır. Üçüncü ve son bileşen Pediatri YBÜ’lerdeki uzun süreli kalışları temsil eder. Uzun süreli kalışlar %55’lik karma oranı ile en yüksek temsile sahip bileşendir. Bu bileşenin ortalama değeri ise 119,1 saattir.



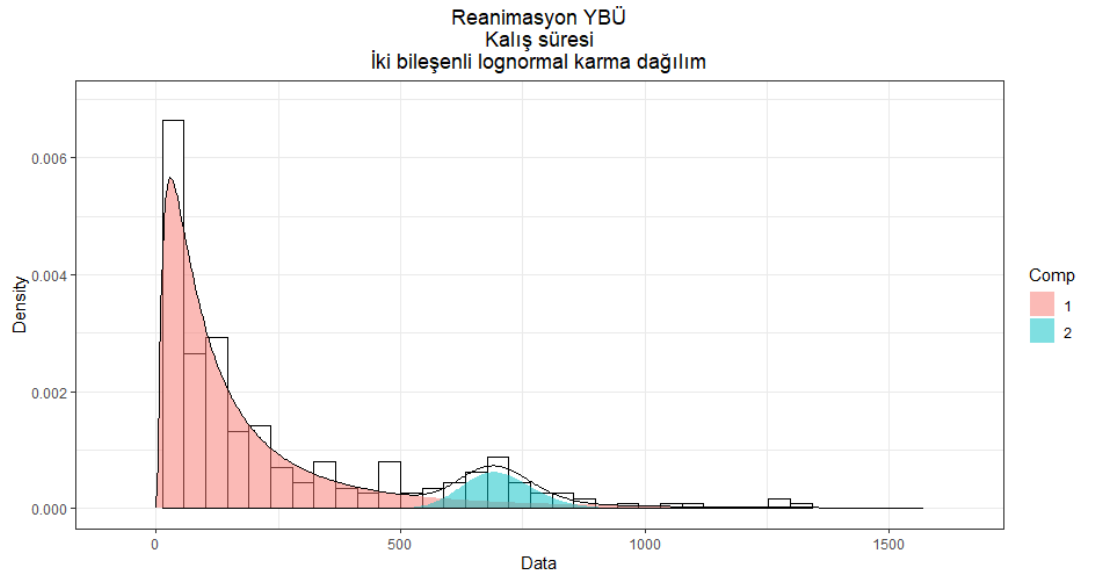
Şekil 4.27. Pediatri YBÜ’deki SS Verisinin Karma Dağılım Yoğunluk Grafiği

Pediatric YBÜ için karma dağılım parametreleri genel olarak yorumlanacak olursa; bu üniteye hastaların gelişi çoğunlukla kısa periyotta olmasına rağmen (~5 saat), bu üniteye kalış süreleri yüksek oranda uzun sürelidir (~119 saat). Bu durum, ciddi bir şekilde yönetilmesi gereken bir problem olarak tanımlanabilir.



Şekil 4.28. Reanimasyon YBÜ'ye GAS Verisinin Karma Dağılım Yoğunluk Grafiği

Ele alınan diğer ünite ise, Reanimasyon YBÜ'dür. Şekil 4.28., Reanimasyon YBÜ'ye GAS verisi için tahmin edilen iki bileşenli gamma dağılımın OYF'sini ve gerçek gözlemlerin histogramını içermektedir. Birinci bileşeni temsil eden kümenin ortalama değeri 34,5 saat ($\alpha\beta$) ve karma oranı %72 olarak hesaplanmıştır. İkinci bileşenin karma oranı %28 ve ortalama değeri ise 28,6 saattir. Pediatric YBÜ'ye benzer şekilde, Reanimasyon YBÜ'de de hastaların GAS değeri yüksek oranda daha kısa sürelidir.



Şekil 4.29. Reanimasyon YBÜ'deki SS Verisinin Karma Dağılım Yoğunluk Grafiği

Şekil 4.29., Reanimasyon YBÜ’de hastaların SS verisi için tahmin edilen iki bileşenli lognormal dağılımın OYF’sini ve gerçek gözlemlerin histogramını içermektedir. Lognormal karma dağılımın birinci bileşenin karma oranı %88’dir ve bu bileşenin ortalama değeri ($\exp(\log \mu) = \mu$) 105 saattir. Lognormal karma dağılımın ikinci ve son bileşenin karma oranı ise kalan %12’lik bölümdür ve yaklaşık 1 aya yakın süreyi temsil eden 690 saatlik bir kalış süresini temsil eder.

Reanimasyon YBÜ için karma dağılım parametreleri genel olarak incelenirse; bu üniteye hastaların gelişi çoğunlukla kısa periyotta olmasına rağmen (~35 saat), bu üniteye kalış süreleri yüksek oranda ikinci bileşen göre daha kısa sürelidir (~105 saat). Pediatri YBÜ’ye benzer şekilde, Reanimasyon YBÜ’de de gelişler ile üniteye kalışlar arasında farklar bulunmaktadır. Bu dengesiz durumun yönetilebilmesi için en önemli kriterlerin başında yatak sayısı belirlenmesi gelir. Hastalar ne kadar uzun süre kalırsa kalsın ya da sıklıkla geliyor olsun, yatak sayısı ve kapasitesi doğru bir şekilde belirlenebilir ve yönetilebilirse, sistem en az problem ile ilerler.

Veri genel olarak çarpık bir dağılım gösterdiğinde ve saf bir istatistiksel dağılımla temsil edilemediğinde ampirik dağılım oluşturulması klasik bir yaklaşımdır. Yapılan istatistiksel analizler sonucu Pediatri ve Reanimasyon YBÜ’lerine hastaların GAS ve bu ünitelerde SS gözlemleri, test edilen hiçbir saf istatistiksel dağılıma uyum göstermemiştir. Bu noktada, çalışmanın en çarpıcı noktalarından birisi olan karma dağılım yaklaşımı kullanılarak veri temsil edilmiştir. Tahmin edilen bu karma dağılımların ve ampirik dağılımların gerçek veriyi temsil yeteneği kıyaslanmak hedefiyle bu süreleri temsil edecek ampirik dağılımlar da oluşturuldu.

4.3. Ampirik Dağılım Oluşturma

Simülasyon modeli için kullanılan veri her zaman bilinen saf bir istatistiksel dağılıma uymaz. Klasik yaklaşımda, verinin bilinen saf dağılımlara uymadığı gibi durumlarda ampirik dağılımlar kullanılır. Bu yöntemde, gerçek sistemden elde edilen veri kullanılarak sistemin giriş parametrelerinin ve çıktıların dağılımı tahmin edilir. Ancak ampirik dağılımları kullanmanın bir takım dezavantajları mevcuttur. Öncelikle, elde edilen verinin sınırlı olması durumunda, tahmin edilen dağılımların gerçek sistemdeki dağılımlardan farklı olabileceği bir risk bulunmaktadır. Ayrıca, özellikle büyük veri setleri kullanıldığında ampirik dağılımların oluşturulması zaman ve kaynak gerektirebilir. Bunun yanı sıra, veri hataları veya yanlışlıkla girilen varsayımlar gibi faktörler ampirik dağılımların doğruluğunu etkileyebilir.

Pediatri ve Reanimasyon YBÜ’lerine hastaların GAS ve bu ünitelerde SS değerleri yapılan istatistiksel analizler sonucu bilinen hiçbir saf dağılıma uymadığı görülmüş klasik yaklaşım olan ampirik dağılım oluşturulmuştur. Her bir veri için oluşturulan ampirik dağılım modelleri Çizelge 4.11.’de verilmiştir.

Çizelge 4.11. Tahmin Edilen Ampirik Dağılım Modellerinin Genel Gösterimi

YBÜ	Tür	Ampirik Dağılım Modeli
Pediatri	GAS	CONT(0.000,0.000,0.299,3.074,0.483,6.148,0.573,9.222,0.631,12.296,0.708,15.370,0.793,18.444,0.875,21.519,0.925,24.593,0.950,27.667,0.959,30.741,0.966,33.815,0.971,36.889,0.979,39.963,0.984,43.037,0.989,46.111,0.991,49.185,0.993,52.259,0.9931,55.333,0.995,58.407,0.9951,61.481,0.9952,64.556,0.997,67.630,0.9971,70.704,0.9972,73.778,0.998,76.852,0.999,79.926,1,83.000)
	SS	CONT(0.000,12.000,0.526,49.000,0.653,86.000,0.759,123.000,0.792,160.000,0.827,197.000,0.855,234.000,0.877,271.000,0.901,308.000,0.922,345.000,0.928,382.000,0.939,419.000,0.946,456.000,0.951,493.000,0.957,530.000,0.966,567.000,0.971,604.000,0.974,641.000,0.978,678.000,0.983,715.000,0.987,752.000,0.988,789.000,0.991,826.000,0.993,863.000,0.994,900.000,0.995,937.000,0.999,974.000,1,1011.000)
Reanimasyon	GAS	CONT(0.000,0.000,0.402,18.533,0.736,37.067,0.854,55.600,0.902,74.133,0.921,92.667,0.953,111.200,0.969,129.733,0.976,148.267,0.980,166.800,0.984,185.333,0.988,203.867,0.992,222.400,0.996,240.933,0.996,259.467,1,278.000)
	SS	CONT(0.000,12.000,0.412,100.667,0.600,189.333,0.694,278.000,0.749,366.667,0.776,455.333,0.824,544.000,0.859,632.667,0.925,721.333,0.957,810.000,0.976,898.667,0.80,987.333,0.984,1076.000,0.988,1164.667,0.99,1253.333,1,1342.000)

Pediatri ve Reanimasyon YBÜ'lerine hastaların GAS ve hastaların bu ünitelerdeki SS verisini temsil eden karma dağılımlar ve ampirik dağılımlar sistematik bir şekilde oluşturuldu. Elde edilen bu iki farklı yaklaşımın gerçek veriyi temsil güçleri karşılaştırıldı ve en iyi uyumu veren yaklaşımın parametrelerinin simülasyon modelinde girdi parametresi olarak kullanılması hedeflendi. Bu kapsamda, bir sonraki adımda, bu iki yaklaşımın performans değerlendirme kriterleri belirlendi ve değerlendirildi.

4.4. Performans Kriterlerini Değerlendirme

Simülasyon modelleri, gerçek bir sistem veya sürecin davranışını taklit etmek için kullanılan matematiksel ve istatistiksel modellerdir. Bu modeller, sistemin karmaşık yapısını ve etkileşimlerini temsil eden bir dizi parametre ve değişken içerir. Girdi parametreleri, modelin çıktılarını etkileyen faktörlerdir ve bu nedenle doğru bir şekilde belirlenmeleri büyük önem taşır.

Bir simülasyon modeli oluşturulurken, gelişler arası süreler ve kalış süreleri gibi girdi parametrelerinin doğru bir şekilde istatistiksel olarak belirlenmesi, modelin geçerliliğini ve sonuç doğruluğunu artırma açısından büyük öneme sahiptir. Girdi parametrelerinin doğru bir şekilde tahmin edilmesi, simülasyon modelinin gerçek dünyayı yansıtmaya yeteneğini ve yapılan analizlerin güvenilirliğini etkiler.

Gelişler arası süreler ve kalış süreleri gibi girdi parametrelerinin doğru bir şekilde tahmin edilmesi, simülasyon modelinin güvenilirlik ve geçerlilik açısından sağlam bir temele oturmasını

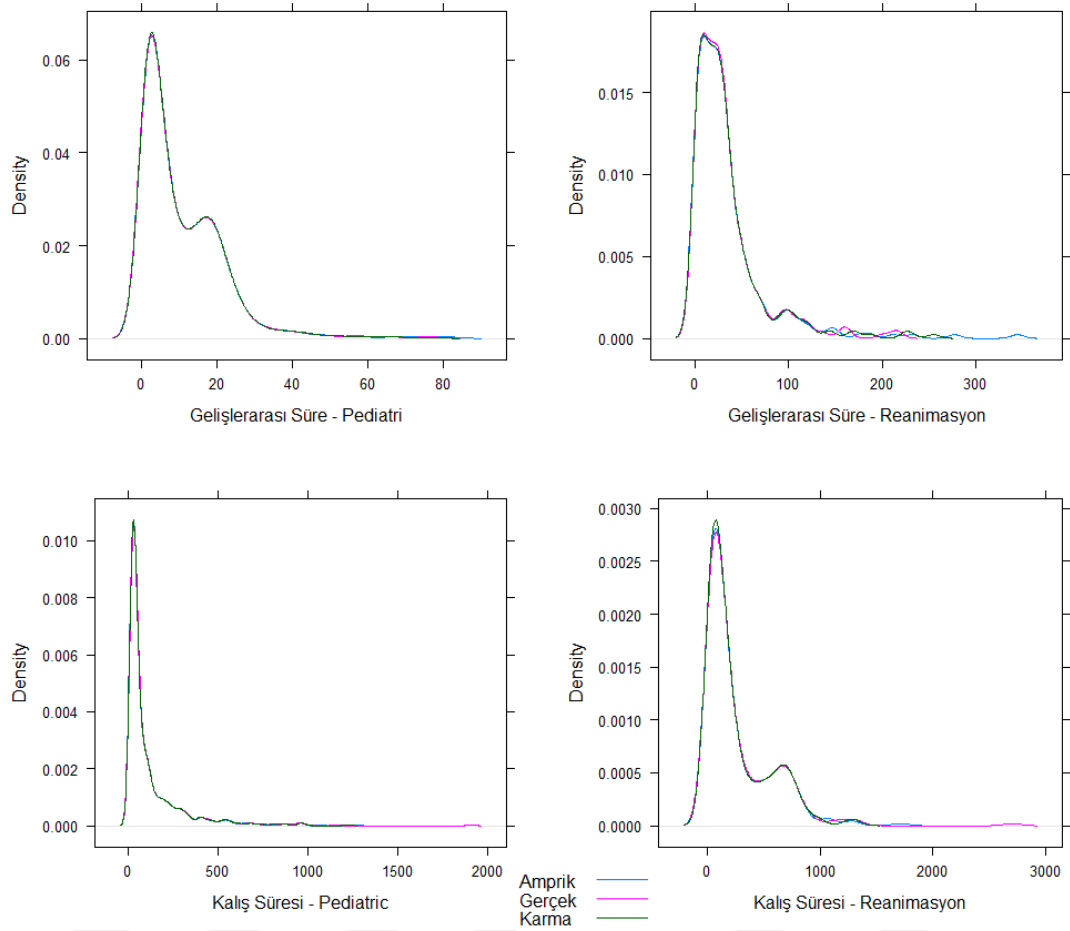
sağlar. Böylece simülasyon modelinin gerçek dünyadaki sistemi doğru bir şekilde temsil yeteneği artar ve yapılan analizler daha gerçekçi sonuçlar üretir.

Bu girdi parametrelerinin belirlenmesi için çeşitli istatistiksel dağılımlar kullanılır ancak yüksek çarpıklık gibi durumlarda saf bilinen istatistiksel dağılımlar veri temsilde yeterli olmayabilir. Bu gibi durumlarda geleneksel olarak ampirik dağılımlar kullanılır ancak ampirik dağılımlar 0.

Ampirik Dağılım başlığında ayrıntılı bir şekilde ele alındığı üzere bir takım dezavantajlara sahiptir. Bu çalışmada, yüksek çarpıklık söz konusu olduğu durumlar söz konusudur ve bu veri yeni bir yaklaşım olarak karma dağılımlar ile temsil edilmek istenmiştir.

Tüm ampirik ve karma dağılım modelleri, farklı derecelerdeki doğrulukla gelişler arası süre ve kalış süresini tahmin eder. Simülasyon modeli için girdi parametresi olarak doğruluğu en yüksek olarak model seçilmek tercih edilmelidir. Bu kapsamda en uygun yaklaşım; grafiksel yöntem, ortalama mutlak yüzde hata (MAPE) ve Mann-Whitney U (MWU) test değerleri kullanılarak değerlendirildi.

Şekil 4.30.'da gerçek verinin OYF'si ile tahmin edilen karma ve ampirik dağılımların OYF'leri çizilmiştir. Her iki yaklaşımın da gerçek veriye uyum sağladığı görülebilir ancak istatistiksel olarak anlamlı bir fark olup olmadığının test edilmesi gerekmektedir.



Şekil 4.30. Gerçek, Ampirik ve Karma Dağılımların Yoğunluk Grafiği

Ampirik ve karma dağılımların gerçek veri seti olan uyumunda performans kriterleri olarak MAPE değeri ve MWU testi p -değeri kullanılmıştır.

MAPE (Mean Absolute Percentage Error), Türkçe olarak Ortalama Mutlak Yüzde Hata olarak adlandırılır ve tahminlerin gerçek değerlere göre yüzde hata oranını ölçen bir performans ölçütüdür. MAPE, tahmin edilen değer ile gerçek değer arasındaki mutlak farkın yüzde olarak ifade edilmesiyle hesaplanır. Her gözlem için mutlak farkın yüzde değeri alınır ve ardından tüm gözlemler için ortalama alınarak MAPE değeri elde edilir. MAPE, pozitif değerlerle ifade edilir ve yüzde olarak ifade edilen bir hata oranıdır. Tahminlerin gerçek değerlere ne kadar yakın olduğunu gösterir. MAPE değeri ne kadar düşükse, tahminlerin gerçek değerlere o kadar yakın olduğu anlamına gelir.

Mann-Whitney U (MWU) testi, iki bağımsız grup arasındaki merkezi eğilim farkını değerlendirmek için kullanılan istatistiksel bir hipotez testidir. Parametrik olmayan bir test olarak bilinir, yani verinin belirli bir dağılıma uymasını gerektirmez. Eğer hesaplanan p -değeri önem düzeyinden küçükse, gruplar arasında istatistiksel olarak anlamlı bir fark olduğu kabul edilir. MWU testi, parametrik varsayımların sağlanmadığı ve özellikle grup verisi normal dağılım

göstermediğinde sıklıkla kullanılan bir testtir. Bu her iki performans kriteri ile ilgili ayrıntılı bilgi 3.6.3. Performans Kriteri başlığı altında verilmiştir.

Ampirik ve karma dağılımların gerçek veriyi temsil etme yeteneklerini kıyaslamak amacıyla hesaplanan MAPE değeri ve MWU testi p -değeri kullanılarak karşılaştırılmıştır ve sonuçlar Çizelge 4.12.'de paylaşılmıştır. Bu karşılaştırma sonucunda, en iyi performansı gösteren yöntemin parametreleri, simülasyon modelinde girdi parametresi olarak kullanılması hedeflenmektedir. Bu şekilde, modelin gerçek veriyle uyumlu sonuçlar üretebileceği ve doğru tahminler yapabileceği öngörülmüştür. Bu sonuçlar, simülasyon modelinin daha doğru ve güvenilir sonuçlar üretmesine yardımcı olacaktır.

Çizelge 4.12. Ampirik ve Karma Dağılımlarının Performans Kriterlerinin Karşılaştırılması

YBÜ	Tür	MAPE (%)		MWU (p -değeri)	
		Ampirik	Karma	Ampirik	Karma
Pediatri	GAS	16,08	8,21	0,48	0,84
	SS	9,11	3,52	0,24	0,93
Reanimasyon	GAS	22,71	8,85	0,49	0,89
	SS	19,89	8,31	0,15	0,83

Pediatri YBÜ için GAS verisi, ampirik dağılım modeliyle %16,08'lik bir hata (MAPE) değeri ile tahmin edilmiştir. Bu tahmin değeri, karma dağılım modeli kullanılarak elde edilen %8,21 MAPE değerinden daha yüksektir. MWU testi p -değeri ampirik dağılım için 0,48 iken karma dağılım için 0,84 olarak hesaplanmıştır. Pediatri YBÜ'ye hastaların gözlenen GAS verisinin MAPE ve MWU testi p -değeri kıyaslandığında karma dağılım ile temsil edilmesi daha yüksek doğruluk sağladığı görülmektedir.

Pediatri YBÜ için SS verisi, ampirik dağılım yaklaşımıyla %9,11, karma dağılım modeli ile ise %3,52 bir hata ile modellenmiştir. MWU testi p -değeri ise ampirik dağılım için 0,24 iken karma dağılım için 0,93 olarak bulunmuştur ve yüksek bir doğrulukla verinin karma dağılım ile temsil edilebileceği söylenir.

Reanimasyon YBÜ için GAS verisi, ampirik dağılım %22,71'lik hata değeri ile temsil edilmiştir. Buna karşı, karma dağılım yaklaşımı Reanimasyon YBÜ'ye hastaların GAS verisi karma dağılım yaklaşımıyla sadece %8,85'lik bir MAPE hatasıyla temsil edilmiştir. MWU p -değerleri değerlendirildiğinde ise; bu değer ampirik dağılım için 0,49 iken karma dağılım için 0,89 olarak bulunmuştur.

Reanimasyon YBÜ için SS verisi, ampirik dağılım %19,89, karma dağılım yaklaşımı ise sadece %8,31'lik bir hata ile modellenmiştir. MWU p -değerleri ise ampirik dağılım için 0,15 iken karma dağılım için 0,83 olarak bulunmuştur.

Pediatri ve Reanimasyon YBÜ'ler için GAS ve SS verisi ampirik ve karma dağılımlar ile temsil edildi ve her iki yaklaşımın performansları izlendi. Tahmin edilen bu parametrelerin gerçek

veri setini temsil yeteneği karşılaştırıldığında, tüm durumlar için karma dağılımın ampirik dağılımdan üstün olduğu görülmüştür.

MAPE değeri ayrıntılı bir şekilde incelendiğinde, tüm durumlarda karma dağılım yüksek doğruluklu bir tahmin olarak nitelendirilen %10'un altında bir hata ile gerçek veri seti temsil edilmiştir. Ampirik dağılım yaklaşımında ise, % 9 ile %22 arasında değerler hesaplanmıştır. Ampirik dağılımda hesaplanan tüm MAPE değerleri tanıma göre kabul edilebilir değerler içindedir ancak karma dağılım parametreleriyle elde edilen MAPE değerleri ile karşılaştırıldığında oldukça yüksektir.

Ampirik ve karma dağılımın performansı MWU testi p -değerleri performans değerlendirme kapsamında karşılaştırılmıştır. Tüm durumlar için p -değerleri kritik değerden büyüktür. Bu, ampirik ve karma dağılım yaklaşımlarının temsil edilen ilgili veri ile aynı kitleden geldiği istatistiksel olarak ispatlar. Ancak Çizelge 4.12.'deki MWU testi p -değerleri ayrıntılı bir şekilde incelenirse, tüm veri için karma dağılım yaklaşımı parametrelerinin p -değerleri, ampirik dağılım yaklaşımı için hesaplanan p -değerlerinden büyük ve neredeyse 1'e yakındır. Bu durum, karma dağılım yaklaşımının istatistiksel olarak ilgili verisi daha güçlü şekilde temsil ettiğini gösterir.

Sonuç olarak, karma dağılım modelleri ampirik dağılım yaklaşımına göre; daha düşük MAPE değerleri ve daha yüksek MWU p -değerleri ile daha iyi performans sergilemektedir. Bu, karma dağılım modellerinin verisi ampirik dağılımdan daha iyi açıkladığını ve daha doğru tahminler sağladığını gösterir.

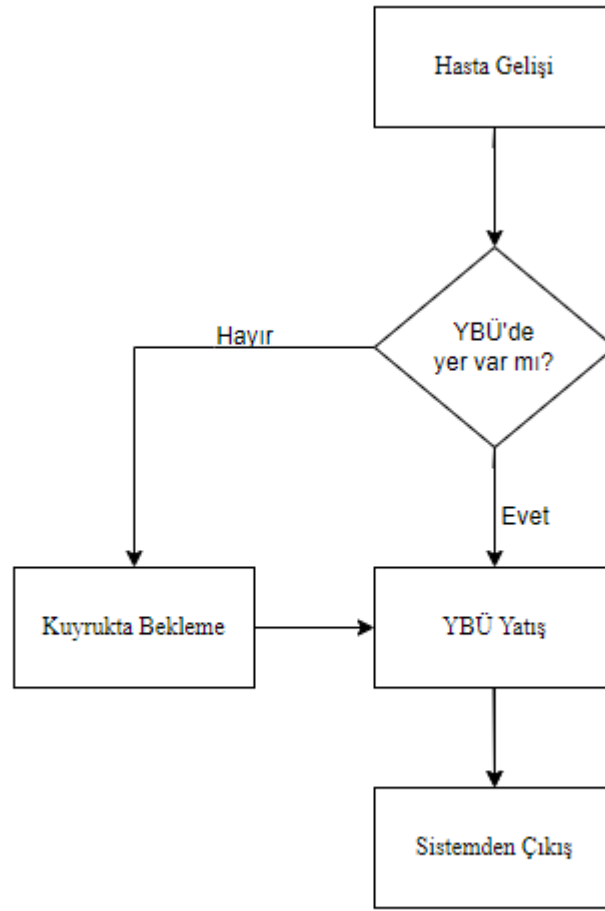
Böylece çalışmanın temelini oluşturan, simülasyon modelinde girdi parametresi olarak kullanılacak olan çarpıklık nedeniyle istatistiksel olarak saf bir dağılımla temsil edilemeyen verinin, klasik olarak ampirik dağılım yaklaşımı kullanılması yerine karma dağılım ile modellenmesi fikrinin oldukça iyi bir çıkış noktası olduğu görülmüştür.

Bir sonraki adımla oluşturulacak simülasyon modelinde tahmin edilen karma dağılım parametreleri girdi olarak kullanılacak ve belirlenecek amaçlar doğrultusunda simülasyon modeli yönetilecektir.

4.5. Simülasyon Modeli Sonuçları

Simülasyon, Endüstri Mühendisliği alanında yaygın olarak kullanılan bir araçtır. Simülasyon, gerçek sistemlerin matematiksel modellerini kullanarak sanal ortamda taklit etmeyi amaçlar. Bu çalışmada kurulan simülasyon modeli için üç farklı senaryo geliştirilmiştir.

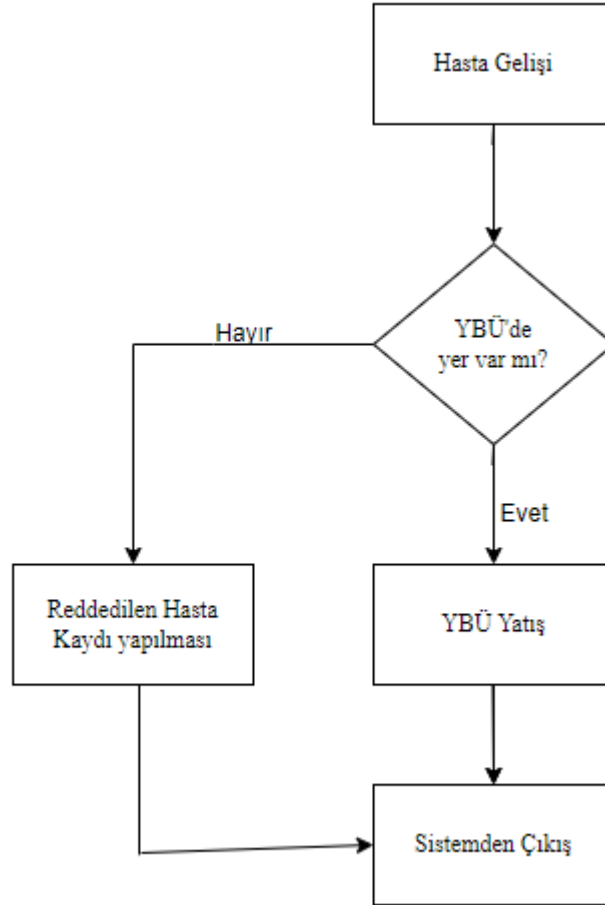
İlk iki senaryoda sistem gerçek yatak sayısı ile değerlendirilmiştir. Birinci senaryoda, hastaların kuyruk beklemesine izin vererek simülasyon modeli çalıştırılmıştır. Geliştirilen bu senaryonun algoritması Şekil 4.31.'de gösterilmiştir.



Şekil 4.31. Senaryo I, Kuyrukta Beklemeye İzin Verilen Simülasyon Modeli

Şekil 4.31.'de akışı verilmiş Senaryo I'de hastalar YBÜ'ye gelir. Eğer hizmet alabileceği boş yatak varsa, ilgili yatağı servis süresi kadar işgal eder ve bu sürenin sonunda sistemden ayrılır. Hizmet alabileceği boş bir yatak yoksa, kuyruğa girer ve FİFO stratejisine göre ilerleyen sırada boş yatak olduğu ilk anda yatağı işgal eder. Servis süresinin sonunda sistemi terkeder.

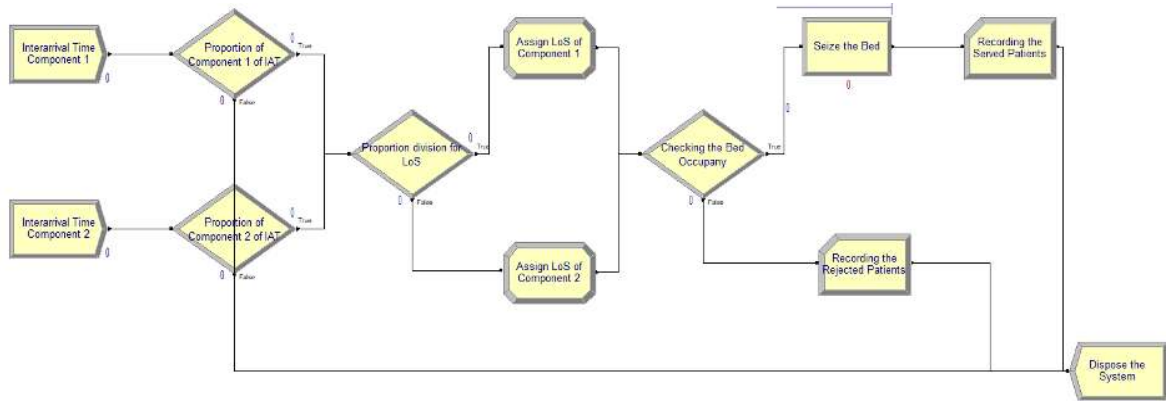
Geliştirilen ikinci senaryoda YBÜ'ye gelen hastaların uygun yatak olmaması durumunda kuyrukta beklemesine izin verilmemiş ve bekleme süresi olmaksızın sistemden çıkması sağlanmıştır. Bu senaryonun algoritması ise Şekil 4.32.'de gösterilmiştir. Senaryo II'de, Senaryo I'den farklı olarak sisteme giren hasta, giriş yaptığı anda uygun bir yatak olmaması durumunda "kaçınma" politikası izleyerek kuyrukta beklemesini ve sistemi terkeder. Bu senaryoda hastanın hizmet görebileceği tek durum, sisteme giriş yaptığı anda boş YBÜ yatağının varlığıdır. Senaryo II'nin bu şekilde kurulmasının nedeni, YBÜ'lerin hayati açıdan oldukça kritik hastalara hizmet vermesi nedeniyle her an uygun yatak bulunması gerekliliği fikri olmuştur.



Şekil 4.32. Senaryo II, Kuyrukta Beklemeye İzin Verilmeyen Simülasyon Modeli

Senaryo III ise, Senaryo II'nin yatak sayısı parametresinin ardışık bir şekilde değiştirilerek gelen tüm hastaların bekleme süresi olmaksızın kabulü için gereken nihai yatak sayısı değerlendirildiği durumdur. Akış olarak Senaryo II ile aynıdır.

Simülasyon modelinin geçerliliği (validation) ve doğruluğu (verification), güvenilir sonuçları elde etmek için esastır. Bu nedenle, kuyruk teorisi formülleri kullanılarak, bilinen her parametre için yaklaşık beklenen sonuçlar hesaplanmıştır. Daha sonra bu bilinen parametreler ile ARENA modeli çalıştırılmış ve beklenen sonuçlar simülasyon sonuçları ile karşılaştırılmıştır. Bu şekilde geliştirilen simülasyon modeli doğrulanmıştır. Senaryo algoritmaları benimsenerek Şekil 4.33.'de ara yüzü görülen geçerliliği ve doğruluğu sınanmış simülasyon modeli oluşturulmuştur.



Şekil 4.33. Simülasyon Modeli Ara Yüzü

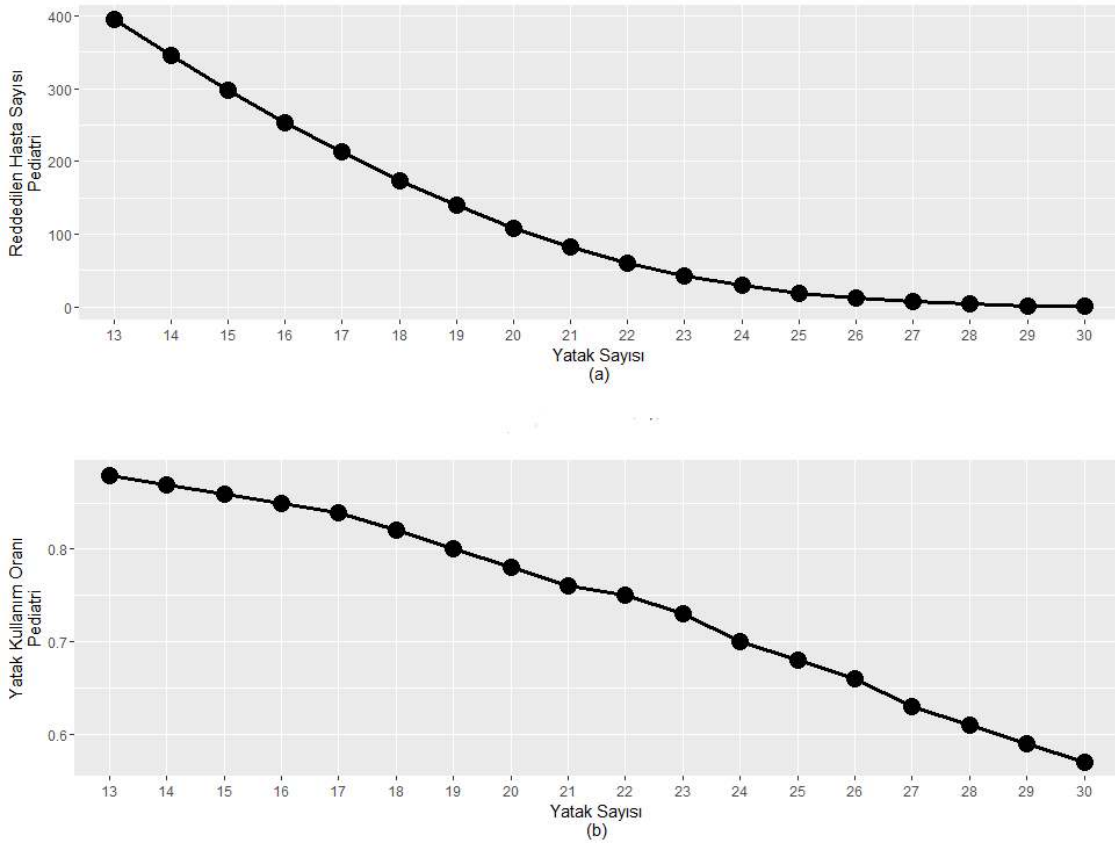
Mevcut sistemde Çukurova Üniversitesi Balcalı Hastanesi Eğitim ve Araştırma Hastanesi Pediatri YBÜ’de yatak sayısı 13, Reanimasyon YBÜ’de ise 8 adettir. Simülasyon modeli her bir replikasyonda rasgele sayılar üretir ve bu sebeple bir çok bağımsız tekrar gerekir. Bu konuyla ilgili ayrıntılı bilgi 3.7.2. Simülasyon Modeli başlığı altında verilmiştir. Güvenilir sonuçlar elde edebilmek için, replikasyon sayısının %95 güven aralığı değeri göz önüne alınarak, 50 olması gerektiği hesaplanmıştır. Isınma periyodu ise toplam simülasyon süresinin %10’u olarak alınması literatüre dayandırılarak uygun bulunmuştur. YBÜ, 7/24 çalışan bir sistem ve bu ünitelerden verinin bir yıl boyunca toplanmış olmasından dolayı toplam simülasyon süresi 8760 (365*24) saattir. Buna göre, ısınma periyodu 876 saat olarak ayarlanmıştır.

Senaryo I’de mevcut yatak sayısı değiştirilmeden sistem değerlendirilmiştir. YBÜ’ye gelen hastalar YBÜ’de boş yatak yoksa kuyrukta beklemekte, uygun yatak olur olmaz hizmet görmektedir. Senaryo I’e göre, Pediatri YBÜ’de yatak doluluk oranı %99, Reanimasyon YBÜ’de ise %91 olarak hesaplanmıştır. YBÜ’lerin hastanedeki kritik öneminden dolayı bu kullanım oranlarının oldukça yüksek olduğu söylenebilir. Reanimasyon YBÜ’de yatakta kalış süresi ortalama 35 saat olarak hesaplanmıştır ve bu değerler gerçek sistem ile uyumludur.

Senaryo II’de mevcut yatak sayısı değiştirilmeden, ancak hastaların uygun yatak olmadığı durumda sistemi terkedecekleri bir algoritma modellenmiştir. Bu durumda Pediatri YBÜ’de hastaların %33, Reanimasyon YBÜ’de ise %24’ünün reddedilerek sistemden çıkış yaptığı simülasyon çıktısı olarak görülmüştür. Senaryo II’de Pediatri YBÜ’de yatak doluluk oranı %88 iken, Reanimasyon YBÜ’de %74’tür.

Senaryo I ve II’de mevcut yatak sayıları değiştirilmeden değerlendirilmiştir. Senaryo III, YBÜ’ye gelen tüm hastaların herhangi bir bekleme süresine maruz kalmadan hizmet alabileceği yatak sayısını belirlemeyi amaçlayan simülasyon tabanlı bir yaklaşımdır. Senaryo II kullanılarak, mevcut yatak sayısından ardışık olarak arttırılmış ve kuyrukta bekleme süresinin olmadığı yatak sayısı belirlenmiştir.

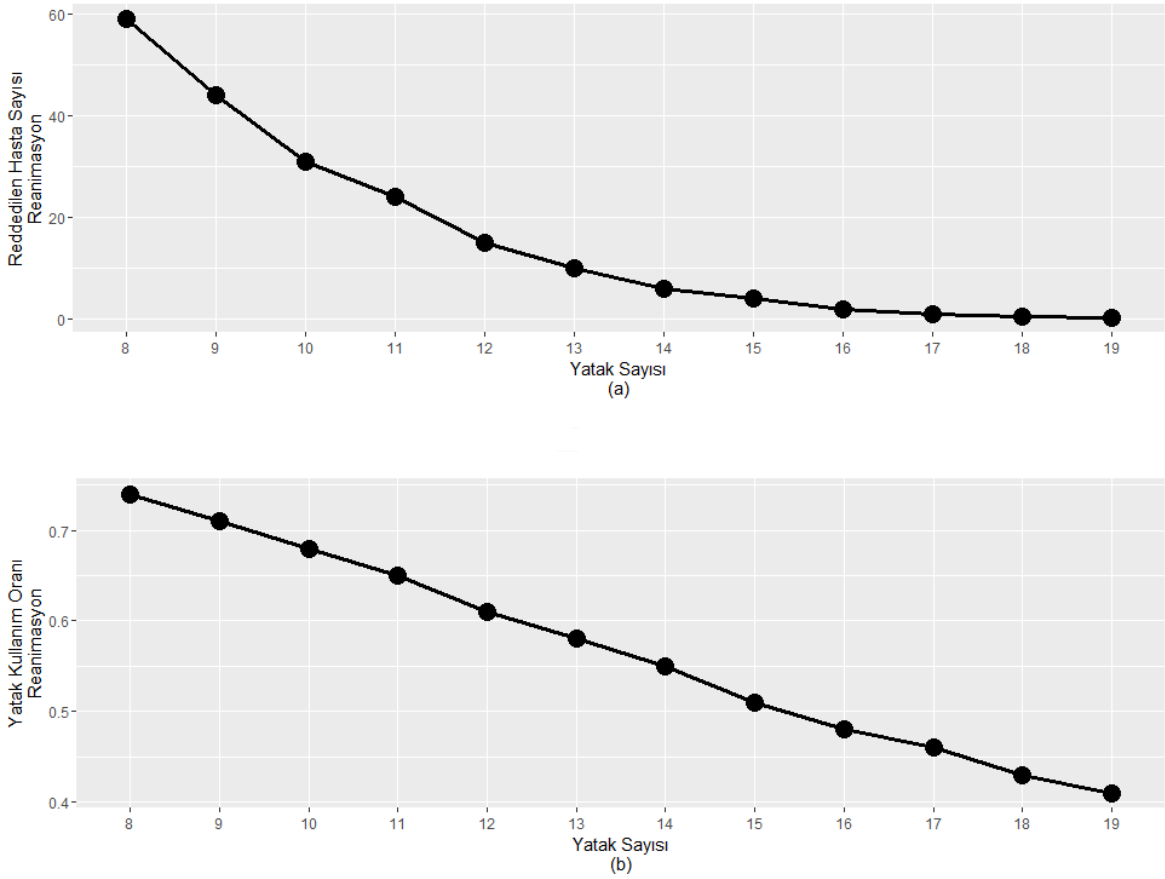
Pediatric ve Reanimasyon YBÜ’lerinde yatak sayısının arttırılarak simülasyon modeli çalıştırılmış ve ortalama reddedilen ve servis verilen hasta sayıları ve yatak kullanım oranları her bir yatak sayısı için kaydedilmiştir.



Şekil 4.34. Senaryo III, Pediatric YBÜ Yatak Sayısına Simülasyon Sonuçları

Şekil 4.34. (a)’da Pediatric YBÜ için kaydedilen ortalama reddedilen hasta sayısı grafik şeklinde paylaşılmıştır. 13 yatak olduğu durumda, gelen hastaların %33’ü reddedilerek hizmet verilememiştir. Yatak sayısı arttırıldıkça reddedilen hasta sayısı dramatik şekilde düşmüştür. Yatak sayısı 25 olduğunda ortalama reddedilen hasta yüzdesi %1,5’e düşer. Yatak sayısının 30’a yükseltiği durumda ise ise, reddedilen hasta sayısı ve yüzdesi 0 olmaktadır.

Yatak sayısı arttıkça yatak kullanım oranının düşmesi beklenen bir durumdur. Pediatric YBÜ için bu değerlerin izlenebileceği grafik Şekil 4.34. (b)’de paylaşılmıştır. Yatak sayısının 13 olduğu durumda ortalama yatak kullanım oranı %88 olarak hesaplanmıştır. Sistematik olarak arttırılan yatak sayısı bu oranın düşmüş ve yatak sayısının 25 olduğu durumda ortalama yatak kullanım oranı %70, 30 olduğu durumda ise %57’ye düşmüştür.



Şekil 4.35. Senaryo III, Reanimasyon YBÜ Yatak Sayısına Simülasyon Sonuçları

Senaryo III, Reanimasyon YBÜ için simülasyon modelinde çalıştırılmış ve ortalama reddedilen hasta ve ortalama yatak kullanım oranı irdelenerek her bir yatak sayısı için sonuçlar Şekil 4.35.'de paylaşılmıştır. Şekil 4.35. (a) grafiği, her bir yatak sayısı için hesaplanan ortalama reddedilen hasta sayısını vermektedir. Buna göre, mevcut yatak sayısı olan 8'de beklemeye izin vermeyen Senaryo III modelinde hastaların %24'ü reddedilerek hizmet görememektedir. Yatak sayısının 15'e yükseltildiği durumda ortalama reddedilen hasta toplam hasta sayısının sadece %1,5'ini oluşturur. Yatak sayısının 19 olduğu durumdaysa gelen tüm hastalara hiçbir bekleme süresi olmadan hizmet verilebilir.

Şekil 4.35. (b)'de, Reanimasyon YBÜ'de yatak sayısı arttırıldıkça yatak kullanım oranının düşüşünün izlediği grafik verilmiştir. Yatak sayısının 8 olduğu mevcut durumda, ortalama yatak doluluğu %74 olarak hesaplanmıştır. Yatak sayısının artmasıyla bu oran düşerek 15 yatak sayısı olan bir sistemde %51'e, 19 olduğu durumda ise %41'e kadar düşmüştür. Yatakların düşük kullanım oranına rağmen, YBÜ'lerin acil tıbbi müdahale gerektiren kritik durumlardaki hastalara hizmet verdiği ve bu nedenle her zaman hazır bir yatağa sahip olması gerektiği akılda tutulmalıdır.

5. SONUÇLAR VE ÖNERİLER

5.1. Çalışmanın Özeti ve Sonuçlar

Bu çalışma, Adana’da bulunan Çukurova Üniversitesi Balcalı Eğitim ve Araştırma Hastanesinin yoğun bakım ünitesi (YBÜ) yatak kapasitesi problemi. Ele alınan Pediatri ve Reanimasyon YBÜ’deki tüm hastalara bekleme süresi olmaksızın hizmet vermek için gereken yatak kapasitesinin yönetilmesi için girdi parametreleri olarak karma dağılımının kullanıldığı simülasyon tabanlı bir model sunulmuştur. Simülasyona dayalı bu model, karar vericilerin hasta reddetme oranını en aza indirmek ve kritik durumdaki hastalara bekleme süresi olmadan hizmet verilmesi için ilave yatak sağlamanın faydasını değerlendirmesini sağlar.

YBÜ’ler, hastanelerde kritik hastaların yoğun gözetim, tedavi ve bakımını sağlamak amacıyla tasarlanmış özel birimlerdir. Bu ünitelerde yaşanacak uzun bekleme süreleri bir dizi potansiyel problemi beraberinde getirir. Örneğin, uzun bekleme süreleri, acil tıbbi müdahale gerektiren hastaların tedavi ve bakımının gecikmesine neden olabilir. Özellikle hayati tehlike arz eden durumlarda hızlı müdahale, hastaların sağkalım şansını artırabilir. Uzun bekleme süreleri, durumu hızla kötüleşebilecek hastalar için ciddi sonuçlara yol açabilir. YBÜ’de uzun süre bekleyen hastalar, potansiyel komplikasyon riski altında olabilir. Örneğin, solunum güçlüğü yaşayan bir hastanın solunum desteği alması gerekiyorsa ve bu desteği beklemek zorunda kalıyorsa, solunum yetmezliği riski artabilir. Gerekli tedavi ve destek geciktikçe, hastaların durumu daha da kötüleşebilir. Uzun bekleme süreleri hastalara ya da yakınlarına sadece fiziki değil psikolojik yönden de zarar verebilir. Bekleme süreleri, endişe, kaygı ve belirsizlik hissiyatını artırabilir. Hasta ve ailelerinin memnuniyetsizlikleri arttıkça, hasta-doktor iletişimi ve hasta tatmini olumsuz etkilenebilir. Sonuç olarak, yoğun bakım ünitesinde uzun bekleme süreleri ciddi sorunlara yol açabilir. Acil tıbbi müdahalelerin gecikmesi, komplikasyon riskinin artması, kaynakların etkin kullanılmaması gibi problemler ortaya çıkar. Bu noktada çalışmanın temel amacını oluşturan, YBÜ’lerde kuyruğa izin vermeyen gerekli yatak kapasitesinin ileri düzeyde planlanabilmesinin pratikteki önemi ortaya çıkmaktadır.

Bu çalışmada odaklanılan Pediatri ve Reanimasyon YBÜ’leri için temel bilgiler şu şekildedir. Pediatri YBÜ’ye toplanan bir yıllık veri dikkate alındığında toplam 758, Reanimasyon YBÜ’ye ise, 254 hasta gelmiştir. Pediatri YBÜ’de 13 yatak bulunurken, Reanimasyon YBÜ’de bu sayı 8’dir.

Yatak sayısının belirlenmesi temelde bir Endüstri Mühendisliği kapasite problemidir. Bu problemin ele alınması için, simülasyon modelinin nasıl kurulacağı, girdi parametrelerinin nasıl belirleneceğine dair literatür çalışmaları incelendi ve özellikle istatistiksel olarak çarpık olarak nitelendirilen hastaların YBÜ’ye GAS ve YBÜ’deki SS verisi parametre tahmini için ileri istatistik yöntemleri araştırıldı.

Bu çalışmada, iki YBÜ'ye gelen tüm hastalara bekleme süresi olmaksızın hizmet verecek yatak kapasitesini belirlemek için simülasyon modeli geliştirildi. Kurulan simülasyon modelinin kullanılabilmesi için öncelikle YBÜ'ye gelen hastaların GAS ve SS verisinin istatistiksel olarak tanımlanması gerekmektedir. Bu aşamadaki temel sorun, GAS ve SS verisinin oldukça çarpık olması problemidir. Bu nedenle, GAS ve SS parametrelerini tanımlamak için literatürden faydalanılarak girdi parametresi belirlemede uygun olduğu ve istatistiksel olarak güçlü olduğu tespit edilen karma dağılım yaklaşımı kullanıldı ve bu parametreler simülasyon modeline girdi olarak dahil edildi. Bilinen kadarıyla bu, YBÜ'ler için yatak kapasitesi problemini belirlemek için karma dağılımı ve simülasyon modelleme yaklaşımlarının bir kombinasyonunu kullanan ilk çalışmalardan biridir.

Geliştirilen simülasyon modelinde her bir YBÜ ayrı ayrı değerlendirildi. Simülasyon modeli yardımıyla, hastaların sistemde beklemeden hizmet görebileceği yatak sayısı belirlenmiştir. Mevcut yatak sayıları kullanılarak elde edilen simülasyon sonuçlarına göre; sistemin doluluk oranı Pediatri YBÜ ve Reanimasyon YBÜ için sırasıyla %99 ve %91'dir. Bu doluluk oranı, literatürde tavsiye edilen %70-75 seviyesinden oldukça fazladır. Bu ünitelere gelen hastalar özellikle Pediatri YBÜ'de çok uzun süre kuyrukta beklemek zorunda kalmaktadır. Böylece, her iki YBÜ için de mevcut yatak sayısının yetersiz olduğu tespit edildi. Geliştirilen simülasyon modelinde üniteye gelen tüm hastaların bekleme süresi olmaksızın hizmet görebilmesi için yatak sayıları sistematik olarak arttırıldı ve iki YBÜ için gereken yatak sayısı belirlendi.

Çukurova Üniversitesi Balcalı Eğitim ve Araştırma Hastanesi Pediatri YBÜ'de mevcut olan 13 yatak sayısının, tüm hastalara bekleme süresi olmaksızın hizmet verebilmesi için 30'a, Reanimasyon YBÜ'de mevcutta bulunan 8 yatak, 19'a yükseltilmesi gerektiği simülasyon modeli sonucunda tespit edildi. Böylece, yüksek kritik öneme sahip bu ünitelerde hastalara en hızlı şekilde hizmet verilebilecektir. Pediatri YBÜ'de 25, Reanimasyon YBÜ'de ise 15 yatak olduğu durumda ise, yıl boyunca gelen tüm hastaların sadece %1,5'i reddedilmektedir.

Bu çalışmanın temel katkısı, yüksek oranda çarpık olan gözlemlerin karma dağılım modelleri kullanılarak yüksek doğrulukla modellenmesi ve simülasyon modelinde girdi parametreleri olarak kullanılabilmesi ve böylece daha güvenilir sonuçlara ulaşılabilmesidir. Ek olarak, bu çalışma, seçilen her karar stratejisi için GAS ve SS verisinin dağılım modellenmesi ve sağlık sistemlerinde yatak kapasitesi gibi kaynak planlamasının daha da geliştirilmesine yardımcı olur.

5.2. Sonraki Çalışmalar için Öneriler

Hastaneler, belirsizliğin çok yüksek olması nedeniyle modellenmesi en zor sistemlerden biridir. Gerçek sistemde, hastane ortamı günlük ve hatta dakika dakika değişikliklere tabidir. Uzun bekleme süreleri, hastaların önceliklendirilmesinde sorunlara neden olabilir. Acil müdahale gerektiren hastaların beklemesi, daha az acil durumları olan hastaların beklemesine yol açabilir. Bu

durumda, hastaların ihtiyalarına gre uygun nceliklendirme yapılamaz ve saėlık hizmetlerinin adil bir ekilde daėıtılması engellenebilir.

Bu alıřmada geliřtirilen simlasyon modeli, sistemin dinamiklerini deėiřtirebilecek acil durum hastalarını veya personel seviyelerini dikkate almamıřtır. Bu nedenle, gelecekteki alıřmalar, yalnızca yatak sayısını deėil, aynı zamanda tıbbi personel ve ekipman sayısını ve/veya hastaların nceliklendirilmesini de dikkate alan ve bu yaklařım kullanılarak yeniden optimize edilebilecek bir sistem eklinde tasarlanabilir. Gelecekte yapılacak alıřmalarda sisteme, artan nfus nedeniyle deėiřen hasta sayıları da entegre edilebilir. Model, blgede bulunan diėer hastaneler de gz nne alınarak geniřletilebilir ya da hastane ii farklı YB'lere transfere msaade eden bir sisteme dnřtrlebilir. Bylece bazı durumlarda yatak sayısı arttırılmadan, kabul edilen seviyelerdeki bir bekleme sresi iersinde hastalara gereken hizmet saėlanabilir. Eklenen yatak maliyetleri eklenerek optimal bir dzey bulunmaya alıřılabilir. Tm bu alıřmalar yapılırken, girdi parametresi olarak karma daėılımların kullanılması, arpık olan verisi istatistiksel olarak iyi bir uyumda temsil edecek ve tm bu belirsizlikler iersinde modelin gerek sisteme biraz daha yakınlařtıracaktır.

Bu alıřmada sadece hastanedeki iki YB ele alınmıřtır. Ancak genelleřtirilmiř bu model minr deėiřikliklerle pek ok sistemi temsil edebilir.



KAYNAKLAR

- Açıkgöz, İ., 2007. Sonlu Karma Dağılımlarda Parametre Tahmini. Doktora Tezi. Ankara Üniversitesi, Ankara, 141 s.
- Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6): 716-723. doi:[10.1109/TAC.1974.1100705](https://doi.org/10.1109/TAC.1974.1100705)
- Akalın, H. E., 2002. Yoğun Bakımda Hizmet Kalitesi. Yoğun Bakım Dergisi 2(1): 35-38.
- Akar, S., ve Uğurlu, B., 2005. Forecasting Production of Fossil Fuel Sources in Turkey using a Comparative Regression and ARIMA Mode. Energy Policy, 34(18): 3836-3848.
- Akçalı, E., Côté, M. J., ve Lin, C., 2006. A Network Flow Approach to Optimizing Hospital Bed Capacity Decisions. Health Care Management Science, 9(4): 391-404. doi:[10.1007/s10729-006-0002-4](https://doi.org/10.1007/s10729-006-0002-4)
- Andreas, V., ve Patrick, F., 2011. Recommendations on Basic Requirements for Intensive Care Units: Structural and Organizational Aspects. Intensive Care Medicine, 37(10): 1575-1587. doi:[10.1007/s00134-011-2300-7](https://doi.org/10.1007/s00134-011-2300-7)
- Antelo, M., Santias, F., ve Calvo, A., 2015. Bed Capacity and Surgical Waiting Lists: A Simulation Analysis. European Journal of Government and Economics, 4(2): 118-133.
- Antmen, Z. F., ve Oğulata, S. N., 2013. The Capacity Planning of Intensive Care Units via Simulation: A Case Study in University Hospital. Centre for Environment Social and Economic Research, 51(21).
- Arabie, P., Hubert, L., ve De Soete, G., 1996. Clustering and Classification. World Scientific, Singapur, 490 s.
- Atienza, N., Garcia-Heras, J., Munoz-Pichardo, J., ve Villa, R., 2008. An Application of Mixture Distributions in Modelization of Length of Hospital Stay. Statistics in Medicine, 27: 1403-1420.
- Banks, J., Carson, J., Nelson, B., ve Nicol, D., 2005. Discrete-Event System Simulation. Prentice Hall, New Jersey, 662 s.
- Belciug, S., ve Gorunescu, F., 2015. Improving Hospital Bed Occupancy and Resource Utilization Through Queuing Modeling and Evolutionary Computation. Journal of Biomedical Informatics, 53: 261-269. doi:[10.1016/j.jbi.2014.11.010](https://doi.org/10.1016/j.jbi.2014.11.010)
- Benzeval, M., ve Judge, K., 1994. The Determinants of Hospital Utilization: Implications for Resource Allocation in England. Health Economics, 3(2): 105-116. doi:[10.1002/hec.4730030207](https://doi.org/10.1002/hec.4730030207)
- Bergman, L., ve Magnusson, D., 1997. A Person-oriented Approach in Research on Developmental Psychopathology. Dev Psychopathol, 291-319.
- Bhattacharya, C. G., 1967. A Simple Method of Resolution of a Distribution Into Gaussian Components. Biometrics, 23: 115-135.

- Blank, S., 1968. Descriptive Statistics. Michigan: Appleton Century Crofts.
- Bozdoğan, H., 1994. Mixture-Model Cluster Analysis Using Model Selection Criteria and a New Information Measure of Complexity. In Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach. Kluwer Academic Publisher, 69-113.
- Bulut, Ş., 2006. Orta Ölçekli Bir İşletmede Talep Tahmin Yöntemlerinin Uygulanması. Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Kırıkkale, 116 s.
- Burchardi, H., ve Moerer, O., 2001. Twenty-Four Hour Presence of Physicians in the ICU. Critical Care, 5(3): 131-137. doi:[10.1186/cc1012](https://doi.org/10.1186/cc1012)
- Buschiazzo, M., Mula, J., ve Campuzano-Bolarin, F., 2020. Simulation Optimization for the Inventory Management of Healthcare Supplies. International Journal of Simulation Modelling, 19(2): 255-266. doi:[10.2507/IJSIMM19-2-514](https://doi.org/10.2507/IJSIMM19-2-514)
- Cassie, R., 1954. Some Uses of Probability Paper in the Analysis of Size Frequency Distributions. Australian Journal of Marine and Freshwater Research, 5: 513-522.
- Chhikara, R., ve Register, D. T., 1979. A Numerical Classification Method for Partitioning of a Large Dimensional Mixed Data Set. Technometrics, 21: 531-538.
- Cleary, P. G., 1991. Variations in Length of Stay and Outcomes for Six Medical and Surgical Conditions in Massachusetts and California. JAMA, 266: 73-85.
- Cohen, A., 1967. Estimation in Mixture of Two Normal Distributions. Technometrics, 9, 15-28.
- Çalış, N., 2005. Karma Dağılım Modellerinde Bileşen Sayısını Tahmin Etmek İçin Yöntemler. Yüksek Lisans Tezi, Çukurova Üniversitesi, Adana, 121 s.
- Deb, P., ve Trivedi, P., 1998. Demand for Medical Care by the Elderly: a Finite Mixture Approach. Journal of Applied Econometrics, 12(3): 313-336.
- Dempster, A., Laird, N., ve Rubin, D., 1977. Maximum Likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society Series B (Statistical Methodology) 39: 1-38.
- DuToit, S. H., Steyn, A. G., ve Stumpf, R. H., 1986. Graphical Exploratory Data Analysis. Springer, New York, 314 s.
- Eirola, E., ve Lendasse, A., 2013. Gaussian Mixture Models for Time Series Modelling, Forecasting, and Interpolation. Lecture Notes in Computer Science
- Ergün, S., ve Şahin, S., 2017. İşletme Talep Tahmini Üzerine Literatür Araştırması. Ulakbilge, 5(10): 469-487.
- Erlang, A. K., 1909. The Theory of Probabilities and Telephone Conversations. Nyt Tidsskrift for Matematik B., 20: 33-39.
- Fisher, R. A., 1922. On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society A, Londra, 61 s.

- Fowlkes, E., 1979. Some Methods for Studying the Mixture of Two Normal (Lognormal) Distributions. *Journal of Classification*, 16: 297-306.
- Fraley, C., ve Raftery, A. E., 1998. How Many Clusters? Which Clustering Method?-Answers via Model-Based Cluster Analysis. *Computer Journal*, 41, 578-588.
- Fryer, J. G., ve Robertson, C. A., 1972. A Comprasion of Some Methods for Estimating Mixed Normal Distributions. *Biometrika*, 59: 639-648.
- Garg, L., McClean, S., Meenan, B., El-Darzi, E., ve Millard, P., 2009. Clustering Patient Length of Stay Using Mixtures of Gaussian Models and Phase Type Distributions. 22nd IEEE International Symposium on Computer-Based Medical Systems, 1-7. doi:[10.1109/CBMS.2009.5255245](https://doi.org/10.1109/CBMS.2009.5255245)
- Gorunescu, F., McClean, S., ve Millard, P., 2002. A Queueing Model for Bed Occupancy Management and Planning of Hospitals. *Journal of the Operational Research Society*, 53(1): 19-24. doi:[10.1057/palgrave/jors/2601244](https://doi.org/10.1057/palgrave/jors/2601244)
- Green, L. V., 2004. Capacity Planning and Management in Hospitals. Brandeau, M.L., Sainfort, F., Pierskalla, W.P. *Operations Research and Health Care. International Series in Operations Research & Management Science*, Springer, Boston. doi:[10.1007/1-4020-8066-2_2](https://doi.org/10.1007/1-4020-8066-2_2)
- Griffiths, J., Price-Lloyd, N., Smithies, M., ve Williams, J., 2006. A Queueing Model of Activities in an Intensive Care Unit. *IMA Journal of Management Mathematics*, 17(3): 277-288. doi:[10.1093/imaman/dpi042](https://doi.org/10.1093/imaman/dpi042)
- Günal, M., ve Pidd, M., 2006. Understanding Accident and Emergency Department Performance using Simulation. Monterey, CA, IEE: In: Perrone LF, Wieland FP, Liu J, Lawson BG, Nicol DM and Fujimoto RM (eds). *Proceedings of the 2006 Winter Simulation Conference*.
- Gürsakal, N., 2018. Çıkarımsal İstatistik. Dora Yayıncılık, Bursa, 526s.
- Hadavandi, E., Shavandi, H., Ghanbari, A., ve Abbasian-Naghneh, S., 2012. Developing a hybrid artificial intelligence model for outpatient visits forecasting in hospitals. *Applied Soft Computing*, 12(2): 700-711.
- Harchol-Balter, M., 2012. Scheduling: Preemptive, Size-Based Policies. M. Harchol-Balter içinde, *Performance Modeling and Design of Computer Systems: Queueing Theory in Action* (s. 508-517). Cambridge: Cambridge University Press. doi:[10.1017/CBO9781139226424.040](https://doi.org/10.1017/CBO9781139226424.040)
- Harding, J., 1948. The use of probability paper for the graphical analysis of polymodal frequency distributions. *Journal of the Marine Biological Association*, 28: 141-153.
- Harper, P., ve Shahani, A., 2002. Modelling for the Planning and Management of Bed Capacities in Hospitals. *Journal of the Operational Research Society*, 53: 11-18. doi:[10.1057/palgrave/jors/2601278](https://doi.org/10.1057/palgrave/jors/2601278)
- Hogg, R. V., ve Craig, A. T., 1989. *Introduction to Mathematical Statistics*. Macmillan Publishing Company, New York, 225 s.

- Horrocks, P., 1986. The Components of a Comprehensive District Health Service for Elderly People – a Personal View. *Age Ageing*, 15(6): 321-342. doi:[10.1093/ageing/15.6.321](https://doi.org/10.1093/ageing/15.6.321)
- Ickowicz, A., ve Sparks, R., 2016. Modelling Hospital Length of Stay using Convolutional Mixture Distributions. *Statistics in Medicine*, 36(1): 122-135. doi:[10.1002/sim.7135](https://doi.org/10.1002/sim.7135)
- Johnson, N. L., Kotz, S., ve Balakrishnan, N., 1994. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. Wiley ve Sons, New York, 756 s.
- Johnson, R. A., ve Wichern, D. W., 1982. *Applied Multivariate Statistical Analysis*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 393 s.
- Jone, A. M., 2010. *Models for Healthcare*. York Üniversitesi, York, 32 s.
- Karaman, E., 2019. Optimal Ölçekleme Teknikleri ve bir Uygulama. Doktora Tezi, İstanbul Üniversitesi, İstanbul, 108 s.
- Kendall, D. G., 1953. Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain. *The Annals of Mathematical Statistics*, 24(3): 338.
- Keribin, C., 2000. Consistent estimation of the order of mixture models. *The Indian Journal of Statistics, Series A*, 62(1): 49-66. doi:[10.2307/25051289](https://doi.org/10.2307/25051289)
- Kessler, D., ve McDowell, A., 2012. *Introducing the FMM Procedure for Finite Mixture Models*. SAS Institute Inc, Cary, North Carolina, 22 s.
- Kingman, J. F., 1961. The single server queue in heavy traffic . *Mathematical Proceedings of the Cambridge Philosophical Society*, 57(4): 902-904. doi:[10.1017/S0305004100036094](https://doi.org/10.1017/S0305004100036094)
- Kingman, J. F., 2009. The first Erlang century—and the next. *Queueing Systems*, 63: 3-4. doi:[10.1007/s11134-009-9147-4](https://doi.org/10.1007/s11134-009-9147-4)
- Klute, B., Homb, A., Chen, W., ve Stelpflug, A., 2019. Predicting Outpatient Appointment Demand Using Machine Learning and Traditional Methods. *Journal of Medical Systems* 43:288, 278-288.
- Kokangül, A., Akcan, S., & Narlı, M., 2017. Optimizing nurse capacity in a teaching hospital neonatal intensive care unit. *Health Care Management Science*, 20: 276-285. doi:[10.1007/s10729-015-9352-0](https://doi.org/10.1007/s10729-015-9352-0)
- Kokangül, A., 2008. A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. *Computer Methods and Programs in Biomedicine*, 90(1): 56–65. doi:[10.1016/j.cmpb.2008.01.001](https://doi.org/10.1016/j.cmpb.2008.01.001)
- Köroğlu, F., Bayrakçı, B., Dursun, O., Kendirli, T., Yıldıztaş, D., ve Karaböcüoğlu, M., 2006. A Guide for Pediatric Intensive Care Units: Propositions From Pediatric Emergency Medicine and Intensive Care Society. *Türk Pediatri Arşivi*, 41: 139-145.
- Kraemer, W., ve Langenbach-belz, M., 1976. Approximate Formulae for the Delay in the Queueing System GI/G/1.

- Kurtulmuş, S., 1998. Sağlık Ekonomisi ve Hastane Yönetimi. Değişim Dinamikleri Yayınları, İstanbul, 185 s.
- Lakshmi, R. V., ve Vaidyanathan, V. S., 2016. Parameter Estimation in Gamma Mixture Model using Normal-based Approximation. *Journal of Statistical Theory and Applications* 15: 25-35.
- Larson, H. J., 1982. *Introduction to Probability Theory and Statistical Inference*. Wiley ve Sons, New York, 656 s.
- Law, A., 2007. *Simulation Modeling and Analysis*. McGraw-Hill, Arizona, 400 s.
- Lee, A., Ng, A., ve Yau, K., 2001. Determinants of maternity length of stay: a gamma mixture risk-adjusted model. *Health Care Management Science*, 4(4): 449-455. doi:[10.1023/a:1011810326113](https://doi.org/10.1023/a:1011810326113)
- Lee, A., Xiao, J., Codde, J., ve Ng, A., 2002. Public Versus Private Hospital Maternity Length of Stay: A Gamma Mixture Modelling Approach. *Health Services Management Research* 15(1): 46-54.
- Lee, M. L., I., Park, I., Park, D. U., ve Park, C., 2017. Constrained Ranking and Selection for Operations of an Emergency Department. *International Journal of Simulation Modelling*, 16(4): 563-575. doi:[10.2507/IJSIMM16\(4\)1.388](https://doi.org/10.2507/IJSIMM16(4)1.388)
- Little, J. D., 1961. A proof of the queueing formula $L = \lambda W$. *Operations Research*, 9(3): 383-387. doi:[10.1287/opre.9.3.383](https://doi.org/10.1287/opre.9.3.383)
- Mackay, M., ve Lee, M., 2005. Choice of Models for the Analysis and Forecasting of Hospital Beds. *Health Care Management Science*, 8(3): 221-230.
- Maharlou, H., Niakan Kalhori, S., Shahbazi, S., ve Ravangard, R., 2018. Predicting Length of Stay in Intensive Care Units after Cardiac Surgery: Comparison of Artificial Neural Networks and Adaptive Neuro-fuzzy System. *Healthcare Informatics Research*, 24(2): 109-117. doi:[10.4258/hir.2018.24.2.109](https://doi.org/10.4258/hir.2018.24.2.109)
- Makridakis, S., Wheelwright, S. C., ve Hyndman, R. J., 1998. *Forecasting: methods and applications*. John Wiley & Sons, Hoboken, 656s.
- Maman, S., 2009. *Uncertainty in the Demand for Service: The Case of Call Centers and Emergency Departments*. Haifa: Israel Institute of Technology.
- Mann, H. B., ve Whitney, D. R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1): 50-60.
- Marazzi, A., Paccaud, F., Ruffieux, C., ve Beguin, C., 1998. Fitting the Distributions of Length of Stay by Parametric Models. *Medical Care*, 36(6): 915-927. doi:[10.1097/00005650-199806000-00014](https://doi.org/10.1097/00005650-199806000-00014)
- Markay, M., ve Lee, M., 2005. Choice of Models for the Analysis and Forecasting of Hospital Beds. *Health Care Management Science*, 8: 221-230.

- Markov, A. A., 1906. Extension of the law of large numbers to dependent quantities. *Izvestiia Fiz.-Matem. Obsch. Kazan Univ.*, 15(1906): 135–156.
- McClean, S., ve Millard, P., 1993. Patterns of Length of Stay after Admission in Geriatric Medicine: An Event History Approach. *The Statistician*, 42(3): 263–274. doi:[10.2307/2348804](https://doi.org/10.2307/2348804)
- McLachlan, G., ve Peel, D., 2000. *Finite Mixture Models*. Wiley ve Sons, New York, 456 s.
- Mendenhall, W., Wackerly, D. D., ve Scheaffer, R., 1990. *Mathematical Statistics with Applications*. PWS-Kent Publishing Company, Boston, 939 s.
- Meyer, A., Zverinski, D., Pfahringer, B., Kempfert, J., Kuehne, T., Sündermann, S., Eickhoff, C., 2018. Machine learning for Real-Time Prediction of Complications in Critical Care: A Retrospective Study. *Lancet Respiratory Medicine*, 6(12): 905-914. doi:[10.1016/S2213-2600\(18\)30300-X](https://doi.org/10.1016/S2213-2600(18)30300-X)
- Millard, P., 1988. *Geriatric Medicine: A New Method of Measuring Bed Usage a Theory for Planning*. Yüksek Lisans Tezi, Londra Üniversitesi, Londra.
- Mills, A., ve Gilson, L., 1988. *Health Economics for Developing Countries: A Survival Kit*. Health Economics&Financing Programme Working Paper Series.
- Montgomery, D., 2009. *Introduction to Statistical Quality Control*. Wiley ve Sons, Inc, Denver, 754 s.
- Ng, S. K., Xiang, L., ve Yau, K. K., 2019. *Mixture Modelling for Medical and Health Sciences*. CRC Press Boca Raton, 300 s.
- NIST/SEMATECH, 2022. *e-Handbook of Statistical Methods*. Measures of Skewness and Kurtosis.
- Ojstersek, R., ve Buchmeister, B., 2021. Simulation Based Resource Capacity Planning with Constraints. *International Journal of Simulation Modelling*, 20(4): 672-683. doi:[10.2507/IJSIMM20-4-578](https://doi.org/10.2507/IJSIMM20-4-578)
- Pallant, J., 2005. *SPSS Survival Manual, Using Graphs to Describe and Explore the Data*. Allen ve Unwin, Sidney, 378s.
- Papi, M., ve Pontecorvi, L., 2015. Hospital Resource Consumption Modelling. *Research in Business and Management*, 1(1): 105-122.
- Pearson, K., 1894. Contributions to the Mathematical Theory of Evolution. *Philosophical Transactions of the Royal Society Series A*, 185, 71-110. doi:[10.1098/rsta.1894.0003](https://doi.org/10.1098/rsta.1894.0003)
- Pollaczek, F., 1930. Über eine Aufgabe der Wahrscheinlichkeitstheorie. I. *Mathematische Zeitschrift*, 32: 64-100. doi:[10.1007/BF01194620](https://doi.org/10.1007/BF01194620)
- Quantin, C., Sauleau, E., Bolard, P., Mousson, C., Kerkri, M., Lecomte Brunet, P., Moreau, T. Ve Dusserre, L., 1999. Modeling of High-cost Patient Distribution within Renal Failure Diagnosis Related Group, *Journal of Clinical Epidemiology*, 52(3): 251-258. doi:[10.1016/S0895-4356\(98\)00164-4](https://doi.org/10.1016/S0895-4356(98)00164-4)

- Rao, C. R., 1952. Advanced Statistical Methods in Biometric Research. Wiley ve Sons, New York, 390 s.
- Rossetti, M., 2021. Simulation Modeling and Arena, 3rd and Open Text Edition, Wiley ve Sons, New York, 719 s.
- Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics*, 461-464.
- Servi, T., 2009. Çok deęişkenli karma dağılım modeline dayalı kümeleme analizi. Doktora tezi, Çukurova Üniversitesi, Adana, 266 s.
- Shorten, A., 1995. Obstetric early discharge versus traditional hospital stay. *Australian Health Review*, 18(2): 19-39.
- Sin, C. Y., ve White, H., 1996. Information criteria for selecting possibly misspecified parametric models. *Journal of Econometrics*, 71(1-2): 207-225. doi:[10.1016/0304-4076\(94\)01701-8](https://doi.org/10.1016/0304-4076(94)01701-8)
- Singh, C., ve Ladusingh, L., 2010. Inpatient Length of Stay: A Finite Mixture Modeling Analysis. *The European Journal of Health Economics*, 11(2): 119-126.
- Singh, N., 2019. Predicting Hospital Length of Stay in Intensive Care Unit. Yüksek Lisans Tezi, Wisconsin-Milwaukee Üniversitesi, Milwaukee.
- Snedecor, G. W., ve Cochran, W. G., 1989. Statistical Methods, 8th Edition. Iowa State University Press, Ames, 503 s.
- Stekler, H. O., 2007. The Future of Macroeconomic Forecasting: Understanding the Forecasting Process. *International Journal of Forecasting*, 23(2): 237-248.
- Stephens, M. A., 1976. Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters. *Annals of Statistics*, 4(2): 357-369. doi:[10.1214/aos/1176343411](https://doi.org/10.1214/aos/1176343411)
- Sundarapandian, V., 2009. 7. Queueing Theory. V. Sundarapandian içinde, *Probability, Statistics and Queueing Theory* (s. 686-749). PHI Learning, Yeni Delfi.
- T.C. Sağlık Bakanlığı, 2008. Yoęun Bakım Üniteleri Standartları Genelgesi.
- Tan, W., ve Chang, W., 1972. Some comparisons of the method of moments and the method of maximum likelihood in estimating parameters of a mixture of two normal densities. *Journal of the American Statistical Association*, 67: 702-708.
- Tarter, M. E., ve Silvers, A., 1975. Implementation and application of bivariate Gaussian mixture decomposition. *Journal of the American Statistical Association*, 70: 47-55.
- Titterington, D., Smith, A., ve Makov, U., 1985. Statistical Analysis of Finite Mixture Distributions. Wiley, New York, 243s.
- TÜİK, 2021. Sağlık Harcamaları İstatistikleri, Sayı 45728.
- Tukey, J. W., 1977. Exploratory Data Analysis. Addison-Wesley. 688s.
- Uncu, N., 2017. Isınma Periyodu Belirleme Yöntemlerinin Etkinliklerinin Analizi. *Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 32(4):201-210.
- Uz, H., 1998. Sağlık Hizmetleri El Kitabı, Sağlık Ekonomisi. Çevik Matbaası, İstanbul, 450s.

- Verburg, I. W., de Keizer, N. F., de Jonge, E., ve Peek, N., 2014. Comparison of Regression Methods for Modeling Intensive Care Length of Stay. *PLoS One*, 9(10). doi:[10.1371/journal.pone.0109684](https://doi.org/10.1371/journal.pone.0109684)
- Veronica, T., ve Bauer, D., 2016. A Note on the Use of Mixture Models for Individual Prediction. *Structural Equation Modeling*, 23(4): 615-631.
- Wang, K., Yau, K., ve Lee, A., 2002. A Hierarchical Poisson Mixture Regression Model to Analyse Maternity Length of Hospital Stay. *Statistics in Medicine*, 21(23): 3639-3654.
- Weibull, W., 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics, Transactions*, 18(3):293-297.
- Weldon, W. F., 1892. Certain correlated variations in *Crangon vulgaris*. *Proceedings of the Royal Society*, 52: 2-21.
- Wengrzik, J., 2012. Parameter Estimation for Mixture Models Given Grouped Data. *Universtat Bremen, Bremen*, 127s.
- Whittle, P., 2002. Applied Probability in Great Britain. *Operations Research*, 50(1): 227-239. doi:[10.1287/opre.50.1.227.17792](https://doi.org/10.1287/opre.50.1.227.17792)
- Wilk, M. B., ve Gnanadesikan, R., 1968. Probability plotting methods for the analysis of data. *Biometrika*, 55:1-17.
- Wright, P. D., ve Bretthauer, K. M., 2010. Strategies for Addressing the Nursing Shortage: Coordinated Decision Making and Workforce Flexibility. *Decision Sciences*, 41(2): 373-401. doi:[10.1111/j.1540-5915.2010.00269.x](https://doi.org/10.1111/j.1540-5915.2010.00269.x)
- Wu, J., Lin, Y., Li, P., Hu, Y., Zhang, L., ve Kong, G., 2021. Predicting Prolonged Length of ICU Stay through machine learning. *Diagnostics*, 11(12):2242. doi:[10.3390/diagnostics11122242](https://doi.org/10.3390/diagnostics11122242)
- Xiao, J., Lee, A., ve Vemuri, S., 1999. Mixture Distribution Analysis of Length of Hospital Stay for Efficient Funding. *Socio-Economic Planning Sciences*, 33: 39-59.
- Yalta, T., 2011. Modellemeye İlişkin Konular-Model Seçim Ölçütleri. *Türkiye Bilimleri Akademisi: Ulusal Açık Ders Malzemeleri Konsorsiyumu*.
- Yavuz, S., 2008. Simülasyonda İstatistiksel Modeller. *Yıldız Teknik Üniversitesi, İstanbul*.
- Yıldırım, C., Yıldırım, S., ve Arı, H. O., 2014. Sağlık Kurumlarında Talep Öngörü Yöntemleri. *Sağlıkta Performans ve Kalite Dergisi*, (8): 77-92.
- Yüzer, A. F., 2006. İstatistik. *Anadolu Üniversitesi, Eskişehir*, 252s.
- Zhou, Z., Matteson, D. S., Woodard, D. B., Henderson, S. G., ve Micheas, A. C., 2015. A Spatio Temporal Point Process Model for Ambulance Demand. *Journal of the American Statistical Association*, 110:509, 6-15. doi: [10.1080/01621459.2014.941466](https://doi.org/10.1080/01621459.2014.941466)
- Zhu, Z., Hen, B. H., ve Teow, K. L., 2012. Estimating ICU bed capacity using discrete event simulation. *International Journal of Health Care Quality Assurance*, 25(2): 134-144. doi:[10.1108/09526861211198290](https://doi.org/10.1108/09526861211198290)

ÖZGEÇMİŞ

Selin SARAÇ GÜLERYÜZ, ilk ve orta öğrenimini İstanbul, lise öğrenimini Hatay'da tamamladı. 2009 yılında Mersin Üniversitesi Makine Mühendisliği bölümünü kazandı. 2014 yılında mezun oldu ve aynı sene Toros Üniversitesi Endüstri Mühendisliği bölümünde Araştırma Görevlisi olarak çalışmaya başladı. 2015 yılında Çukurova Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı'nda yüksek lisans eğitime başladı. 2018 yılında yüksek lisans eğitimini tamamladıktan sonra aynı Anabilim Dalı'nda doktora eğitime başladı. Halen Toros Üniversitesi Endüstri Mühendisliği bölümünde Araştırma Görevlisi kadrosunda görevini sürdürmektedir.

