

**DAĞILIM PARAMETRELERİNİN  
TAHMİNİNDE KULLANILAN  
ROBUST YÖNTEMLER**

İbrahim ARIK  
Yüksek Lisans Tezi

İstatistik Anabilim Dalı

Ocak – 2014

## JÜRİ VE ENSTİTÜ ONAYI

İbrahim ARIK'ın "Dağılım Parametrelerinin Tahmininde Kullanılan Robust Yöntemler" başlıklı İstatistik Anabilim Dalındaki, Yüksek Lisans Tezi 26.12.2013 tarihinde, aşağıdaki jüri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

	Adı Soyadı	İmza
Üye (Tez Danışmanı) :	Doç. Dr. Yeliz MERT KANTAR	
Üye	: Prof. Dr. Embiya AĞAOĞLU	
Üye	: Prof. Dr. Birdal ŞENOĞLU	

Anadolu Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun  
..... tarih ve ..... sayılı kararıyla onaylanmıştır.

Enstitü Müdürü



## ÖZET

Yüksek Lisans Tezi

### DAĞILIM PARAMETRELERİNİN TAHMİNİNDE KULLANILAN ROBUST YÖNTEMLER

İbrahim ARIK

Anadolu Üniversitesi  
Fen Bilimleri Enstitüsü  
İstatistik Anabilim Dalı

Danışman: Doç. Dr. Yeliz MERT KANTAR  
2013, 104 sayfa

En çok olabilirlik ve momentler tahmincileri dağılımsal parametrelerin tahmininde en yaygın kullanılan yöntemlerdir. Ancak bilinmektedir ki bu yöntemler veride aykırı değer olması durumunda güvenilir olmayan sonuçlar verebilmektedir.

Bu tez çalışmasının amacı, veride aykırı değer olması durumunda, dağılımsal parametrelerin tahmininde kullanılan çeşitli sağlam (robust) tahmincilerin incelenmesidir. Ele alınan tahminciler; doğrusal regresyon modeli için robust tahminciler, kantillere dayalı robust tahminci, konum-ölçek ailesine dönüştürülebilen dağılımlar için robust tahminciler, medyanlar metodu, kırılmış momentler metodu şeklinde sıralanabilir.

Ele alınan robust tahminciler Pareto ve Weibull dağılımlarının parametrelerinin tahmini için ayrıca elde edilmiş olup veride aykırı değerler olması ve olmaması durumunda bu tahmincilerin performansı kurgulanan benzetim çalışması yardımıyla incelenmiştir. Ayrıca gerçek yaşamdan alınan iki farklı veri seti üzerinde ele alınan tüm tahminciler uygulanmış ve elde edilen sonuçlar tartışılmıştır.

**Anahtar Kelimeler :** Dağılımsal Parametre, Sağlamlık, Robust Tahminci, Pareto Dağılımı, Weibull Dağılımı, Şekil Parametresi

## ABSTRACT

Master of Science Thesis

### ROBUST ESTIMATION METHODS FOR ESTIMATING DISTRIBUTIONAL PARAMETERS

İbrahim ARIK

Anadolu University  
Graduate School of Sciences  
Statistics Program

Supervisor: Assoc. Prof. Yeliz MERT KANTAR  
2013, 104 pages

The most commonly used methods in the estimation of distributional parameters are the maximum likelihood and moments estimators. However, it is well-known that these estimators can be very unreliable in the case of outliers in the data.

The objective of this thesis is examination of the robust estimators, which are used in the estimation of distributional parameters, in the case of outliers in the data. The considered robust estimators are given as follow: robust estimators for the linear regression model, the estimator based on quantiles, the robust estimators for transformed location-scale family, the method of medians, the method of trimmed moments.

The considered robust estimators are also obtained for estimating the parameters of the Weibull and Pareto distributions and the performance of the considered robust estimators are evaluated via a simulation study for the case of outliers and non-outliers in data. Also, the considered estimators are applied on two real life data and the obtained results are discussed.

**Keywords :** Distributional Parameter, Robustness, Robust Estimator, Pareto distribution, Weibull distribution, Shape parameter

## TEŞEKKÜR

Yüksek lisans eğitimi ve tezin hazırlanması sürecinde değerli bilgilerini ve tecrübesini benimle paylaşan danışman hocam Sayın *Doç Dr. Yeliz Mert Kantar*'a,

Çalışma süresince destekleriyle yanımda olan *Prof. Dr. Embiya Ağaoğlu*, *Prof. Dr. Memmedaga Memmedli* ve *Doç. Dr. İlhan Usta*'ya,

Bu tezin yazımı esnasında bana zaman ayıran ve yardımlarını esirgemeyen *Araş. Gör. Dr. Şükrü Acıtaş*'a,

Beni hayatımın her aşamasında destekleyen aileme; annem *Safiye Arık*, babam *Kadir Arık* ve ağabeyim *Hikmet Arık*'a,

Sabrı ve özverisi ile her zaman yanımda olduğunu hissettiğim eşim *Özlem Arık*'a,

En içten teşekkürü bir borç bilirim.

İbrahim ARIK

Aralık 2013

## İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
ÇİZELGELER DİZİNİ	vii
ŞEKİLLER DİZİNİ	ix
1. GİRİŞ	1
2. DAĞILIM PARAMETRELERİNİN TAHMİNİ	4
2.1. En Küçük Kareler (Least Squares-LS) Tahmincisi . . . . .	4
2.2. Momentler Tahmincisi . . . . .	5
2.3. En Çok Olabilirlik (Maximum Likelihood-ML) Tahmincisi . . . . .	6
3. ROBUST TAHMİNLEME	9
3.1. Robust Tahminciler için Değerlendirme Kriterleri . . . . .	9
3.1.1. Etki Fonksiyonu (Influence Function-IF) . . . . .	10
3.1.2. Kırılma Noktası (Breakdown Point-BP) . . . . .	11
3.2. Konumun Robust Tahmini . . . . .	12
3.2.1. M-Tahminciler . . . . .	14
3.2.2. Kırılmış Ortalama . . . . .	18
3.3. Ölçeğin Robust Tahmini . . . . .	19
3.3.1. Medyan Mutlak Sapma (Median Absolute Deviation (MAD) Tahmincisi . . . . .	19
3.3.2. $S_n$ Tahmincisi . . . . .	20
3.3.3. $Q_n$ Tahmincisi . . . . .	21
4. DAĞILIM PARAMETRELERİNDE ROBUST YÖNTEMLER	22
4.1. Doğrusal regresyon modeli yardımıyla robust tahminleme . . . . .	22
4.1.1. M-Tahminciler . . . . .	23
4.1.2. MM Tahmincisi . . . . .	26

4.1.3.	En Küçük Mutlak Sapmalar (Least Absolute Deviation-LAD) Tahmincisi . . . . .	27
4.1.4.	En Küçük Medyan Kareler (Least Median of Squares-LMS) Tahmincisi . . . . .	27
4.1.5.	Kırılmış En Küçük Kareler (Least Trimmed Squares-LTS) Tahmincisi . . . . .	28
4.1.6.	Tekrarlı Medyanlar (Repeated Median-RepMed) Tahmincisi . . . . .	28
4.1.7.	Theil Tahmincisi . . . . .	29
4.2.	Kantillere Dayalı Tahminleme . . . . .	29
4.3.	Konum-Ölçek Ailesine Dönüştürülebilen Dağılımlar için Robust Tahminleme . . . . .	30
4.4.	Medyanlar Metodu . . . . .	31
4.5.	Kırılmış Momentler Metodu (Method Of Trimmed Moments - MTM) . . . . .	32
<b>5.</b>	<b>DAĞILIMLAR İÇİN TAHMİNCİLERİN HESAPLANMASI</b>	<b>34</b>
5.1.	Weibull Dağılımı . . . . .	34
5.1.1.	Doğrusal Regresyon Modeli Yardımıyla Weibull Dağılımının Parametrelerinin Tahmini . . . . .	35
5.1.2.	Weibull Dağılımı İçin Kantillere Dayalı Tahminleme . . . . .	36
5.1.3.	Konum-Ölçek Ailesine Dönüştürerek Weibull Dağılımının Parametrelerinin Tahmini . . . . .	37
5.1.4.	Medyanlar Metodu ile Weibull Dağılımının Parametrelerinin Tahmini . . . . .	37
5.1.5.	Kırılmış Momentler Metodu ile Weibull Dağılımının Parametrelerinin Tahmini . . . . .	38
5.2.	Pareto Dağılımı . . . . .	40
5.2.1.	Doğrusal Regresyon Modeli Yardımıyla Pareto Dağılımının Parametrelerinin Tahmini . . . . .	41
5.2.2.	Pareto Dağılımı İçin Kantillere Dayalı Tahminleme . . . . .	42

5.2.3.	Konum-Ölçek Ailesine Dönüştürerek Pareto Dağılımının Parametrelerinin Tahmini . . . . .	43
5.2.4.	Medyanlar Metodu ile Pareto Dağılımının Parametrelerinin Tahmini . . . . .	43
5.2.5.	Kırılmış Momentler Metodu ile Pareto Dağılımının Parametrelerinin Tahmini . . . . .	44
<b>6.</b>	<b>SİMULASYON ÇALIŞMASI</b>	<b>47</b>
6.1.	Aykırı Değer Bulunmadığı Durumda Tahmincilerin Performanslarının İncelenmesi . . . . .	48
6.1.1.	Weibull Dağılımı . . . . .	48
6.1.2.	Pareto Dağılımı . . . . .	59
6.2.	Aykırı Değer Bulunduğu Durumda Tahmincilerin Performanslarının İncelenmesi . . . . .	70
6.2.1.	Weibull Dağılımı . . . . .	70
6.2.2.	Pareto Dağılımı . . . . .	80
<b>7.</b>	<b>UYGULAMALAR</b>	<b>91</b>
<b>8.</b>	<b>SONUÇLAR</b>	<b>98</b>
	<b>KAYNAKLAR</b>	<b>100</b>



## ÇİZELGELER DİZİNİ

2.1. Üstel (1)'den üretilmiş veri . . . . .	7
3.1. Kepekli un verisi . . . . .	13
3.2. M-Tahminciler . . . . .	17
3.3. $n \leq 9$ için $c_n$ değerleri . . . . .	20
3.4. $n \leq 9$ için $d_n$ değerleri . . . . .	21
4.1. Uygulamada Kullanılan Çeşitli Dağılımlar için $Med(X)$ ve $MAD(X)$ değerleri . . . . .	31
4.2. Çizelge 4.1 teki dağılımlar için robust tahminciler . . . . .	31
5.1. Şekil parametresinin farklı değerleri için Weibull dağılımının çarpıklık ve basıklık değerleri . . . . .	35
5.2. $F_0$ : Log-Weibull Dağılımı için $c_1$ değeri . . . . .	39
5.3. $F_0$ : Log-Weibull Dağılımı için $c_2$ değeri . . . . .	39
5.4. Şekil parametresinin farklı değerleri için Pareto dağılımının çarpıklık ve basıklık değerleri . . . . .	41
5.5. $F_0$ : Pareto Dağılımı için $c_1$ değeri . . . . .	46
5.6. $F_0$ : Pareto Dağılımı için $c_2$ değeri . . . . .	46
6.1. Weib ( $\gamma = 1$ , $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	49
6.2. Weib ( $\gamma = 1$ , $\alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	51
6.3. Weib ( $\gamma = 1$ , $\alpha = 2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	53
6.4. Weib ( $\gamma = 1$ , $\alpha = 4$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	55
6.5. Weib ( $\gamma = 1$ , $\alpha = 8$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	57
6.6. Pareto ( $\sigma = 1$ , $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	60
6.7. Pareto ( $\sigma = 1$ , $\alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	62

6.8. Pareto ( $\sigma = 1, \alpha = 2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	64
6.9. Pareto ( $\sigma = 1, \alpha = 4$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	66
6.10. Pareto ( $\sigma = 1, \alpha = 8$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	68
6.11. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 0.2, \alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	72
6.12. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	74
6.13. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	76
6.14. Weib ( $\gamma = 1, \alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri (%10 100 değeri aykırı değer olarak atılmış) . . . . .	78
6.15. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	82
6.16. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 0.2, \alpha = 5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	84
6.17. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 U (0, 20) dağılımı için tahmincilerin MSE ve Bias değerleri . . . . .	86
6.18. Pareto ( $\sigma = 1, \alpha = 5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri (%10 100 değeri aykırı değer olarak atılmış) . . . . .	88
7.1. Radyasyona maruz kalan farelerin yaşam süresi verisi . . . . .	91
7.2. Yaşam süresi verisi için tahminler . . . . .	93
7.3. Sulama verisi . . . . .	95
7.4. Sulama verisi için tahminler . . . . .	96

## ŞEKİLLER DİZİNİ

2.1. Üstel dağılımdan türetilen veri için tahminler ve bu tahminlere ilişkin üstel olasılık yoğunluk fonksiyonları . . . . .	8
3.1. M-tahmincilerin $\psi$ fonksiyonları grafiği . . . . .	18
5.1. Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği . . . .	34
5.2. Pareto dağılımının olasılık yoğunluk fonksiyonları grafiği . . . .	41
6.1. Weib ( $\gamma = 1, \alpha = 0.5$ ) dağılımı için tahmincilerin MSE grafiği	50
6.2. Weib ( $\gamma = 1, \alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği .	52
6.3. Weib ( $\gamma = 1, \alpha = 2$ ) dağılımı için tahmincilerin MSE grafiği .	54
6.4. Weib ( $\gamma = 1, \alpha = 4$ ) dağılımı için tahmincilerin MSE grafiği .	56
6.5. Weib ( $\gamma = 1, \alpha = 8$ ) dağılımı için tahmincilerin MSE grafiği .	58
6.6. Pareto ( $\sigma = 1, \alpha = 0.5$ ) dağılımı için tahmincilerin MSE grafiği	61
6.7. Pareto ( $\sigma = 1, \alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği	63
6.8. Pareto ( $\sigma = 1, \alpha = 2$ ) dağılımı için tahmincilerin MSE grafiği	65
6.9. Pareto ( $\sigma = 1, \alpha = 4$ ) dağılımı için tahmincilerin MSE grafiği	67
6.10. Pareto ( $\sigma = 1, \alpha = 8$ ) dağılımı için tahmincilerin MSE grafiği	69
6.11. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 0.2, \alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği . . . . .	73
6.12. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE grafiği . . . . .	75
6.13. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE grafiği . . . . .	77
6.14. Weib ( $\gamma = 1, \alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği (%10 100 değeri aykırı değer olarak atılmış) . . . . .	79
6.15. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE grafiği . . . . .	83
6.16. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 0.2, \alpha = 5$ ) dağılımı için tahmincilerin MSE grafiği . . . . .	85
6.17. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE grafiği . . . . .	87

6.18. Pareto ( $\sigma = 1, \alpha = 5$ ) dağılımı için tahmincilerin MSE grafiği (%10 100 değeri aykırı değer olarak atılmış) . . . . .	89
7.1. Yaşam süresi verisi için Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği . . . . .	92
7.2. Aykırı değer içeren yaşam süresi verisi için Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği . . . . .	94
7.3. Sulama verisi için Pareto dağılımının olasılık yoğunluk fonksi- yonları grafiği . . . . .	96
7.4. Aykırı değer içeren sulama verisi için Pareto dağılımının olasılık yoğunluk fonksiyonları grafiği . . . . .	97

## 1. GİRİŞ

Dağılım fonksiyonunun parametrelerinin doğru bir şekilde tahmin edilmesi büyük öneme sahiptir. En çok olabilirlik, en küçük kareler ve momentler yöntemi gibi teknikler parametre tahmininde kullanılabilen en yaygın metotlardır. En çok olabilirlik tahmincisi sahip olduğu istatistiksel özelliklerin sonucu olarak en popüler metottur. Ancak verinin doğası gereği ya da ölçüm hatası sebebiyle veri setinde bulunabilecek aykırı değerlerden, yukarıda ifade edilen klasik yöntemler olumsuz yönde etkilenmekte ve bu yöntemlerle elde edilen tahminlerin etkinliği düşmektedir (Boudt ve ark., 2011; He ve Fung, 1999; Shier ve Lawrence, 1984).

Varsayımlarda sapmaların olduğu durumlarda etkin tahmin edicilerin elde edilebilmesi için, aykırı değerlerden ve modeldeki sapmalardan etkilenmeyen "robust (sağlam)" tahmin ediciler tercih edilir. Bu tahmincilerde bulunması beklenen belli özellikler; küçük sapmalarda aşırı tepki vermeyerek sapmanın büyük olduğu durumda da güvenilir olması, veri setindeki aykırı değerleri teşhis edebilmesi, verinin asıl kısmını temsil edebilecek tahminler üretebilmesi ve etkinliğinin olabildikçe yüksek olması şeklinde sıralanabilir (Karagöz, 2012). Son yıllarda büyük gelişmeler gösteren ve sık kullanılan robust yöntemlere, genellikle konum modeli ve regresyon modeli için önerilen, en çok olabilirlik tipi tahminci olarak bilinen M-tahminciler (Huber, 1964; Huber, 1973), Rousseeuw (1984) tarafından önerilen en küçük medyan kareler (LMS) ve en küçük kırılmış kareler (LTS) tahmincileri, sıralı istatistiklerin doğrusal kombinasyonundan hesaplanan L-tahminciler ve dağılım parametrelerinin tahmini için önerilen, klasik moment tahmincisinden yaralanarak elde edilen kırılmış moment tahmincisi (MTM) (Bajorunaite ve Brazauskas, 2008) ve kitlenin ve örneklemin skor fonksiyonlarının birbirine eşitlenmesi ile tahmincilerin elde edildiği medyanlar metodu (He ve Fung, 1999) örnek olarak verilebilir.

Dağılımsal parametrelerin robust tahmine ilişkin literatür aşağıdaki gibi sıralanabilir:

Shier ve Lawrence (1984) Weibull dağılımının parametrelerini robust lineer regresyon yöntemleri ile tahmin etti.

Gather and Schultze (1999) üstel dağılımın ölçek parametresinin robust tahmini olarak medyanı kullandı.

He and Fung (1999), Weibull dağılımının parametreleri için M-tahminciler sınıfında robust medyanlar metodunu önerdi.

Bajorunaite ve Brazauskas (2008) dağılım aileleri için kırılmış momentler metodunu önerdi.

Brazauskas ve Kleefeld (2009), genelleşmiş Pareto dağılımı için pek çok robust tahminciyi inceledi.

Wang ve Cheng (2010) Burr XII dağılımının parametrelerinin tahmini için M-tahmincileri kullandı.

Boudt ve ark. (2011) Weibull dağılımı için kantillere dayalı tahminciler, medyan ve robust ölçek tahmincisine dayalı pek çok tahminciyi inceledi.

Wang ve Lee (2010) Extended Burr Type III dağılımının parametrelerinin tahmini için M-tahmincileri kullandı.

Zyl (2012) üç parametrelili pareto dağılımı için medyan regresyon modelini kullandı.

Yavuz (2012) Weibull dağılımının parametrelerini robust lineer regresyon yöntemleri ile karşılaştırdı.

Zyl ve Schall (2012) Weibull and Gumbel dağılımı için ağırlıklandırılmış en küçük kareler rank regresyon uyguladı.

Kantar ve Arik (2014) Log-logistic dağılımı için M-tahmincilerin performansını inceledi.

Kantar ve Yildirim (2014) Extended Burr III dağılımının robust tahmincilerini inceledi, şekil parametresi için robust kantillere dayalı tahminciyi önerdi.

Çalışmanın ikinci bölümünde dağılım parametrelerinin tahmini için kullanılan temel yöntemlerden en çok olabilirlik tahmincisi, momentler tahmincisi ve en küçük kareler tahmincisi kısaca ele alınacaktır.

Üçüncü bölümde robust tahminleme konusundan bahsedilip robust tahmincilerin değerlendirme kriterlerinden olan etki fonksiyonu ve kırılma noktası konuları ele alınacak, istatistikte dağılımın önemli iki parametresi olan konum ve ölçek parametrelerinin literatürde yaygın kullanılan robust tahmincileri anlatılacaktır.

Dördüncü bölümde ilk olarak doğrusal regresyon yardımıyla dağılım parametrelerinin elde edildiği robust yöntemler tanıtılacak olup, sırasıyla kantillere dayalı tahminleme, konum-ölçek ailesine dönüştürülebilen dağılımlar için robust tahminleme, medyanlar metodu ve kırılmış momentler metodu anlatılacaktır.

Beşinci bölümde bir önceki bölümde ele alınan robust tahminciler Weibull ve Pareto dağılımları için hesaplanacaktır.

Altıncı bölümde MATLAB programında yazılan özel kodlar yardımıyla simulasyon çalışması yapılacak ve tahmincilerin performansları incelenecektir.

Yedinci bölümde gerçek veri seti üzerinde robust tahminciler uygulanacak ve tahmincilerin performansları karşılaştırılacaktır.

Sonuç bölümünde ise simulasyon ve uygulama kısmında elde edilen bulgular özetlenecektir.

## 2. DAĞILIM PARAMETRELERİNİN TAHMİNİ

Parametre tahmini istatistiğin temel konularından biridir. Tez çalışmasının bu bölümünde de dağılım parametreleri için kullanılan bazı temel yöntemler ele alınacaktır. Bu yöntemler en küçük kareler tahmincisi, momentler tahmincisi ve en çok olabilirlik tahmincisidir. Ayrıca şu da vurgulanmalıdır ki bu tahmincilere ek olarak pek çok metod literatürde bulunmaktadır (Kantar ve Şenoğlu, 2008).

### 2.1. En Küçük Kareler (Least Squares-LS) Tahmincisi

1805 yılında Legendre tarafından önerilen en küçük kareler metodu gözlem noktalarına eğri uydurmada kullanılan bir yaklaşırma yöntemi olarak bilinmektedir. Genellikle, doğrusal ve doğrusal olmayan regresyon modellerinde parametre tahmininde sık kullanılan bir yöntemdir.  $X$  ile  $Y$  gibi iki değişken için  $X = x$  olarak verildiğinde  $Y$ 'nin koşullu dağılımının beklenen değeri,

$$E(Y|X = x) = \beta_0 + \beta_1 x \quad (2.1)$$

şeklinde olsun ve  $\beta_0, \beta_1$  katsayıları tahmin edilmek istensin.  $X_i = x_i$  değeri sabit olmak üzere bu  $x_i$ 'lere karşılık gözlenen  $Y_i$  değerleri için

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.2)$$

doğrusal modeli yazılabilir. (2.2) eşitliğindeki  $\varepsilon_i$ 'ler sıfır ortalamalı ve aynı varyanslı bağımsız gözlenemeyen rassal değişkenlerdir.

$$\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.3)$$

ifadesinde parametrelere göre kısmi türevler alınıp sıfıra eşitlenirse,

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n Y_i \quad (2.4)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n Y_i x_i \quad (2.5)$$

denklemleri elde edilir. (2.4) ve (2.5) denklemlerine normal denklemler denir. Normal denklemlerin çözümleri olan,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{n \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.6)$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (2.7)$$



değerleri sırasıyla  $\beta_1$  ve  $\beta_0$ ' in en küçük kareler tahmincileridir (Montgomery ve ark., 2001).

(2.2) eşitliğindeki model matris formunda yazılırsa,

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

olmak üzere

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \quad (2.8)$$

şeklinde ifade edilebilir. Bu ifade yardımıyla normal denklemler  $X'X\hat{\beta} = X'Y$  olarak elde edilir ve bu denklemlerin çözümü sonucu katsayıların en küçük kareler tahmincisi,

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.9)$$

şeklinde (Akdi, 2010).

## 2.2. Momentler Tahmincisi

Parametre tahmininde kullanılan yöntemlerden biri olan momentler yöntemi, 1800'lü yılların sonlarında Karl Pearson'a dayanan en eski yöntemlerden birisidir. Parametreleri tahmin edilecek dağılımın kitle momentleri var ise momentler tahmincisi elde edilebilir. Momentler tahmincisi, kitle momentleri ile örneklem momentlerinin birbirine eşitlenmesi ile elde edilir.

$\theta = (\theta_1, \theta_2, \dots, \theta_r)' \in \Theta \subset \Re^r$  olan  $\theta$  parametre vektörü tahmin edilmek istensin.  $X_1, X_2, \dots, X_n$  olasılık yoğunluk fonksiyonu  $f(x, \theta), \theta \in \Theta$  olan dağılımdan çekilen bir örneklem için, dağılımın kitle momentleri,

$$\gamma_k = E_\theta(X^k), \quad k = 1, 2, 3, \dots, r \quad (2.10)$$

ve örneklem momentleri,

$$\hat{\gamma}_k = \frac{1}{n} \sum_{i=1}^n X_i^k, \quad k = 1, 2, 3, \dots, r \quad (2.11)$$

olmak üzere kitle ve örneklem momentlerinin ilk r tanesinin eşitlenmesi ile oluşan denklem sisteminin, bilinmeyen parametrelere göre çözümü sonucu elde

edilen

$$\hat{\theta} = (\hat{\theta}_1(X_1, X_2, \dots, X_n), \hat{\theta}_2(X_1, X_2, \dots, X_n), \dots, \hat{\theta}_r(X_1, X_2, \dots, X_n))'$$

ifadesine  $\theta$ 'nın momentler tahmincisi denir (Öztürk ve ark., 2006).

### 2.3. En Çok Olabilirlik (Maximum Likelihood-ML) Tahmincisi

En çok olabilirlik tahmincisi asimptotik özellikleri gereği parametre tahmininde kullanılan en popüler yöntemdir. En çok olabilirlik tahmincileri olabilirlik fonksiyonunu maksimum yapan değerlerdir.

$X = (X_1, X_2, \dots, X_n)$  olasılık veya olasılık yoğunluk fonksiyonu  $f(x; \theta)$  olan bir dağılımdan çekilen örneklem olsun.  $\theta$ 'nın olabilirlik fonksiyonu,

$$L(\theta; X = x) = f(x; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (2.12)$$

şeklindedir. (2.12) fonksiyonunu maksimum yapan değer  $\theta$ 'nın en çok olabilirlik tahminidir. Bir başka deyişle,

$$\hat{\theta}_n(X) = \arg \max_{\theta} L(\theta|X = x) \quad (2.13)$$

dir. Logaritma fonksiyonunun monoton olması sebebiyle genellikle olabilirlik fonksiyonunun logaritması maksimize edilir. Yani bir dağılımın en çok olabilirlik tahmincisi,

$$\hat{\theta}_n(X) = \arg \max_{\theta} (\log(L(\theta|X = x))) \quad (2.14)$$

şeklindedir. En çok olabilirlik tahmincilerini elde etmek için (2.14) denkleminin parametrelere göre kısmi türevleri alınır ve sıfıra eşitlenir. Bilinmektedir ki en çok olabilirlik tahmincileri, regülerlik koşulları altında, asimptotik yansız ve tutarlı tahmincilerdir (Kantar ve Şenoğlu, 2008; Öztürk ve ark., 2006).

Ancak burada bahsedilen tahminciler veri setinde aykırı değer bulunduğu zaman tutarsız sonuçlar verebilmektedir (Boudt ve ark., 2011; He ve Fung, 1999; Zhang ve ark., 2006; Zyl, 2012; Yavuz, 2012; Mutan ve Şenoğlu, 2009; Ergül, 2006). Bu durum, Çizelge 2.1'de verilen, örneklem boyutu  $n = 20$  olan ve parametresi  $\mu = 1$  olan üstel dağılımdan üretilen veri seti üzerinde incelenmiştir.

**Çizelge 2.1.** Üstel (1)'den üretilmiş veri

3.3323	0.1635	0.0683	0.3875	0.2774	0.2969	0.9359	0.4224	1.7650	0.3481
3.4473	1.2840	3.0754	2.3317	0.1942	0.3641	1.1485	0.0511	3.3684	0.8238

Üstel dağılımın olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right), \quad x > 0, \quad \lambda > 0 \quad (2.15)$$

şeklinde. Dağılımın parametresi olan  $\lambda$ 'nın en çok olabilirlik ve moment tahmincileri bilindiği üzere örneklemin ortalamasına eşittir. Çizelge 2.1'de verilen örneklem yardımıyla dağılımın parametresi, aşağıdaki gibi tahmin edilmiştir.

$$\hat{\lambda} = \frac{\sum x_i}{n} = 1.2043 \quad (2.16)$$

Daha sonra örneklemden 1.1485 değerinin bir hata sonucu 114.85 olarak kaydedildiği düşünülmüştür. Bu durumda tahminci,

$$\hat{\lambda} = \frac{\sum x_i}{n} = 6.8894 \quad (2.17)$$

olarak elde edilmiştir. Görüldüğü üzere dağılım parametresinin en çok olabilirlik tahmincisi, bir hata sonucu ortaya çıkan tek bir aykırı değerden oldukça etkilenmiştir.

Üstel dağılımın teorik medyanının  $\lambda \log 2$  olduğu bilinmektedir. Ayrıca medyan robust tahmincidir. Momentler yöntemine paralel olarak Üstel dağılımın teorik medyanını, verinin medyanına eşitleyerek  $\lambda$  parametresi aşağıdaki gibi tahmin edilebilir.

$$Med(X) = Med(X_i) \quad (2.18)$$

$$\lambda \log 2 = 0.6231 \quad (2.19)$$

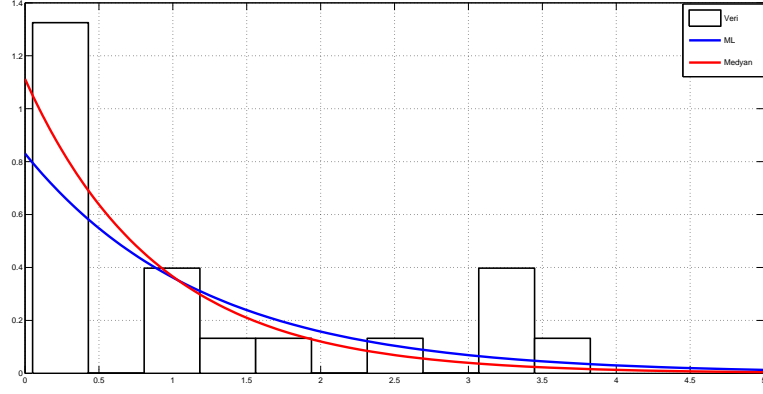
$$\hat{\lambda} = 0.6231 / \log 2 \quad (2.20)$$

$$\hat{\lambda} = 0.8990 \quad (2.21)$$

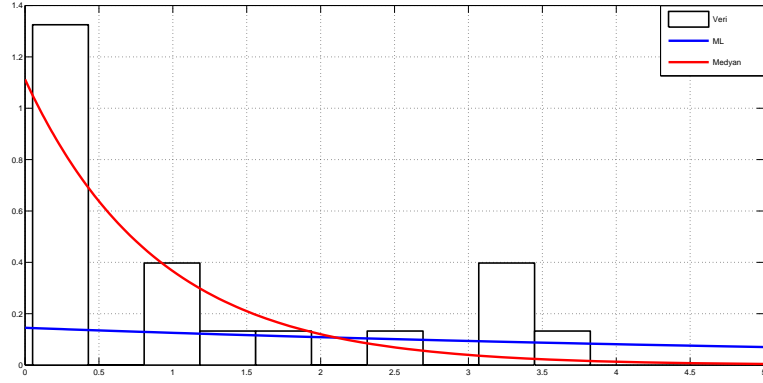
Hatalı kayıt olması durumunda tahmin,

$$\hat{\lambda} = 0.8990 \quad (2.22)$$

şeklinde bulunmuş olup görüldüğü üzere medyanın robust tahminci olması sebebiyle, ilk ve ikinci durum için aynı sonuca ulaşılmış, yani tahmin üzerinde aykırı değer etkisi olmamıştır. Bu durum Şekil 2.1’de verilen grafiklerle de net bir şekilde görülebilir.



(a) Orjinal veri



(b) Aykırı değer içeren veri

Şekil 2.1. Üstel dağılımdan türetilen veri için tahminler ve bu tahminlere ilişkin üstel olasılık yoğunluk fonksiyonları

### 3. ROBUST TAHMİNLEME

Veri setinin geneline nazaran farklılık gösteren gözleme aykırı değer denir ve klasik istatistiksel metotların çoğunluğu tek bir aykırı değer noktasından bile oldukça olumsuz etkilenir. Bu metotlar arasından en çok olabilirlik, momentler ve en küçük kareler tahmincilerinin veride aykırı değer olduğunda tutarsız sonuçlar verdiği bilinmektedir (Maronna ve ark., 2006; Montgomery ve ark., 2001; Hampel ve ark., 1986).

Sağlıklı, dayanıklı, gürbüz gibi anlamlar taşıyan robust sözcüğü ilk kez 1953 yılında Box tarafından istatistik literatüründe kullanılmış olup varsayımların sağlanmadığı durumda sağlıklı sonuçlar veren istatistiksel yöntemler robust kelimesi ile adlandırmıştır (Kayhan, 2006). İstatistiksel modellemede ve veri analizinde kullanılan robust yöntemlerle varsayımların sağlandığı ve veri setinde aykırı değerlerin bulunmadığı durumların yanı sıra varsayımların sağlanmadığı ve ölçüm hatasından ya da verinin doğasından kaynaklanan aykırı değerlerin bulunduğu durumlarda da güvenilebilir tahminler üretilebilir ve bu tahminlere bağlı olarak sağlıklı güven aralıkları elde edilebilir (Maronna ve ark., 2006). İyi bir robust tahmincisi,

- (i) Varsayımlar sağlandığında ve veri setinde aykırı değerler bulunmadığı zaman klasik tahmincilere yakın sonuç vermesi,
- (ii) Aykırı değerlerin bulunduğu durumlarda klasik yöntemlerden daha iyi performans sergilemesi,
- (iii) Anlaşılabilir ve kolay hesap edilebilir olması,

gibi özellikleri sağlamalıdır (Zhang ve ark., 2006; Wu, 1985).

Bu bölümde robust yöntemlerin temel özellikleri üzerinde durulacak, konum ve ölçek parametrelerinin robust tahmincileri tanıtılacaktır.

#### 3.1. Robust Tahminciler için Değerlendirme Kriterleri

Robust yöntemler sonucu elde edilen tahminin genel olarak değerlendirilmesinde kullanılan hata kareler ortalamasının (Mean Squared Error-MSE) yanında

bu tahmincilerin dayanıklılık açısından değerlendirilmesine olanak sağlayan, tek bir noktanın tahmin ediciye etkisini ölçen "etki fonksiyonu" ve tahmin edicinin maksimum ne kadar kirlilik oranına dayanabileceğinin ölçüsü olan "kırılma noktası" robust tahmincilerin değerlendirilmesi için kullanılan önemli ölçütlerdir.

### 3.1.1. Etki Fonksiyonu (Influence Function-IF)

IF, büyük örneklem durumunda, küçük bir orandaki kirliliğin tahmin ediciye olan etkisinin ölçüsüdür.  $F$  varsayılan dağılım ve  $\varepsilon$  küçük bir sayı olmak üzere,

$$F_\varepsilon = (1 - \varepsilon)F + \varepsilon\delta_x \quad (3.1)$$

eşitliği  $\varepsilon$  oranında kirlenmiş dağılımı ifade eder. Burada  $\delta_x$ ,  $x$  noktasında bir olasılık kütlelerini ifade eden kümülatif dağılım fonksiyonudur.  $T$  tahminci olmak üzere  $T$  için etki fonksiyonu aşağıdaki gibi tanımlanır (Hampel ve ark., 1986).

$$IF(x; T, F) = \lim_{\varepsilon \rightarrow 0} \frac{T(F_\varepsilon) - T(F)}{\varepsilon} \quad (3.2)$$

(3.2) eşitliğinden anlaşıldığı üzere varsayılan  $F$  dağılımının tahmincisi ile bu dağılımın çok küçük bir oranda kirlenmesi sonucu elde edilen  $F_\varepsilon$  dağılımının tahmin edicisi arasındaki fark etki fonksiyonu yardımıyla elde edilir. Etki fonksiyonu, yapılan çok küçük bir kirliliğin tahmin edici üzerindeki etkisini ölçer. Ayrıca etki fonksiyonu yardımıyla büyük örneklem için bir tahmincinin değerlendirilmesinde önemli bir kriter olan asimptotik varyans da hesaplanabilir (Kayhan, 2006).

Sonlu örneklem durumunda duyarlılık eğrisi etki fonksiyonu yardımıyla (sensitivity curve- $SC$ ) elde edilir. Etki fonksiyonu duyarlılık eğrisinin limit durumu olarak düşünülebilir.  $(x_1, x_2, \dots, x_n)$  örnekleme  $x_0$  gözlemi dahil edildiğinde kirlilik oranı  $\frac{1}{(n+1)}$  ve duyarlılık eğrisi

$$SC_n(x_0) = \frac{T_{n+1}(x_1, x_2, \dots, x_n, x_0) - T_n(x_1, x_2, \dots, x_n)}{\frac{1}{(n+1)}} \quad (3.3)$$

$$= (n+1)[T_{n+1}(x_1, x_2, \dots, x_n, x_0) - T_n(x_1, x_2, \dots, x_n)] \quad (3.4)$$

şeklinde (Maronna ve ark., 2006).

Literatürde sık kullanılan ortalama ve medyan tahmincilerinin duyarlılıkları kolayca hesaplanabilir. Belirli bir dağılımdan çekilen  $(X_1, X_2, \dots, X_n)$  örneğine farklı bir  $Y$  gözlemi eklensin ve ortalamanın duyarlılık eğrisine olan etkisi incelensin:

$$T_n = \bar{X}_n, T_{n+1} = \frac{1}{(n+1)}(n\bar{X}_n + Y)$$

$$SC_n(Y) = (n+1)\left[\left(\frac{n}{(n+1)} - 1\right)\bar{X}_n + \frac{1}{(n+1)}Y\right] = (n+1)\left[\frac{1}{(n+1)}(Y - \bar{X}_n)\right]$$

$$SC_n(Y) = Y - \bar{X}_n, \quad SC_n(Y) = \sup |Y - \bar{X}_n| = \infty$$

Dolayısıyla örnekleme farklı bir değer eklenmesi sonucunda ortalamanın sınırsız duyarlılığa sahip olabileceği görülmektedir.

Şimdi  $n = 2m + 1$  ve  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  sıralı gözlem değerleri için medyan tahmincisi ele alınsın.  $T_n = T_n(X_1, X_2, \dots, X_n) = X_{(m+1)}$  olur. Örnekleme  $Y$  değeri eklendiğinde  $T_{n+1} = T_{n+1}(X_1, X_2, \dots, X_n, Y)$  tahmincisinin değeri  $Y$ 'nin örneklem içindeki konumuna göre aşağıdaki değerleri alabilir:

$$T_{n+1} = \begin{cases} \frac{X_{(m)} + X_{(m+1)}}{2} & Y \leq X_{(m)} \\ \frac{X_{(m+1)} + X_{(m+2)}}{2} & Y \geq X_{(m+2)} \\ \frac{Y + X_{(m+1)}}{2} & X_{(m)} \leq Y \leq X_{(m+2)} \end{cases} \quad (3.5)$$

Bu değerlere karşılık duyarlılık eğrisi de aşağıdaki değerleri alabilir:

$$SC_n(Y) = \begin{cases} \frac{X_{(m)} - X_{(m+1)}}{2}, & Y \leq X_{(m)} \\ \frac{X_{(m+2)} - X_{(m+1)}}{2}, & Y \geq X_{(m+2)} \\ \frac{Y - X_{(m+1)}}{2}, & X_{(m)} \leq Y \leq X_{(m+2)} \end{cases} \quad (3.6)$$

Görüldüğü üzere duyarlılık eğrisinin değerlerinin hepsi sonlu değerlerdir. Medyan tahmincisi sınırlı duyarlılık eğrisine sahiptir (Jureckova ve Picek, 2006).

### 3.1.2. Kırılma Noktası (Breakdown Point-BP)

Sınırlı örneklem için kırılma noktası, bir tahmin ediciyi bozan minimum kirlilik oranıdır (Montgomery ve ark., 2001). Basit bir ifade ile bir tahmincinin kırılma noktası, tahmincinin gerçek parametre hakkında bilgi vermesini engellemeyecek maksimum kirlilik oranıdır. Yani tahminci veride aykırı değerlerin bulunmasına karşın normal veriler hakkında doğru bilgi verebilmektedir (Maronna ve ark., 2006).

$n$  tane veriden oluşan  $x^0 = (x_1, x_2, \dots, x_n)$  örnekleminin tahmincisi  $T_n(x^0)$  olsun.  $x^0$  örneklemindeki  $m(m < n)$  tane değer, aykırı değer noktaları ile değiştirilsin. Yeni oluşan örneklem  $x^m$  olmak üzere tahmincisi  $T_n(x^m)$  olsun.  $T_n$  tahmincisi için kırılma noktası,

$$\varepsilon_n^*(T_n, x^0) = \frac{m^*(x^0)}{n} \quad (3.7)$$

değeridir. (3.7) denklemindeki  $m^*(x^0)$  değeri,

$$\sup_{x^m} \|T_n(x^m) - T_n(x^0)\| = \infty \quad (3.8)$$

ifadesini sağlayan en küçük aykırı değer sayısıdır (Jureckova ve Picek, 2006).

Tahmincinin parametre hakkında bilgi verebilmesi için örneklemden oluşan gözlemlerin sayısının aykırı değer sayısından daha fazla olması gerekir. Bu sebeple kırılma noktası  $\varepsilon^* \leq \frac{1}{2}$  olmalıdır.

### 3.2. Konumun Robust Tahmini

Öncelikle veri setinin merkezini veya konumunu saptamakta kullanılan klasik yöntemler incelenmelidir. Bu amaçla, ortalama hesabında sık kullanılan iki tahminci ele alınacaktır.

$x = (x_1, x_2, \dots, x_n)$  gözlem değerleri olmak üzere bu gözlem değerleri için örneklem ortalaması,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.9)$$

şeklinde hesaplanır ve görüldüğü üzere, hesaplanan değer verinin aritmetik ortalamasıdır. Örneklem ortalaması verinin merkezini veya konumunu hesaplamak için kullanılır. Bu amaçla sıklıkla kullanılan diğer klasik tahminci ise örneklem medyanıdır. Bir örneklem için medyan değeri veri sıralandığında tam ortada kalan değerdir. Yani  $x = (x_1, x_2, \dots, x_n)$  örneklemini için  $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$  sıralı istatistikleri olmak üzere, eğer  $n$  örneklem sayısı  $n = 2m - 1$  şeklinde tek sayı ise  $x$  örnekleminin medyanı,

$$Med(x) = x_{(m)} \quad (3.10)$$



değerine karşılık gelir. Şayet  $n$  örneklem sayısı  $n = 2m$  gibi çift sayı ise  $x$  örnekleminin medyanı,

$$Med(x) = \frac{x_{(m)} + x_{(m+1)}}{2} \quad (3.11)$$

şeklinde hesap edilir.

Örneklem ortalamasının ve örneklem medyanının veri setinde bulunan tek bir aykırı değere karşılık nasıl değişim gösterdikleri bir örnek üzerinde incelenir. Bu amaçla kepekli unun içerdiği bakır oranları değerlerinin bulunduğu Çizelge 3.1'deki küçükten büyüğe sıralanmış veriler için ortalama ve medyan değerleri hesaplanmıştır (Maronna ve ark., 2006).

**Çizelge 3.1.** Kepekli un verisi

2.20	2.20	2.40	2.40	2.50	2.70	2.80	2.90
3.03	3.03	3.10	3.37	3.40	3.40	3.40	3.50
3.60	3.70	3.70	3.70	3.70	3.77	5.28	28.95

Çizelge 3.1'deki 24 tane gözlem için örneklem ortalamasının değeri  $\bar{x} = 4.28$ 'dir. Verinin merkezini bulmak amacıyla hesaplanan örneklem ortalaması, veride bulunan yalnızca iki değerden daha küçüktür yani verinin merkezi hakkında sağlıklı bilgi verememektedir. Veri seti incelendiğinde diğer gözlem değerlerinden oldukça farklı olan 28.95 gözlemi aykırı değer olarak göze çarpmaktadır. Bu değer göz ardı edilerek örneklem ortalaması tekrar hesaplandığında sonuç  $\bar{x} = 3.21$  olmaktadır ve bu değer örneklemin merkezinde yer alır. Dolayısıyla örneklem ortalaması tek bir aykırı değerden bile oldukça etkilenmekte ve sağlıklı sonuç üretememektedir.

Aynı gözlemler için örneklem medyanı hesaplandığında  $n = 24$  olduğundan  $Med(x) = \frac{x_{(12)} + x_{(13)}}{2} = 3.38$  değerini alır. Aykırı değer olan 28.95 gözlemi göz ardı edilerek örneklem medyanı tekrar hesaplandığında  $Med(x) = 3.37$  değerini alır. Görüldüğü üzere medyan aykırı değer varlığından etkilenmemiş ve her iki durumda da medyan ile elde edilen sonuçlar verinin merkezinde yer almaktadır (Maronna ve ark., 2006).

Yukarıdaki örnekte aykırı değer olarak değerlendirilen gözlemin yerine  $-\infty < x < \infty$  aralığında değişen, keyfi bir  $x$  değeri koyulduğunda örneklem ortalaması da  $(-\infty, \infty)$  aralığında değişir. Yani bölüm 3.1’de de belirtildiği gibi örneklem ortalaması sınırsız duyarlılığa sahiptir ve kırılma noktası bir tahmin edici için olabilecek en küçük değer olan  $\frac{1}{n}$ ’dir. Başka bir ifade ile tek bir aykırı değer bile tahmin ediciyi bozmaktadır (Montgomery ve ark., 2001). Aynı durum örneklem medyanı için düşünüldüğünde örneklem ortalamasının aksine tahmin edicinin değerinde çok küçük değişiklik olmaktadır. Dolayısıyla örneklem medyanı örneklem ortalaması yerine kullanılacak çok iyi bir robust alternatiftir. Bu çalışmada kullanılacak diğer robust konum tahmincileri olan M-tahminciler ve kırılmış ortalamalar izleyen bölümde ele alınacaktır.

### 3.2.1. M-Tahminciler

En çok olabilirlik tipi tahminciler olarakta adlandırılan M-tahminciler, 1964 yılında Huber tarafından önerilmiştir.  $T_n, \theta'$  nın M-tahmincisi olmak üzere,

$$\min_{\theta} \sum_{i=1}^n \rho(X_i, \theta) \quad (3.12)$$

şeklinde tanımlanır (Jureckova ve Picek, 2006). (3.12) ifadesindeki  $\rho$  örneklemin yapısına uygun olarak seçilen amaç fonksiyonudur ve aşağıdaki özelliklere sahiptir:

- (i)  $|x|$ 'in azalmayan bir fonksiyonudur
- (ii)  $\rho(0) = 0$
- (iii)  $0 < x_1 < x_2$  için  $\rho(x_1) < \rho(x_2)$
- (iv) Eğer  $\rho$  sınırlı ise  $\rho(\infty) = 1$  kabul edilir (Maronna ve ark., 2006; Rousseeuw ve Leroy, 1987).

$\rho$  fonksiyonu,  $f(x, \theta)$  yoğunluk fonksiyonu olmak üzere,  $\rho = -\log(f)$  olarak seçilirse M-tahmincisi,

$$\min_{\theta} \sum_{i=1}^n (-\log f(x_i, \theta)) \quad (3.13)$$

şeklinde elde edilir ve bu denklemin sonucu en çok olabilirlik tahmincisini verir. Dolayısıyla M-tahminciler sınıfı en çok olabilirlik tahmincisini de kapsar.

Eğer  $\rho$  amaç fonksiyonu türevlenebilir ise, (3.12) denkleminin  $\theta$ 'ya göre türevi alındığında,

$$\sum_{i=1}^n \psi(x_i, \theta) = 0 \quad (3.14)$$

eşitliği elde edilir.  $\rho$  fonksiyonunun türevi olan  $\psi$ ,

(i)  $\psi(-x) = -\psi(x)$  (Tek fonksiyondur.)

(ii)  $x \geq 0$  için  $\psi(x) \geq 0$

özelliklerine sahiptir (Maronna ve ark., 2006; Rousseeuw ve Leroy, 1987).

**Konumun M-Tahmincisi:**  $x_1, x_2, \dots, x_n$  değerleri  $\mu$  ortalamalı bir kitleden gözlemler olmak üzere  $x_i$  gözlemi,

$$x_i = \mu + u_i \quad (3.15)$$

eşitliği yardımıyla ifade edilebilir. Bu eşitlikteki  $u_1, u_2, \dots, u_n$  hata terimleri rassal olup, bu model **konum modeli** olarak adlandırılır. Eğer aynı kitleden çekilen  $x_1, x_2, \dots, x_n$  gözlem değerleri bağımsız ise, hatalar da bağımsızdır ve aynı dağılım fonksiyonuna sahiptir.

$F_0$  hata terimlerinin dağılım fonksiyonu olmak üzere, yoğunluk fonksiyonu  $f_0 = F_0'$  şeklindedir. Hataların en çok olabilirlik fonksiyonu,

$$L(\mu; x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_0(x_i - \mu) \quad (3.16)$$

olmak üzere en çok olabilirlik tahmincisi,

$$\hat{\mu} = \arg \max_{\mu} L(\mu; x_1, x_2, \dots, x_n) \quad (3.17)$$

şeklinde elde edilir. Hataların dağılımı olan  $F_0'$  in tam olarak bilindiği durumlarda en çok olabilirlik tahmincisi en iyi konum tahmincisidir. Hataların dağılımı net olarak bilinmediği için amaç en uygun konum tahmincisini M-tahminciler yardımıyla elde etmektir. (3.12) denkleminde yararlanarak konumun M-tahmincisi,

$$\hat{\mu} = \arg \min_{\mu} \sum_{i=1}^n \rho(x_i - \mu) \quad (3.18)$$



eşitliği yardımıyla elde edilir. Bu eşitlikte  $\rho$  nun yerine türevi olan  $\psi$  fonksiyonu yazılırsa,

$$\sum_{i=1}^n \psi(x_i - \hat{\mu}) = 0 \quad (3.19)$$

elde edilir. Bu eşitliğin çözümü konum parametresinin M-tahminini verir (Maronna ve ark., 2006).

İteratif yollarla çözülebilen bu eşitliğin kökü (ya da köklerinden biri) konumun M-tahminidir.  $\psi$  fonksiyonu sürekli ve artan ise (3.14) eşitliği tek bir çözüme sahiptir aksi taktirde birden fazla çözüm olabilir. Şayet fonksiyon sürekli değil ise (3.14) eşitliğinin çözümü olmayabilir (Maronna ve ark., 2006).

**M-Tahmincilerin Sınıflandırılması:** M-tahminciler  $\psi$  fonksiyonunun yapısına göre aşağıdaki gibi sınıflandırılabilir (Montgomery ve ark., 2001).

(i) **Monoton M-Tahminciler:** Orijinden itibaren belirlenen bir noktaya kadar doğrusal ya da azalarak artan, bu noktadan sonra ise sabit ya da çok az artışlarla ağırlık veren tahmincilerdir. Monoton M-tahmincileri veri setindeki hiçbir değere sıfır ağırlık vermez.

(ii) **Yeniden Azalan (Redescending) M-Tahminciler:** Büyük aykırı değerlere karşı daha dayanıklıdırlar. Yeniden azalan M-tahmincileri  $\psi$  fonksiyonuna göre ikiye ayrılırlar:

$\psi$  fonksiyonu, belirli bir noktadan sonra sıfıra yakın çok küçük ağırlık veriyorsa **yavaş yeniden azalan (soft redescending)**, belirli bir noktadan sonra sıfır ağırlık veriyorsa **hızlı yeniden azalan (hard redescending)** M-tahminciler adını alır. Hızlı yeniden azalan M-tahminciler aykırı değer teşhisinde de kullanılabilirler (Yıldırım, 2012).

Bu çalışmanın dağılım parametrelerinin robust tahminlemesi kısmında kullanılacak M-tahmincileri Çizelge 3.2'de verilmiştir.

Çizelge 3.2'de verilen M-tahmincilerinin  $\psi$  fonksiyonu grafikleri Şekil 3.1'de verilmiştir.

**Çizelge 3.2.** M-Tahminciler

M-Tahmincisi	Türü	Ayarlama Sabiti	$\psi(x)$
Andrews Sinus	Hızlı Yeniden Azalan	$c = 1.339$	$\psi(x) = \begin{cases} \sin(x/c), &  x  \leq c\pi \\ 0, &  x  > c\pi \end{cases}$
Tukey Biweight	Hızlı Yeniden Azalan	$c = 4.685$	$\psi(x) = \begin{cases} x[1 - (\frac{x}{c})^2]^2, &  x  \leq c \\ 0, &  x  > c \end{cases}$
Cauchy	Yavaş Yeniden Azalan	$c = 2.385$	$\psi(x) = \frac{x}{1+(\frac{x}{c})^2}, \quad  x  < \infty$
Welsch	Yavaş Yeniden Azalan	$c = 2.985$	$\psi(x) = x \exp(-(x/c)^2), \quad  x  < \infty$
Huber	Monoton	$c = 1.345$	$\psi(x) = \begin{cases} x, &  x  \leq c \\ c \operatorname{sign}(x), &  x  > c \end{cases}$
Fair	Monoton	$c = 1.400$	$\psi(x) = \frac{x}{1+ x /c}, \quad  x  < \infty$

$\psi$  fonksiyonu gözlemdaki her bir verinin ne kadar ağırlıklandırılacağını ağırlık fonksiyonu aracılığıyla kontrol eder. Ağırlık fonksiyonu,

$$W(x) = \begin{cases} \frac{\psi(x)}{x}, & x \neq 0 \text{ ise,} \\ \psi'(0), & x = 0 \text{ ise.} \end{cases} \quad (3.20)$$

şekindedir. (3.19) eşitliği ağırlık fonksiyonu yardımı ile yazılırsa,

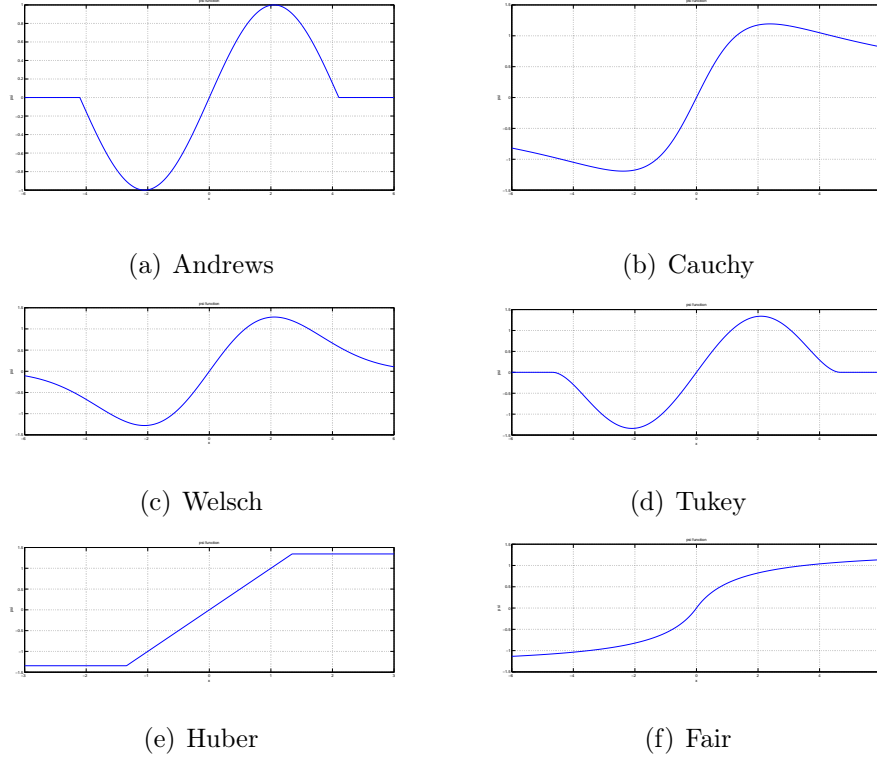
$$\sum_{i=1}^n W(x_i - \hat{\mu})(x_i - \hat{\mu}) = 0 \quad (3.21)$$

eşitliği elde edilir. Bu eşitlik aracılığıyla konumun M-tahmincisi,

$$\hat{\mu} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}, \quad w_i = W(x_i - \hat{\mu}) \quad (3.22)$$

olarak elde edilir. Bu eşitlikten de görüldüğü üzere konumun M-tahmincisi ağırlıklı ortalama şeklinde ifade edilebilir. Eğer  $W(x)$ ,  $|x|$  in artmayan fonksiyonu ise aykırı değerler daha az ağırlık alır.

Ayrıca Çizelge 3.2'nin üçüncü sütununda verilen ayarlama sabiti (tuning constat) tahmin edicinin robustluğunun ve etkinliğinin belirlenmesinde rol oynar. Başka bir ifade ile etkinlik ve sağlamlık arasındaki dengenin ayarlanmasında yardımcıdır. Ayarlama sabiti belirlenirken tahmin edicinin dayanıklılığının yüksek olması sağlanırken, normal dağılımda etkinliğinin yüksek olmasına da dikkat edilmelidir. Çizelge 3.2'de verilen ayarlama sabiti değerleri, standart normal dağılımda tahmincilerin etkinliğinin %95 olmasını sağlayacak şekilde belirlenmiştir (Maronna ve ark., 2006; Yıldırım, 2012).



Şekil 3.1. M-tahmincilerin  $\psi$  fonksiyonları grafiği

### 3.2.2. Kırılmış Ortalama

Konumun robust tahmini için bir diğer yaklaşım da en küçük ve en büyük değerlerin hesaba katılmadan tahmincinin elde edilmesidir.  $x = (x_1, x_2, \dots, x_n)$  örnekleme için  $\alpha \in [0, 1/2)$  ve  $m = [(n - 1)\alpha]$  olmak üzere kırılmış ortalama tahmincisi,

$$\bar{x}_\alpha = \frac{1}{n - 2m} \sum_{i=m+1}^{n-m} x_{(i)} \quad (3.23)$$

şeklinde hesaplanır. Denklemdaki  $x_{(i)}$  örneklemin sıralı istatistiği ve  $\alpha$  da atılan veri oranıdır.  $\alpha = 0$  ve  $\alpha \rightarrow 0.5$  olarak alındığında (3.23) ifadesinin sonucu sırasıyla örneklem ortalamasını ve medyanını verir (Maronna ve ark., 2006). Burada kırma simetrik olarak yapılmıştır. Dağılım simetrik ise bu durum uygundur ancak dağılım sağa ya da sola çarpık ise bu durum uygun değildir.

### 3.3. Ölçeğin Robust Tahmini

Standart sapma (standart deviation(SD)) ve ortalama mutlak sapma (mean absolute deviation (MD)) veri setindeki deęişkenlięi ölçmede sıklıkla kullanılan klasik ölçek tahmincilerindendir.  $x = (x_1, x_2, \dots, x_n)$  örneklemini için örneklem standart sapması,

$$SD(X) = \left[ \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (3.24)$$

formülü yardımıyla hesap edilir. (3.24) yardımıyla Çizelge 3.1'de verilen kepekli un verisinin standart sapması hesaplandığında  $SD(x) = 5.30$  deęerini almaktadır. Aykırı deęer olarak deęerlendirilen 28.95 deęeri çıkarıldıktan sonra geri kalan gözlemler için örneklem standart sapması hesaplandığında  $SD(x) = 0.69$  olmaktadır. Görüldüğü üzere deęişkenliğin ölçümünde kullanılan klasik tahmincilerden biri olan standart sapma yöntemi tek bir aykırı deęerden oldukça olumsuz bir şekilde etkilenmiştir.

$x = (x_1, x_2, \dots, x_n)$  örneklemini için örneklem ortalama mutlak sapması,

$$MD(x) = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| \quad (3.25)$$

şeklinde hesaplanır. Kepekli un örneğindeki gözlemler için (Çizelge 3.1) ortalama mutlak sapma deęeri hesaplandığında  $MD(x) = 2.14$  olmaktadır. Aykırı deęer olan 28.95 gözlemi çıkarılarak tekrar hesaplandığında ortalama mutlak sapma deęeri  $MD(x) = 0.52$  olmaktadır. Bu tahminci de tek bir aykırı deęerden oldukça fazla etkilenmiştir.

İzleyen bölümde,  $SD$  ve  $MD$  tahmincilerine alternatif olarak önerilen robust ölçek tahmincileri incelenecektir.

#### 3.3.1. Medyan Mutlak Sapma (Median Absolute Deviation (MAD) Tahmincisi

Literatürde yaygın olarak kullanılan robust ölçek tahmincilerinden biri olan  $MAD$ ,  $x = (x_1, x_2, \dots, x_n)$  örneklemini için,

$$MAD(x) = \frac{med|x - med(x)|}{0.6745} \quad (3.26)$$

şeklinde hesaplanır. (3.26)'daki 0.6745 sabiti, hataların normal dağıldığı ve  $n$ 'nin büyük olduğu durumda  $MAD$  değerinin,  $\sigma$ 'nın yansız bir tahmincisi olması için kullanılır (Stuart, 2011). Bu tahminci olabilecek en yüksek kırılma noktası değerine (%50) sahip olup, basit yapısı ve kolay hesaplanabilirliğinden dolayı yaygın olarak kullanılır. Belirtilen avantajlarına nazaran simetrik dağılımlara yönelik olması ve Gaussian dağılımlar için etkinliğinin (%37) düşük olması bu tahmincinin dezavantajlarıdır (Rousseeuw and Croux, 1993).

### 3.3.2. $S_n$ Tahmincisi

$MAD$ 'ye göre daha yüksek etkinliğe (%58) sahip olup asimetric dağılımlar içinde uygun bir ölçek tahmincisi olan  $S_n$  tahmincisi,

$$S_n = c_n 1.1926 med_i \{ med_{j \neq i} |x_i - x_j| \} \quad (3.27)$$

şeklinde hesap edilir. Bu eşitlikten de görüldüğü üzere  $S_n$  tahmincisi hesaplanırken örneklemdaki değerlerin kendi aralarındaki uzaklıkları kullanılır. Dolayısıyla  $MAD$  tahmincisinin hesap edilebilmesi için gerekli olan örneklemin konum tahminine,  $S_n$  tahmincisi hesaplanırken ihtiyaç yoktur. Bu durum tahmincinin asimetric dağılımlar için geçerli olmasını sağlar (Rousseeuw ve Croux, 1993). (3.27) eşitliğindeki  $c_n$  ve 1.1926 sabitleri sonlu örneklem durumunda tahmincinin yansız olmasını sağlar.  $c_n$  sabitinin değeri aşağıdaki gibidir (Croux ve Rousseeuw, 1992).

**Çizelge 3.3.**  $n \leq 9$  için  $c_n$  değerleri

n	2	3	4	5	6	7	8	9
$c_n$	0.743	1.851	0.954	1.351	0.993	1.198	1.005	1.131

$n > 9$  için,  $n$  tek ise,

$$c_n = \frac{n}{n - 0.9} \quad (3.28)$$

ve  $n$  çift olduğunda da  $c_n = 1$  değerini alır (Rousseeuw ve Croux, 1993).



### 3.3.3. $Q_n$ Tahmincisi

MAD ve  $S_n$  tahmincilerine göre daha yüksek etkinliğe (%82) sahip olmasının yanında etki fonksiyonu sürekli olan  $Q_n$  tahmincisi,

$$Q_n = d_n 2.2219 \{|x_i - x_j|; i < j\}_{(k)} \quad (3.29)$$

şeklindedir. Eşitliğindeki  $k$  değeri  $h = [n/2] + 1$  olmak üzere  $k = \binom{h}{2} \approx \binom{n}{2}/4$  dir.  $d_n$  sabitinin değerleri ise,  $n \leq 9$  olduğunda Çizelge 3.4 ten yararlanarak,

**Çizelge 3.4.**  $n \leq 9$  için  $d_n$  değerleri

n	2	3	4	5	6	7	8	9
$d_n$	0.399	0.994	0.512	0.844	0.611	0.857	0.669	0.872

$n$  tek olduğunda,

$$d_n = \frac{n}{n + 1.4} \quad (3.30)$$

ve  $n$  çift olduğunda da,

$$d_n = \frac{n}{n + 3.8} \quad (3.31)$$

şeklinde hesap edilir (Croux ve Rousseeuw, 1992; Rousseeuw ve Croux, 1993).

## 4. DAĞILIM PARAMETRELERİNDE ROBUST YÖNTEMLER

### 4.1. Doğrusal regresyon modeli yardımıyla robust tahminleme

Açık olarak ifade edilen dağılım fonksiyonu  $F(x)$ 'in doğrusal forma dönüştürülebildiği dağılımlar için parametre tahmini, doğrusal regresyon modelinin katsayılarının tahminine dönüşür. Doğrusal regresyon modeli,

$$y = X\beta + u \quad (4.1)$$

şeklinde ifade edilir. Bu eşitlikte  $X$  ilk sütunu 1 değerini alan (modelde sabit terimin olduğu durumda) ve geriye kalan  $p$  tane sütununu bağımsız değişkenlerin oluşturduğu  $n \times (p + 1)$  boyutlu matris,  $y = y_1, y_2, \dots, y_n$  olmak üzere  $n$  boyutlu bağımlı değişken vektörü,  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  olmak üzere  $(p + 1)$  boyutlu katsayı vektörü ve  $u = (u_1, u_2, \dots, u_n)$  olmak üzere hata terimlerinin oluşturduğu vektördür. (4.1) yardımıyla tahmin edilen  $y$  değerleri,

$$\hat{y} = X\hat{\beta} \quad (4.2)$$

şeklinde hesap edilir.  $e = (e_1, e_2, \dots, e_n)$  artıklar vektörü olmak üzere,

$$e = y - \hat{y} \quad (4.3)$$

olarak elde edilir.

Dağılım fonksiyonu doğrusal hale getirildiğinde bağımlı ya da bağımsız değişkenlerden biri  $F(x)$  e bağlı olarak elde edilecektir. Bu yüzden  $F(x)$  değerinin de regresyon modelinin katsayılarının tahmininden önce tahmin edilmesi gerekir. Dağılım fonksiyonunun  $F(x) \sim U(0, 1)$  özelliğine bağlı olarak literatürde parametrik olmayan çeşitli tahminleri mevcuttur (Tiryakioğlu ve Hudak, 2007). Bu tahminler,

$$\hat{F} = \frac{i - a}{n + b} \quad (4.4)$$

genel formunda yazılabilir. Burada  $i$ , veri seti küçükten büyüğe sıralandıktan sonra, gözlem noktasının sırasını,  $n$  gözlem sayısını,  $a$  ve  $b$  ise  $0 \leq a \leq 0.5$ ,

$0 \leq b \leq 1$  özelliklerinin sağlayan sayılardır (Tiryakioğlu ve Hudak, 2007). Bu tezde  $\hat{F}$  değeri, örneklem tam olduğunda etkili sonuç veren ve literatürde sıklıkla kullanılan  $\hat{F} = \frac{i-0.3}{n+0.4}$  olarak alınmıştır (Zhang ve ark., 2006). Dağılımın lineer bir modele dönüşmesi sayesinde, doğrusal regresyon modeli için geliştirilen tüm robust yöntemler kolayca dağılım parametrelerinin tahmininde de kullanılmaktadır (Zhang ve ark., 2006; Boudt ve ark., 2011; Brazauskas ve Serfling, 2000; Altın, 2012; Skinner ve ark., 2001; Brazauskas ve Serfling, 2001; Zyl, 2012; Zyl ve Schall, 2012; Yıldırım ve Kantar, 2014). Ayrıca eğer dağılım fonksiyonu lineer olmayan bir modele dönüşürse lineer olmayan model için kullanılan robust yöntemler uygulanabilir (Wang ve Lee, 2010; Kantar ve Yıldırım, 2014).

#### 4.1.1. M-Tahminciler

Üçüncü bölümde anlatılan M-tahminciler, (4.1) modelinin parametrelerinin tahmininde kullanılabilir. Bu modelin M-tahmincisi  $M_n$  olmak üzere,

$$\min \sum_{i=1}^n \rho(e_i) \quad (4.5)$$

minimizasyon probleminin modelin katsayılarına göre çözümü ile elde edilir. Burada  $\rho$  fonksiyonu sürekli ve çoğunlukla konveks bir fonksiyon olup türevi  $\psi$  fonksiyonudur. Bu problemin çözümü sonucu elde edilen  $M_n$ ,

$$M_n(Y + Xb) = M_n(Y) + b \quad \forall b \in \mathbb{R}^p \quad (4.6)$$

eşitliğini sağlar. Dolayısıyla regresyon equivaryantlık özelliğine sahiptir. Fakat  $M_n$  tahmincisi genellikle,

$$M_n(cY) = cM_n(Y) \quad c > 0 \quad (4.7)$$

eşitliğini sağlamaz yani  $M_n$  scale equivaryantlık özelliğine sahip değildir. Bu özelliğe sahip M-tahmincisi elde etmek için (4.5) problemi,

$$\min \sum_{i=1}^n \rho\left(\frac{e_i}{S_n}\right) \quad (4.8)$$

şeklinde ele alınarak yani artık değerler standartlaştırılarak çözülmesi gerekir. Burada  $S_n$  uygun ölçek tahmincisi olup regresyon ve scale equivariantlık özelliklerini sağlayan  $M_n$  elde edebilmek için bu ölçek tahmincisinin scale equivariant ve regresyon invariant özelliklerine sahip olması gerekir. Bu da  $S_n$ 'nin,

$$S_n(c(Y + Xb)) = cS_n(Y) \quad \forall b \in \mathbb{R}^p, c > 0 \quad (4.9)$$

eşitliğini sağlaması anlamına gelir.

(4.5) probleminde  $\rho$  yerine bu fonksiyonun türevi olan  $\psi$  fonksiyonu yazılırsa,

$$\sum_{i=1}^n x_i \psi\left(\frac{e_i}{S_n}\right) = 0 \quad (4.10)$$

eşitliği elde edilmiş olur ve  $M_n$  bu sistemin köklerinden biridir. Bu eşitlikte çözüm, 3. bölümde gösterilen monoton  $\psi$  fonksiyonlarından biri kullanıldığında monoton regresyon M-tahmincisi, yeniden azalan  $\psi$  fonksiyonlarından biri kullanıldığında ise yeniden azalan regresyon M-tahmincisi adını alır. (4.10) eşitliği için Monoton  $\psi$  kullanıldığında elde edilen çözümlerin tamamı (4.5) problemi için de geçerlidir. Eğer  $\psi$  artan veya monoton ise sistem tek bir çözüme sahiptir (Jureckova ve Picek, 2006; Maronna ve ark., 2006).

(4.10) eşitliğinde  $\psi$  fonksiyonu yerine,  $r_i = \frac{e_i}{S_n}$  olmak üzere,  $w_i = \frac{\psi(r_i)}{r_i}$  ağırlık fonksiyonu yazıldığında eşitlik,

$$\sum_{i=1}^n x_i w_i r_i = 0 \quad (4.11)$$

ağırlıklandırılmış ortalama şekline dönüşür. Bu eşitlik iteratif yöntemlerle çözülebilir. Bu yöntemler içerisinde en yaygın kullanılanı Yeniden Ağırlıklandırılmış En Küçük Kareler (Iteratively Reweighted Least Squares-IRLS) yöntemidir (Holland ve Welsch, 1977; Yıldırım, 2012).

**Yeniden Ağırlıklandırılmış En Küçük Kareler (Iteratively Reweighted Least Squares-IRLS) Tekniği:** Bu teknik adından da anlaşılacağı üzere gözlem değerlerinin ağırlıklarının her adımda hesaplanarak en küçük kareler yönteminin uygulanmasıdır. Kullanılan bu tekrarlı yöntemin ilk adımında gerekli olan ağırlıkların hesap edilebilmesi için bir başlangıç tahmincisine ihtiyaç vardır. Bu başlangıç tahmincisi yardımıyla artık değerler hesap edilir

ve bu artık değerler aracılığıyla tahmin ediciye uygun ölçek tahmincisi hesap edilerek kullanılan M-tahmincisinin (3.20) eşitliğinde verilen ağırlık fonksiyonu ile ilk ağırlıklar bulunur. Bu ağırlıklar ile IRLS tekniği kullanılarak doğrusal regresyon modelinin katsayıları tahmin edilir. Bu işlem, ardışık olarak tahmin edilen, iki adımın katsayıları arası farkın, belirli bir değer altına düşmesiyle sonlandırılır. Bu yöntemin algoritması aşağıdaki gibi özetlenebilir (Andersen, 2008; Stuart, 2011):

**Adım I:** Başlangıç tahmincisi aracılığıyla ilk artık değerleri hesaplanır.

**Adım II:** Ölçek tahmincisi hesap edilir.

**Adım III:** Ölçek tahmincisi aracılığıyla artıklar standartlaştırılır.

**Adım IV:** M-tahmincilerinin ağırlık fonksiyonu aracılığıyla artıkların ağırlıkları bulunur.

**Adım V:** Hesap edilen ağırlıklar kullanılarak en küçük kareler tekniği aracılığıyla regresyon modelinin katsayıları,  $\hat{\beta} = (X'WX)^{-1}X'WY$  yardımıyla hesap edilir. Burada  $W$ , köşegenleri artık değerlerin ağırlıklarından oluşan köşegen matristir.

**Adım VI:** Bu işlem, ardışık iki adımın katsayıları arasındaki fark belirli bir değerden küçük ise sonlandırılır değilse bulunan katsayılarla yeni artık değerler hesaplanarak 3. adımdan itibaren tekrar edilir.

**M-Tahmincisi için Etki Fonksiyonu:** (4.1) modelindeki  $X$  bağımsız değişken matrisinin rassal olduğu durumu ele alalım. Bu durumda  $(x'_i, Y_i)'$ ,  $i = 1, 2, \dots, n$  olmak üzere, dağılım fonksiyonu  $F(x, y)$  olan aynı dağılımlı rassal vektör olur. M-tahmincisinin amaç fonksiyonu olan  $\rho$  nun türevi  $\psi$  sürekli ve (4.1) modelinin katsayıları tahmincisi  $T(F)$  olmak üzere bu tahminci,

$$\int x\psi(y - x'T(P))dF(x, y) = 0 \quad (4.12)$$

sisteminin çözümü ile elde edilir.

$$F_\varepsilon = (1 - \varepsilon)F + \varepsilon\delta(x_0, y_0) \quad 0 \leq \varepsilon \leq 1 \quad (4.13)$$

olmak üzere  $F_\varepsilon$   $\varepsilon$  oranında  $\delta(x_0, y_0)$  ile kirletilmiş dağılım fonksiyonudur. Bu dağılım için regresyon modelinin parametre tahmincisi  $T(F_\varepsilon)$ ,

$$(1 - \varepsilon) \int x\psi(y - x'T(F_\varepsilon))dF(x, y) + \varepsilon x_0\psi(y_0 - x'_0T(F_\varepsilon)) = 0 \quad (4.14)$$

sisteminin çözümü ile elde edilir. Etki fonksiyonu  $IF(x_0, y_0; T, F) = \frac{dT(F_\varepsilon)}{d\varepsilon} |_{t=0}$  olarak elde edildiğinden (4.14) eşitliğinde  $\varepsilon$  a göre türev alındığında,

$$\begin{aligned} & - \int x\psi(y - x'T(F_\varepsilon))dF(x, y) - (1 - \varepsilon) \int x'x \frac{dT(F_\varepsilon)}{d\varepsilon} \psi'(y - x'T(F_\varepsilon))dF(x, y) \\ & + x_0\psi(y_0 - x'_0T(F_\varepsilon)) - tx'_0x_0 \frac{dT(F_\varepsilon)}{d\varepsilon} \psi'(y_0 - x'_0T(F_\varepsilon)) = 0 \end{aligned}$$

denklemi elde edilir.  $\varepsilon$  yerine 0 koyulup, (4.12) eşitliği dikkate alındığında yukarıdaki denklem,

$$IF(x_0, y_0; T, F) \int x'x\psi'(y - x'T(F))dF(x, y) = x_0\psi(y_0 - x'_0T(F)) \quad (4.15)$$

haline dönüşür. Sonuç olarak etki fonksiyonu  $B = \int x'x\psi'(y - x'T(F))dF(x, y)$  olmak üzere,

$$IF(x_0, y_0; T, F) = B^{-1}x_0\psi(y_0 - x'_0T(F)) \quad (4.16)$$

olarak elde edilir. Elde edilen etki fonksiyonu incelendiğinde  $y_0$  değeri  $\psi$  fonksiyonu içerisinde yer almaktadır. M-tahminciler sınırlı  $\psi$  fonksiyonuna sahip olduklarından tahminci  $y_0$ 'a göre sınırlıdır dolayısıyla dayanıklıdır. Fakat aynı durum  $x_0$  için geçerli değildir. Etki fonksiyonu incelendiğinde  $x_0$ 'ın, tahmin edicinin üzerinde sınırsız etkiye sahip olabileceği açıkça görülmektedir. Dolayısıyla M-tahmincisi  $X$ -yönlü aykırı değere göre dayanıklı değildir.

#### 4.1.2. MM Tahmincisi

Yüksek kırılma noktasına sahip bir tahminci ile M-tahmincilerin birleşimi olan bu yöntem Yohai (1987) tarafından önerilmiştir. MM-tahmincileri,

- (i) Öncelikle yüksek kırılma noktalı bir tahmin edici yardımıyla  $\beta^*$  elde edilir. Bu aşamada kullanılan tahmincinin etkinliği dikkate alınmaz.
- (ii) İlk aşamada elde edilen tahminci ile artıklar hesaplanır ve bu artıklar aracılığıyla  $MAD$  ya da  $S_n$  gibi dayanıklı (kırılma noktaları %50) ölçek tahmincisi elde edilir.

(iii) Son olarak MM-tahmincisi  $\hat{\beta}$

$$\sum_{i=1}^n x_i \psi\left(\frac{e_i}{S_n}\right) = 0 \quad (4.17)$$

eşitliğinin çözümü ile elde edilir.

(4.17) eşitliğinde kullanılan  $\psi$  fonksiyonu uygun olarak seçilen M-tahmincisinin  $\psi$  fonksiyonudur. Dolayısıyla yüksek etkinliğe sahiptir. Böylece yüksek kırılma noktasına sahip başlangıç tahmincisi ile tahmin sürecine başlanıp etkinliği yüksek olan M-tahmincisi de son aşamada kullanılarak MM-tahmincisi hesaplanmıştır. Dolayısıyla yüksek kırılma noktasına sahip ve etkin bir tahminci elde edilmiştir.

#### 4.1.3. En Küçük Mutlak Sapmalar (Least Absolute Deviation-LAD) Tahmincisi

$L_1$  tahmincisi olarak anılan bu yöntem, hataların mutlak değerlerinin minimize edilmesi prensibine dayanır. (4.1) modelinin katsayılarının LAD tahmincileri,

$$\min_{\beta} \sum_{i=1}^n |e_i| \quad (4.18)$$

denklemleri yardımıyla elde edilir. En küçük mutlak sapmalar tahmincisi  $y$ 'nin içerdiği aykırı değerlerden en küçük kareler tahmincisine göre daha dayanıklı olmasına karşın  $x$  yönlü aykırı değerlere (kaldıraç noktası) karşı dayanıklı değildir. Bu yüzden kırılma noktası 0'dır. En küçük kareler yöntemine göre aykırı değerlerden daha az etkilenen bu metot M-tahmincilerinin başlangıç tahmincisi olarak da kullanılabilir (Andersen, 2008; Yıldırım, 2012).

#### 4.1.4. En Küçük Medyan Kareler (Least Median of Squares-LMS) Tahmincisi

İlk olarak Rousseeuw (1984) tarafından önerilen bu tahminci hata karelerinin medyanını minimize etme prensibine dayanır. En küçük medyan kareler tahmincisinin amaç fonksiyonu,

$$\min_{\beta} \{Med(e_i^2)\} \quad (4.19)$$

şeklinde. En küçük kareler yönteminin aksine toplam operatörünün yerine medyanın kullanılması tahminciyi aykırı değerlere karşı dayanıklı hale getirmiştir. Dolayısıyla en küçük medyan kareler tahmincisi medyanın sahip olduğu %50 kırılma noktası değerine sahiptir. Bu avantajlarının olmasına karşın düşük etkinliğe sahip olmasının yanında iyi tanımlanmış etki fonksiyonu da bulunmamaktadır. En küçük medyan kareler tahmincisi yüksek kırılma noktasına sahip olduğu için MM-tahmincilerinin elde edilmesinde önemli bir yere sahiptir (Andersen, 2008).

#### 4.1.5. Kırılmış En Küçük Kareler (Least Trimmed Squares-LTS) Tahmincisi

1984 yılında Rousseeuw tarafından önerilen bu yöntemde kırılmış hata karelerinin toplamı minimize edilerek regresyon modelinin katsayıları tahminine ulaşılır. Bu yöntemin amaç fonksiyonu,

$$\min_{\beta} \sum_{i=1}^h e_{(i)}^2 \quad (4.20)$$

şeklinde. Burada  $e_{(i)}$  artık terimlerin  $i$ 'inci sıralı istatistik değeri olup  $h = [n(1 - \alpha) + 1]$  tahmin sürecine dahil edilen gözlem sayısını ve  $\alpha$  ise kırılma yüzdesini ifade etmektedir.  $h = n/2$  alındığında kırılma noktasının değeri % 50 olmakta ve tahminci olabilecek en yüksek dayanıklılık düzeyine ulaşmaktadır. Fakat daha etkin bir tahminci elde edebilmek için daha yüksek bir  $h$  değeri alınması gerekir, yani daha fazla gözlem tahmin sürecine dahil edilmelidir. Kırılmış en küçük kareler tahmincisi de MM-tahmincilerinin elde edilmesi için kullanılan başlangıç tahmincisi olarak seçilebilir (Andersen, 2008).

#### 4.1.6. Tekrarlı Medyanlar (Repeated Median-RepMed) Tahmincisi

1982 yılında Siegel tarafından geliştirilmiş olan tekrarlı medyanlar tahmincisi %50 kırılma noktası değerine sahiptir.  $Y = \beta_0 + \beta_1 X + \varepsilon$  modelindeki  $\beta_1$  parametresi için tekrarlı medyanlar tahmincisi,

$$\hat{\beta}_1 = \text{Med}_j \text{Med}_{i \neq j} \frac{y_j - y_i}{x_j - x_i} \quad (4.21)$$



$\beta_0$  için,

$$\hat{\beta}_0 = Med_j Med_{i \neq j} \frac{x_j y_i - x_i y_j}{x_j - x_i} \quad (4.22)$$

ya da  $\hat{\beta}_1$  yardımıyla

$$Med_i (Y_i - \hat{\beta}_1 X_i) \quad (4.23)$$

denklemleri yardımıyla elde edilir.

#### 4.1.7. Theil Tahmincisi

Theil tarafından 1950 yılında aykırı değerlerin olduğu durumlarda diğer robust tahmincilere alternatif olarak önerilen bu yöntem, parametrik değildir ve hesaplanırken hata dağılımları dikkate alınmaz. Veri sayısı büyüdükçe ve değişken sayısı ikiden fazla olduğunda bu tekniğin uygulaması zordur.  $x_1, \dots, x_n$  ve  $y_1, \dots, y_n$  gözlem değerleridir ve  $x$  küçükten büyüğe sıralandığında  $1 \leq i < j \leq n$  olmak üzere,

$$b_{ij} = \frac{y_j - y_i}{x_j - x_i} \quad (4.24)$$

şeklinde hesaplanan  $n$ 'in 2 lisi kadar eğim değerlerinin medyanının alınması ile,

$$\hat{\beta}_1 = Med\{b_{ij} | x_i \neq x_j\} \quad (4.25)$$

eğim katsayısının Theil tahmincisi elde edilir. Sabit terimin Theil tahmincisi ise,

$$\hat{\beta}_0 = Med\{y_i - \hat{\beta}_1 x_i\} \quad (4.26)$$

şeklinde hesap edilir. Bilindiği üzere medyan dayanıklı bir tahmincidir dolayısıyla Theil tahmincisi de veri setinde bulunan aykırı değerlere karşı dayanıklıdır (Mutan ,2004).

#### 4.2. Kantillere Dayalı Tahminleme

Kantillere dayalı tahminleme yöntemi Marks (2005) tarafından örneklemden alınan simetrik yüzdeliğe dayalı olarak hesap edilmiştir. Dağılım ve kantil fonksiyonu,

$$F(x) = P \quad (4.27)$$

$$x = F^{-1}(P) \quad (4.28)$$

şeklinde. Simetrik  $P_1$  ve  $P_2$  değerleri için,

$$x_{P_1} = F^{-1}(P_1) \quad (4.29)$$

$$x_{P_2} = F^{-1}(P_2) \quad (4.30)$$

bu denklemler yardımıyla dağılım parametreleri tahmin edilebilir. Açıktır ki denklem sayısı bilinmeyen parametre sayısı kadar olmalıdır.

### 4.3. Konum-Ölçek Ailesine Dönüştürülebilen Dağılımlar için Robust Tahminleme

$f(x)$  olasılık yoğunluk fonksiyonu,  $\sigma \in \mathbb{R}^+$  ve  $\mu \in \mathbb{R}$  olmak üzere  $g(x; \mu, \sigma) = (1/\sigma)f((x - \mu)/\sigma)$  özelliğini sağlayan  $g(x; \mu, \sigma)$  olasılık yoğunluk fonksiyonlarının

$$C^* = \{g(x; \mu, \sigma) : g(x; \mu, \sigma) = (1/\sigma)f((x - \mu)/\sigma), \sigma > 0, \mu \in \mathbb{R}\} \quad (4.31)$$

şeklindeki sınıfına olasılık yoğunluk fonksiyonlarının konum-ölçek ailesi denir. Burada  $\mu$  konum parametresi  $\sigma$  ise ölçek parametresidir. Konum-ölçek ailesine dahil olan dağılımların parametreleri dahil olunan ailenin konum ve ölçek parametrelerinin tahmin edilmesi ile elde edilebilir. Bunun için de robust konum ve ölçek tahmincileri kullanılabilir. Olive (2006)'daki çalışmasında konum için robust tahminci olan medyanı, ölçek için ise yine robust tahmincisi olan MAD'yi kullanmış ve bu tahminciyi *Med/MAD* tahmincisi olarak adlandırmıştır (Boudt ve ark., 2011). Cauchy  $C(\mu, \sigma)$ , Laplace  $L(\theta, \lambda)$ , üstel  $EXP(\lambda)$ , iki parametrelili üstel  $EXP(\theta, \lambda)$ , Gumbel  $G(\theta, \sigma)$ , logistik  $L(\mu, \sigma)$ , normal  $N(\mu, \sigma)$ , log-Weibull  $LW(\theta, \sigma)$  ve düzgün  $U(\theta_1, \theta_2)$  dağılımlarının medyan ve *MAD* değerleri ile parametrelerinin robust tahmincileri Çizelge 4.1 ve Çizelge 4.2'de sırasıyla verilmiştir (Olive, 2006).

**Çizelge 4.1.** Uygulamada Kullanılan Çeşitli Dağılımlar için  $Med(X)$  ve  $MAD(X)$  değerleri

İsim	$f(y)$	$Med(X)$	$MAD(X)$
$C(\mu, \sigma)$	$1/(\pi\sigma[1 + (\frac{x-\mu}{\sigma})^2])$	$\mu$	$\sigma$
$L(\theta, \lambda)$	$\frac{1}{2\lambda} \exp(-\frac{ x-\theta }{\lambda})$	$\theta$	$0.6931\lambda$
$EXP(\lambda)$	$\frac{1}{\lambda} \exp(-\frac{x}{\lambda})I(x \geq 0)$	$0.6931\lambda$	$\lambda/2.0781$
$EXP(\theta, \lambda)$	$\frac{1}{\lambda} \exp(-\frac{(x-\theta)}{\lambda})I(x \geq 0)$	$\theta + 0.6931\lambda$	$\lambda/2.0781$
$G(\theta, \sigma)$	$\frac{1}{\sigma} \exp(-\frac{x-\theta}{\sigma}) \exp[-\exp(-\frac{x-\theta}{\sigma})]$	$\theta + 0.3665\sigma$	$0.7671\sigma$
$L(\mu, \sigma)$	$\frac{\exp(-(x-\mu)/\sigma)}{\sigma[1+\exp(-(x-\mu)/\sigma)]^2}$	$\mu$	$1.0986\sigma$
$N(\mu, \sigma)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$	$\mu$	$0.6745\sigma$
$LW(\theta, \sigma)$	$\frac{1}{\sigma} \exp(\frac{x-\theta}{\sigma}) \exp[-\exp(\frac{x-\theta}{\sigma})]$	$\theta - 0.3665\sigma$	$0.7671\sigma$
$U(\theta_1, \theta_2)$	$\frac{1}{\theta_2-\theta_1} I(\theta_1 < x < \theta_2)$	$(\theta_1 + \theta_2)/2$	$(\theta_2 - \theta_1)/4$

**Çizelge 4.2.** Çizelge 4.1 teki dağılımlar için robust tahminciler

$C(\mu, \sigma)$	$\hat{\mu} = Med(n)$	$\hat{\sigma} = MAD(n)$
$L(\theta, \lambda)$	$\hat{\theta} = Med(n)$	$\hat{\lambda} = 1.443MAD(n)$
$EXP(\lambda)$	$\hat{\lambda}_1 = 1.443Med(n)$	$\hat{\lambda}_2 = 2.0781MAD(n)$
$EXP(\theta, \lambda)$	$\hat{\theta} = Med(n) - 1.440MAD(n)$	$\hat{\lambda} = 2.0781MAD(n)$
$G(\theta, \sigma)$	$\hat{\theta} = Med(n) - 0.4778MAD(n)$	$\hat{\sigma} = 1.3037MAD(n)$
$L(\mu, \sigma)$	$\hat{\mu} = Med(n)$	$\hat{\sigma} = 0.9102MAD(n)$
$N(\mu, \sigma)$	$\hat{\mu} = Med(n)$	$\hat{\sigma} = 1.483MAD(n)$
$LW(\theta, \sigma)$	$\hat{\theta} = Med(n) + 0.4778MAD(n)$	$\hat{\sigma} = 1.3037MAD(n)$
$U(\theta_1, \theta_2)$	$\hat{\theta}_1 = Med(n) - 2MAD(n)$	$\hat{\theta}_2 = Med(n) + 2MAD(n)$

Çizelge 4.2'deki  $Med(n)$  ve  $MAD(n)$  değerleri sırasıyla örneklemin medyan ve MAD değerlerini ifade etmektedir. Bu tahmin yapılırken, dağılımın teorik  $MAD$  ve  $Med$  değeri verinin  $MAD$  ve  $Med$  değeri ile eşleştirilir.

#### 4.4. Medyanlar Metodu

Sınırlı etki fonksiyonuna sahip ve aykırı değerlere karşı oldukça dayanıklı olan bu yöntemin uygulanmasında skor fonksiyonu kullanılır.  $f(x)$  herhangi bir olasılık yoğunluk fonksiyonu olmak üzere skor fonksiyonu,

$$\varphi(x) = \frac{\partial \log f(x)}{\partial \theta} \quad (4.32)$$

şekindedir. Bu denklemdeki  $\theta$  dağılımın parametre vektörüdür.

Medyanlar metoduna bağılı olarak tahminçiler, örneklemin skor fonksiyonunun medyanı ile karşılık gelen kitlenin skor fonksiyonunun medyanının eşitlenmesi ile elde edilir. Yani  $X = (x_1, x_2, \dots, x_n)$  olasılık yoğunluk fonksiyonu  $f(x)$  olan dağılımdan çekilmiş örneklem olmak üzere bu dağılımın parametreleri medyanlar metodu ile

$$Med_i(\varphi(x_i)) = Med(\varphi(x)) \quad (4.33)$$

eşitliğinin çözümü ile elde edilir. Sadece tek bir eşitliğin çözümüne dayalı olduğu için hesaplanması oldukça kolay bir tahminçidir.

Medyanlar metodu,

$$\sum_i h(\psi(x_i)) = 0 \quad (4.34)$$

yardımıyla M-tahminçileri formunda da yazılabilir. Bu eşitlikte  $h(x) = x$  olarak alındığında en çok olabilirlik tahminçisi,  $h(x) = \text{sign}(x)$  olarak alındığında ise medyanlar metodu tahminçisi elde edilir (He ve Fung, 1999).  $Med(\varphi(x))$  in hesaplanması her dağılım için mümkün olmadığından bu metodun her dağılım için uygulanması zordur.

#### 4.5. Kırpılmış Momentler Metodu (Method Of Trimmed Moments - MTM)

Bajorunaite ve Brazauskas (2008) tarafından önerilen bu yöntemde tahminçiler, parametre sayısı kadar kırpılmış örneklem momentlerinin karşılık gelen kırpılmış kitle momentlerine eşitlenmesi ile elde edilir. Sıralı istatistikleri  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  olan  $X_1, X_2, \dots, X_n, \theta_1, \theta_2, \dots, \theta_k$  parametrelili  $F$  dağılımından çekilen bağımsız rassal değişkenler olsun.  $\theta_1, \theta_2, \dots, \theta_k$  parametrelerinin MTM tahminçileri genel olarak aşağıdaki adımları izleyerek elde edilir.

(i)  $k$  tane örneklem kırpılmış momentleri hesaplanır.

$$\hat{\mu}_j = \frac{1}{n - m_n(j) - m_n^*(j)} \sum_{i=m_n(j)+1}^{n-m_n^*(j)} h_j(X_{(i)}), \quad j = 1, 2, \dots, k \quad (4.35)$$

Bu denklemde  $m_n(j)$  ve  $m_n^*(j)$  değerleri,  $0 \leq m_n(j) \leq n - m_n^*(j) \leq n$  ve  $n \rightarrow \infty$  iken  $m_n(j)/n \rightarrow a_j$  ve  $m_n^*(j) \rightarrow b_j$  özelliklerini sağlayan tam sayılardır.  $a_j$  ve  $b_j$  ise  $0 \leq a_j + b_j < 1$  olmak üzere kırpılma oranlarıdır.

- (ii) Birinci adımda hesaplanan kırılmış örneklem momentlerine karşılık gelen kırılmış kitle momentleri hesaplanır.

$$\mu_j = \frac{1}{1 - a_j - b_j} \int_{a_j}^{1-b_j} h_j(F^{-1}(u)) du, \quad j = 1, 2, \dots, k \quad (4.36)$$

Bu denklemde  $F^{-1}$  kantil fonksiyonudur.

- (iii) Elde edilen kitle ve örneklem kırılmış momentleri,

$$\begin{cases} \mu_1(\theta_1, \theta_2, \dots, \theta_k) = \hat{\mu}_1 \\ \vdots \\ \mu_k(\theta_1, \theta_2, \dots, \theta_k) = \hat{\mu}_k \end{cases} \quad (4.37)$$

şeklinde eşitlenerek parametrelere göre çözüldüğünde MTM tahmincileri elde edilir.

(4.35) ve (4.36) denklemlerindeki  $h_j$  fonksiyonu genellikle  $h_j(x) = x^j$  olarak seçilir ki bu durumda  $a_j = b_j = 0$  seçildiğinde klasik momentler metodu tahmincileri elde edilir. Örneğin konum-ölçek ailelerinden olan dağılımlar için MTM tahmincileri elde edileceği zaman  $h_1(x) = x$  ve  $h_2(x) = x^2$  olarak alınacaktır. Logaritmik dönüşüm ile konum-ölçek ailesine dönüştürülebilen dağılımlar için genellikle  $h_1(x) = \log(x)$  ve  $h_2(x) = (\log(x))^2$  olarak alınır (Bajrunaite ve Brazauskas, 2008). Seçilen momentler genelde dağılımın yeterli istatistiği olmaktadır.

## 5. DAĞILIMLAR İÇİN TAHMİNCİLERİN HESAPLANMASI

Bu bölümde dördüncü bölümde anlatılan robust tahmincilerin Weibull ve Pareto dağılımları için uygulanabilirliği incelenecektir. Ayrıca, Kantar ve Arık (2014) çalışmasında bu tezde ele alınan bazı tahmincileri Log-lojistik dağılımı için de incelemiştir.

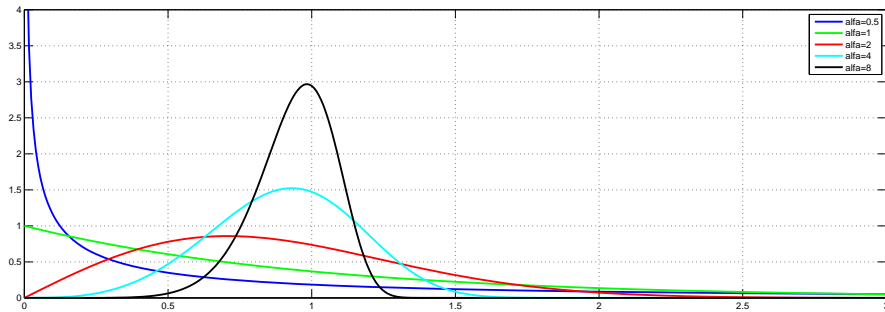
### 5.1. Weibull Dağılımı

Weibull dağılımı güvenilirlik analizi, yaşam analizi ve ekonomi alanlarında en çok kullanılan dağılımdır. Bu sebeple dağılımın robust veya robust olmayan pek çok tahmincisi önerilmiştir (Kantar, 2006; Kantar ve Şenoğlu, 2008; Boudt ve ark., 2011). Weibull dağılımının olasılık yoğunluk ve dağılım fonksiyonları sırasıyla;

$$f(x) = \frac{\alpha}{\gamma} (x/\gamma)^{\alpha-1} \exp[-(x/\gamma)^\alpha], \quad x, \gamma, \alpha > 0 \quad (5.1)$$

$$F(x) = 1 - \exp[-(x/\gamma)^\alpha] \quad (5.2)$$

şeklinde olup  $\gamma$  ve  $\alpha$  parametreleri sırasıyla ölçek ve şekil parametresi olarak adlandırılır. Ölçek parametresi  $\gamma = 1$  olan Weibull dağılımının farklı şekil parametresi değerleri ile oluşturulan olasılık yoğunluk fonksiyonu grafikleri şekil 5.1'de ve şekil parametresinin farklı değerleri için Weibull dağılımının çarpıklık ve basıklık değerleri Çizelge 5.1'de verilmiştir.



Şekil 5.1. Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği

**Çizelge 5.1.** Şekil parametresinin farklı değerleri için Weibull dağılımının çarpıklık ve basıklık değerleri

	$\alpha$				
	0.5	1	2	4	8
Çarpıklık	6.6188	2	0.6311	-0.0872	-0.5337
Basıklık	87.7200	9	3.2451	2.7478	3.3277

### 5.1.1. Doğrusal Regresyon Modeli Yardımıyla Weibull Dağılımının Parametrelerinin Tahmini

(5.2) eşitliğinde verilen Weibull dağılım fonksiyonu,

$$F(x) = 1 - \exp[-(x/\gamma)^\alpha] \quad (5.3)$$

$$1 - F(x) = \exp[-(x/\gamma)^\alpha] \quad (5.4)$$

$$\ln(1 - F(x)) = -(x/\gamma)^\alpha \quad (5.5)$$

$$\ln[-\ln(1 - F(x))] = \alpha \ln(x) - \alpha \ln(\gamma) \quad (5.6)$$

$P = F(x)$  olmak üzere;

$$\ln[-\ln[1 - P]] = \alpha \ln(x) - \alpha \ln(\gamma) \quad (5.7)$$

şeklinde doğrusal formda yazılabilir.

Doğrusal regresyon modeli yardımıyla Weibull dağılımının parametrelerini tahmin etmek için, (5.7) denklemini,  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$   $n$  boyutlu örneklemin sıralı istatistik değerleri olmak üzere,

$$\ln[-\ln(1 - P_{(i)})] = \alpha \ln(x_{(i)}) - \alpha \ln(\gamma) \quad (5.8)$$

$$Y_i = \ln[-\ln(1 - P_{(i)})] \quad (5.9)$$

$$X_i = \ln(x_{(i)}) \quad (5.10)$$

$$\beta_1 = \alpha, \quad \beta_0 = -\alpha \ln(\gamma) \quad (5.11)$$

$$Y_i = \beta_0 + \beta_1 X_i \quad (5.12)$$

şeklinde basit doğrusal regresyon modeli elde edilir. Bu model aracılığıyla dördüncü bölümde doğrusal regresyon modeli, robust tahminleme başlığı altında verilen tahminciler ile regresyon parametreleri tahmin edilebilir. Tahmin

edilen regresyon parametreleri ve (5.11) eşitlikleri aracılığıyla ters dönüşüm yapılarak Weibull dağılımı için tahminler,

$$\hat{\alpha} = \hat{\beta}_1 \quad (5.13)$$

$$\hat{\gamma} = \exp[-(\hat{\beta}_0/\hat{\alpha})] \quad (5.14)$$

olarak elde edilmiş olur.

### 5.1.2. Weibull Dağılımı İçin Kantillere Dayalı Tahminleme

Weibull dağılımı için bu yöntem kullanılarak parametreler hesaplanmak istendiğinde (5.2)' de verilen Weibull dağılım fonksiyonu yardımıyla,

$$\exp[-(x/\gamma)^\alpha] = 1 - F(x) \quad (5.15)$$

$$-(x/\gamma)^\alpha = \ln(1 - F(x)) \quad (5.16)$$

$$x = \gamma[-\ln(1 - F(x))]^{1/\alpha} \quad (5.17)$$

şeklinde kantil değeri elde edilir. Parametre tahmini için birinci ve üçüncü kantil değerlerinin ( $P_{25}$  ve  $P_{75}$ ) kullanıldığı varsayımı altında,

$$P_{75} = \gamma[-\ln(0.25)]^{1/\alpha} \quad (5.18)$$

$$P_{25} = \gamma[-\ln(0.75)]^{1/\alpha} \quad (5.19)$$

(5.18) ve (5.19) denklemleri taraf tarafa oranlandığında,

$$(P_{75}/P_{25})^\alpha = \ln(0.25)/\ln(0.75) \quad (5.20)$$

(5.20) denkleminde her iki tarafın logaritması alınır,

$$\hat{\alpha} = \frac{\ln[\ln(0.25)/\ln(0.75)]}{\ln(P_{75}/P_{25})} \quad (5.21)$$

(5.21) eşitliği ile  $\alpha$  şekil parametresinin tahminci elde edilmiş olur. Ölçek parametresi ise (5.18) ve (5.19) denklemlerinden herhangi birisi kullanılarak,

$$\hat{\gamma} = P_{75}/[-\ln(0.25)]^{1/\hat{\alpha}} \quad (5.22)$$

şeklinde elde edilir. Simetrik yüzdeler kullanılması ile elde edilen bu tahminci, 10'uncu ve 90'uncü kantiller ile en iyi sonucu vermektedir (Marks, 2005). Ayrıca tahmincinin Weibull dağılımı için kırılma noktası en yüksek değerine 1/3'üncü ve 2/3'üncü kantiller kullanıldığında ulaşmaktadır (Boudt ve ark., 2011).



### 5.1.3. Konum-Ölçek Ailesine Dönüştürerek Weibull Dağılımının Parametrelerinin Tahmini

$X \sim Weibull(\gamma, \alpha)$  olan bir rassal değişken olmak üzere  $\log(X) \sim \log - Weibull(\mu = \log(\gamma), \sigma = 1/\alpha)$  dağılımına sahiptir. Yani log-Weibull dağılımı konum parametresi  $\mu = \log(\gamma)$  ve ölçek parametresi  $\sigma = 1/\alpha$  olan konum-ölçek ailesinde yer alır. Dolayısıyla Weibull dağılımının parametrelerinin tahmini, gözlemlerin logaritmasının dahil olduğu konum-ölçek ailesinin konum ve ölçek parametrelerinin tahmini problemi gibi düşünülebilir. Olive (2006)'daki çalışmasında konumun robust tahmincisi için median tahmincisini ve ölçeğin robust tahmincisi için  $MAD$  tahmincisini kullanmış ve  $Med/MAD$  tahmincisi olarak adlandırmıştır. Bu tahminci kullanılarak konum ve ölçeğin robust tahmincileri sırasıyla,

$$\hat{\mu} = Med(\log(X)) + 0.4778MAD(\log(X)) \quad (5.23)$$

$$\hat{\sigma} = 1.3037MAD(\log(X)) \quad (5.24)$$

olarak elde edilir. Robust tahmini yapılan konum ve ölçek parametrelerinden ters dönüşüm yapılarak Weibull dağılımının parametrelerinin tahmincileri,

$$\hat{\gamma} = \exp(\hat{\mu}) \quad (5.25)$$

$$\hat{\alpha} = 1/\hat{\sigma} \quad (5.26)$$

şeklinde elde edilir.

### 5.1.4. Medyanlar Metodu ile Weibull Dağılımının Parametrelerinin Tahmini

Dördüncü bölümde belirtildiği gibi bu metoda bağlı olarak tahminciler, örneklemin skor fonksiyonunun medyanı ile karşılık gelen kitlenin skor fonksiyonunun medyanının eşitlenmesi ile elde edilir. Olasılık yoğunluk fonksiyonunun parametrelere göre kısmi türevi alınarak elde edilen skor fonksiyonu,

$$\varphi(x) = \frac{\partial \log f(x)}{\partial \theta} = \left( \begin{array}{c} \frac{\partial f(x)}{\partial \gamma} \\ \frac{\partial f(x)}{\partial \alpha} \end{array} \right) = \left( \begin{array}{c} \frac{\alpha}{\gamma} [(x/\gamma)^\alpha - 1] \\ \frac{1}{\alpha} [1 + (\alpha \log(x/\gamma))(1 - (x/\gamma)^\alpha)] \end{array} \right)$$

şeklinde dir. Örneklemin skor fonksiyonunun medyanı ile kitleninki eşitlendiğinde,

$$Med(\varphi(X_i)) = Med(\varphi(X)) \quad (5.27)$$

$$Med\left\{\left(\frac{X_i}{\gamma}\right)^\alpha\right\} = \log(2) \quad (5.28)$$

$$Med\{(1 - (X_i/\gamma)^\alpha) \log(X_i/\gamma)\} = c \quad (5.29)$$

eşitlikleri elde edilir. (5.28) eşitliğinin sağ tarafındaki  $Y = (X/\gamma)^\alpha$  ifadesi parametresi bir olan üstel dağılıma sahiptir ve bu dağılımın medyanı  $\log(2)$ 'dir. (5.29) eşitliğindeki  $c$  değeri ise  $Med((1-Y)\log(Y)) \approx -0.51$  olarak hesaplanır. Bu eşitliklerin iteratif olarak çözümü ile Weibull dağılımının medyan metoduna bağlı olarak parametreleri elde edilmiş olur (He ve Fung, 1999) .

### 5.1.5. Kırpılmış Momentler Metodu ile Weibull Dağılımının Parametrelerinin Tahmini

Bölüm 5.1.3'te de ifade edildiği gibi Weibull dağılımının logaritmik dönüşümü sonucu elde edilen dağılım konum-ölçek ailesine dahildir. Dolayısıyla dönüşüm sonucu elde edilen konum-ölçek ailesinin parametrelerinin kırpılmış momentler tahmincisinin hesaplanmasının ardından ters dönüşüm yapılarak Weibull dağılımının parametreleri elde edilebilir.  $X \sim Weibull(\gamma, \alpha)$  olan bir rassal değişken olmak üzere  $Y = \log(X) \sim log - Weibull(\theta = \log(\gamma), \sigma = 1/\alpha)$  dağılımına sahiptir. Yani log-Weibull dağılımı konum parametresi  $\mu = \log(\gamma)$  ve ölçek parametresi  $\sigma = 1/\alpha$  olan konum ölçek ailesinde yer alır. Dolayısıyla,

$$F(y) = F_0\left(\frac{y - \theta}{\sigma}\right) \quad (5.30)$$

şeklinde ifade edilebilir. Bu denklemdeki  $F_0$ , konum parametresi  $\theta = 0$  ve ölçek parametresi  $\sigma = 1$  olan standart dağılım fonksiyonudur.  $F(y)$  dağılımının kantil fonksiyonu standart dağılım fonksiyonunun kantil fonksiyonu yardımıyla,

$$F^{-1}(t) = \theta + \sigma F_0^{-1} \quad (5.31)$$

şeklinde ifade edilebilir.  $F(y)$  dağılım fonksiyonu iki parametrelili olduğu için iki tane kırpılmış moment hesaplanması parametre tahmini için yeterlidir. Bölüm 4.5'te anlatılan adımlar  $h_1(t) = t$  ve  $h_2(t) = t^2$  alınarak uygulandığında ilk olarak örneklemin kırpılmış momentleri,

$$\hat{\mu}_1 = \frac{1}{n - m_n(1) - m_n^*(1)} \sum_{i=m_n(1)+1}^{n-m_n^*(1)} X_{(i)} \quad (5.32)$$

$$\hat{\mu}_2 = \frac{1}{n - m_n(2) - m_n^*(2)} \sum_{i=m_n(2)+1}^{n-m_n^*(2)} X_{(i)}^2 \quad (5.33)$$

şeklinde hesaplanır. Bu denklemlerdeki  $m_n(1) = m_n(2) \rightarrow a$  ve  $m_n^*(1) = m_n^*(2) \rightarrow b$  olacak şekilde belirlenir.

Bir sonraki adımda kitlenin kırılmış momentleri aşağıdaki gibi elde edilir:

$$\mu_1 = \frac{1}{1-a-b} \int_a^{1-b} F^{-1}(u) du \quad (5.34)$$

$$= \frac{1}{1-a-b} \int_a^{1-b} (\theta + \sigma F_0^{-1}) du = \theta + \sigma c_1 \quad (5.35)$$

$$\mu_2 = \frac{1}{1-a-b} \int_a^{1-b} (F^{-1}(u))^2 du \quad (5.36)$$

$$= \frac{1}{1-a-b} \int_a^{1-b} (\theta + \sigma F_0^{-1})^2 du = \theta^2 + 2\theta\sigma c_1 + \sigma^2 c_2 \quad (5.37)$$

Yukarıdaki denklemlerdeki  $c_1$  ve  $c_2$  değerleri,

$$c_k = \frac{1}{1-a-b} \int_a^{1-b} [F^{-1}(u)]^k du \quad (5.38)$$

işleminin sonucu olup farklı kırılma oranları için  $c_1$  ve  $c_2$  değerleri Çizelge 5.2 ve 5.3'de verilmiştir.

**Çizelge 5.2.**  $F_0$ : Log-Weibull Dağılımı için  $c_1$  değeri

		$b$							
$a$	0	0.05	0.10	0.15	0.25	0.49	0.70	0.85	
0	-0.5772	-0.6791	-0.7700	-0.8585	-1.0367	-1.5218	-2.1217	-2.8579	
0.05	-0.3980	-0.4956	-0.5810	-0.6632	-0.8263	-1.2543	-1.7494	-2.2954	
0.10	-0.2773	-0.3735	-0.4567	-0.5360	-0.6921	-1.0938	-1.5441	-2.0201	
0.15	-0.1747	-0.2706	-0.3524	-0.4300	-0.5814	-0.9652	-1.3855	-	
0.25	0.0033	-0.0935	-0.1743	-0.2500	-0.3956	-0.7555	-1.1358	-	
0.49	0.3757	0.2688	0.1849	0.1087	-0.0335	-0.3666	-	-	
0.70	0.7103	0.5805	0.4865	0.4041	0.2556	-	-	-	
0.85	1.0165	0.8450	0.7338	-	-	-	-	-	

**Çizelge 5.3.**  $F_0$ : Log-Weibull Dağılımı için  $c_2$  değeri

		$b$							
$a$	0	0.05	0.10	0.15	0.25	0.49	0.70	0.85	
0	1.9779	1.9822	2.0412	2.1294	2.3818	3.4855	5.5883	9.2067	
0.05	1.1939	1.1550	1.1688	1.2080	1.3467	2.0304	3.3313	5.3734	
0.10	0.8908	0.8317	0.8262	0.8451	0.9387	1.4669	2.5015	4.0964	
0.15	0.7022	0.6276	0.6081	0.6129	0.6756	1.1017	1.9699	-	
0.25	0.4881	0.3876	0.3467	0.3304	0.3491	0.6378	1.2939	-	
0.49	0.3979	0.2352	0.1517	0.0975	0.0438	0.1346	-	-	
0.70	0.6497	0.4005	0.2706	0.1803	0.0670	-	-	-	
0.85	1.1190	0.7309	0.5415	-	-	-	-	-	

Tahmin sürecinin son adımında kitle ve örneklem momentleri eşitlendiğinde,

$$\hat{\mu}_1 = \theta + \sigma c_1 \quad (5.39)$$

$$\hat{\mu}_2 = \theta^2 + 2\theta\sigma c_1 + \sigma^2 c_2 \quad (5.40)$$

denklemleri elde edilir. Bu denklemlerin çözümü sonucu log-Weibull dağılımının kırılmış moment tahmincileri (MTM),

$$\hat{\theta}_{MTM} = \hat{\mu}_1 - c_1 \hat{\sigma}_{MTM} \quad (5.41)$$

$$\hat{\sigma}_{MTM} = \sqrt{(\hat{\mu}_2 - \hat{\mu}_1^2)/(c_2 - c_1^2)} \quad (5.42)$$

olarak elde edilir. Bu tahminciler yardımıyla Weibull dağılımının ölçek ve şekil tahmincileri ters dönüşüm yapılarak,

$$\hat{\gamma}_{MTM} = \exp(\hat{\theta}_{MTM}) \quad (5.43)$$

$$\hat{\alpha}_{MTM} = 1/\hat{\sigma}_{MTM} \quad (5.44)$$

şeklinde elde edilir (Bajrunaite ve Brazauskas, 2008).

## 5.2. Pareto Dağılımı

Aktüerya, telekomünikasyon ve güvenilirlik analizi gibi alanlarda kullanılan Pareto dağılımının olasılık yoğunluk ve dağılım fonksiyonları sırasıyla,

$$f(x) = \frac{\alpha\sigma^\alpha}{x^{\alpha+1}}, \quad \sigma, \quad \alpha > 0, \quad x \geq \sigma \quad (5.45)$$

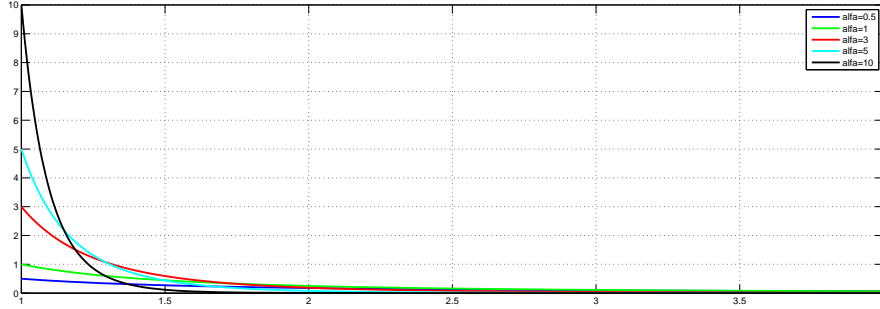
$$F(x) = 1 - \left(\frac{\sigma}{x}\right)^\alpha \quad (5.46)$$

biçiminde ifade edilip  $\sigma$  ve  $\alpha$  parametreleri sırasıyla ölçek ve şekil parametresi olarak adlandırılır. Ölçek parametresi  $\sigma = 1$  olan Pareto dağılımının farklı şekil parametresi değerleri ile oluşturulan olasılık yoğunluk fonksiyonu grafikleri Şekil 5.2’de verilmiştir.

Pareto dağılımının basıklık katsayısı şekil parametresi  $\alpha > 4$  için tanımlı olup,  $\alpha$ ’nın farklı değerleri için bu dağılımın çarpıklık ve basıklık değerleri Çizelge 5.4’te verilmiştir.

**Çizelge 5.4.** Şekil parametresinin farklı değerleri için Pareto dağılımının çarpıklık ve basıklık değerleri

	$\alpha$				
	4.5	5	6	8	10
Çarpıklık	5.4659	4.6476	3.8103	3.1177	2.8111
Basıklık	146.4444	70.8000	35.6667	19.7250	14.8286



**Şekil 5.2.** Pareto dağılımının olasılık yoğunluk fonksiyonları grafiği

### 5.2.1. Doğrusal Regresyon Modeli Yardımıyla Pareto Dağılımının Parametrelerinin Tahmini

(5.46) denkleminde verilen Pareto dağılımının dağılım fonksiyonu,

$$F(x) = 1 - \left(\frac{\sigma}{x}\right)^\alpha \quad (5.47)$$

$$1 - F(x) = \left(\frac{\sigma}{x}\right)^\alpha \quad (5.48)$$

$$\ln(1 - F(x)) = \alpha \ln \sigma - \alpha \ln x \quad (5.49)$$

$F(x) = P$  olmak üzere

$$\ln(1 - P) = \alpha \ln \sigma - \alpha \ln x \quad (5.50)$$

şeklinde doğrusal formda yazılabilir.

Doğrusal regresyon modeli yardımıyla Pareto dağılımının parametrelerini tahmin etmek için, (5.50) denklemi,  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$   $n$  boyutlu örneklemin sıralı

istatistik değerleri olmak üzere,

$$\ln(1 - P_{(i)}) = \alpha \ln \sigma - \alpha \ln x_{(i)} \quad (5.51)$$

$$Y_i = \ln(1 - P_{(i)}) \quad (5.52)$$

$$X_i = \ln(x_{(i)}) \quad (5.53)$$

$$\beta_1 = -\alpha, \quad \beta_0 = \alpha \ln(\sigma) \quad (5.54)$$

$$Y_i = \beta_0 + \beta_1 X_i \quad (5.55)$$

basit doğrusal regresyon modeli şeklinde elde edilir. Bu model aracılığıyla dördüncü bölümde doğrusal regresyon modeli yardımıyla robust tahminleme başlığı altında verilen tahminciler ile regresyon parametreleri elde edilebilir. Tahmin edilen regresyon parametreleri ve (5.54) eşitlikleri aracılığıyla ters dönüşüm yapılarak Pareto dağılımının şekil ve ölçek parametreleri sırasıyla;

$$\hat{\alpha} = -\hat{\beta}_1 \quad (5.56)$$

$$\hat{\sigma} = \exp[(\hat{\beta}_0/\hat{\alpha})] \quad (5.57)$$

olarak elde edilmiş olur.

### 5.2.2. Pareto Dağılımı İçin Kantillere Dayalı Tahminleme

Pareto dağılımı için bu yöntem kullanılarak parametreler hesaplanmak istendiğinde (5.46)' de verilen Pareto dağılımının dağılım fonksiyonu yardımıyla,

$$(\sigma/x)^\alpha = 1 - F(x) \quad (5.58)$$

$$(\sigma/x) = (1 - F(x))^{1/\alpha} \quad (5.59)$$

$$x = \sigma[(1 - F(x))]^{-1/\alpha} \quad (5.60)$$

şeklinde kantil değeri elde edilir. Parametre tahmini için birinci ve üçüncü kantil değerlerinin ( $P_{25}$  ve  $P_{75}$ ) kullanıldığı varsayımı altında,

$$P_{75} = \sigma(1 - 0.75)^{-1/\alpha} \quad (5.61)$$

$$P_{25} = \sigma(1 - 0.25)^{-1/\alpha} \quad (5.62)$$

(5.61) ve (5.62) denklemleri taraf tarafa oranlandığında,

$$(P_{75}/P_{25}) = \frac{0.25^{-1/\alpha}}{0.75^{-1/\alpha}} \quad (5.63)$$

(5.63) denkleminde her iki tarafın logaritması alınırsa,

$$\hat{\alpha} = -\frac{[\ln(0.25/0.75)]}{\ln(P_{75}/P_{25})} \quad (5.64)$$

(5.64) eşitliği ile  $\alpha$  şekil parametresinin tahmincisi elde edilmiş olur. Ölçeğin tahmincisi ise (5.61) ve (5.62) denklemlerinden herhangi birisi kullanılarak,

$$\hat{\sigma} = P_{75}/(1 - 0.75)^{-1/\hat{\alpha}} \quad (5.65)$$

şeklinde elde edilir. Bu tahminlemede elde edilen tahminciler açık olarak ifade edilmiştir.

### 5.2.3. Konum-Ölçek Ailesine Dönüştürerek Pareto Dağılımının Parametrelerinin Tahmini

$X \sim Pareto(\sigma, \alpha)$  olan bir rassal değişken olmak üzere  $\log(X) \sim Ustel(\mu = \log(\sigma), \sigma = 1/\alpha)$  iki parametrelili üstel dağılıma sahiptir. Yani log-Pareto dağılımı konum parametresi  $\mu = \log(\sigma)$  ve ölçek parametresi  $\sigma = 1/\alpha$  olan konum ölçek ailesinde yer alır.  $Med/MAD$  tahmincisi kullanılarak konum ve ölçeğin robust tahmincileri sırasıyla,

$$\hat{\mu} = Med(\log(X)) - 1.440MAD(\log(X)) \quad (5.66)$$

$$\hat{\sigma} = 2.078MAD(\log(X)) \quad (5.67)$$

olarak elde edilir. Robust tahmini yapılan konum ve ölçek parametrelerinden ters dönüşüm yapılarak Pareto dağılımı için tahminciler,

$$\hat{\sigma} = \exp(\hat{\mu}) \quad (5.68)$$

$$\hat{\alpha} = 1/\hat{\sigma} \quad (5.69)$$

şeklinde elde edilir.

### 5.2.4. Medyanlar Metodu ile Pareto Dağılımının Parametrelerinin Tahmini

Pareto dağılımının medyanlar metodu tahmincileri için önce skor fonksiyonu hesaplanmalıdır.

$$\varphi(x) = \frac{\partial \log f(x)}{\partial \theta} = \begin{pmatrix} \frac{\partial f(x)}{\partial \sigma} \\ \frac{\partial f(x)}{\partial \alpha} \end{pmatrix} = \begin{pmatrix} \frac{\alpha}{\sigma} \\ \frac{1}{\alpha} + \log \sigma - \log x \end{pmatrix}$$

Örneklemin skor fonksiyonunun medyanı ile kitleninki eşitlendiğinde,

$$Med(\varphi(X_i)) = Med(\varphi(X)) \quad (5.70)$$

$$Med\{\log(X_i)\} = Med\{\log(X)\} \quad (5.71)$$

denklemleri elde edilir. Görüldüğü üzere  $\sigma$ 'ya göre kısmi türev alınarak elde edilen denklemler için örneklem ve kitlenin medyan değerleri aynı olmakta ve sonuç

olarak (5.71) eşitliği elde edilmektedir. Bu eşitliğin sağ tarafındaki  $Y = \log X$  ifadesi iki parametrelili üstel dağılıma sahip olduğu kolayca gösterilebilir. Yani  $\log X \sim Exp(X)$  ve bu dağılım,

$$Med\{Y\} = \log\sigma + \alpha \log 2 \quad (5.72)$$

medyan değerine sahiptir. (5.71) eşitliğinin sol tarafındaki örneklemin medyan değeri, iki parametrelili üstel dağılımın medyanına eşitlendiğinde Pareto dağılımının şekil parametresi olan  $\alpha$ 'nın medyanlar metoduna bağlı tahmincisi,

$$\hat{\alpha} = \frac{Med\{\log x\} - \log\sigma}{\log 2} \quad (5.73)$$

olarak elde edilir. Skor fonksiyonlarının medyanlarının eşitlenmesi sonucu tek bir denklem elde edildiğinden, bu metotla dağılımın şekil parametresinin tahmincisi elde edilebilmiştir.

### 5.2.5. Kırılmış Momentler Metodu ile Pareto Dağılımının Parametrelerinin Tahmini

Kırılmış momentler metoduna bağlı olarak tahmincilerin elde edilmesi için bölüm 4.5'te verilen adımlar Pareto dağılımı için uygulansın. Pareto dağılımının ilk iki kırılmış momentinin hesaplanması için gerekli olan  $h_1(t) = \log t$  ve  $h_2(t) = (\log t)^2$  olarak alındığında örneklemin kırılmış momentleri,

$$\hat{\mu}_1 = \frac{1}{n - m_n(1) - m_n^*(1)} \sum_{i=m_n(1)+1}^{n-m_n^*(1)} \log(X_{(i)}) \quad (5.74)$$

$$\hat{\mu}_2 = \frac{1}{n - m_n(2) - m_n^*(2)} \sum_{i=m_n(2)+1}^{n-m_n^*(2)} (\log X_{(i)})^2 \quad (5.75)$$

şeklinde hesaplanır. Bu denklemlerdeki  $m_n(1) = m_n(2) \rightarrow a$  ve  $m_n^*(1) = m_n^*(2) \rightarrow b$  olacak şekilde belirlenir.

Pareto dağılımının kantil fonksiyonu  $F^{-1}(u) = \sigma(1 - u)^{-1/\alpha}$  kullanılarak



kitlenin kırılmış momentleri aşağıdaki gibi elde edilir,

$$\mu_1 = \frac{1}{1-a-b} \int_a^{1-b} \log(F^{-1}(u)) du \quad (5.76)$$

$$= \frac{1}{1-a-b} \int_a^{1-b} \log(\sigma(1-u)^{-1/\alpha}) du \quad (5.77)$$

$$= \frac{1}{1-a-b} \left\{ \int_a^{1-b} \log(\sigma) du - 1/\alpha \int_a^{1-b} \log(1-u) du \right\} \quad (5.78)$$

$$= \log(\sigma) + \frac{1}{1-a-b} (-1/\alpha) c_1 \quad (5.79)$$

$$\mu_2 = \frac{1}{1-a-b} \int_a^{1-b} (\log(F^{-1}(u)))^2 du \quad (5.80)$$

$$= \frac{1}{1-a-b} \int_a^{1-b} (\log(\sigma(1-u)^{-1/\alpha}))^2 du \quad (5.81)$$

$$= \frac{1}{1-a-b} \int_a^{1-b} (\log \sigma - (1/\alpha) \log(1-u))^2 \quad (5.82)$$

$$= \frac{1}{1-a-b} \left\{ \int_a^{1-b} (\log(\sigma))^2 du + (1/\alpha^2) \int_a^{1-b} (\log(1-u))^2 du \right. \quad (5.83)$$

$$\left. - \frac{2 \log(\sigma)}{\alpha} \int_a^{1-b} \log(1-u) du \right\} \quad (5.84)$$

$$= (\log(\sigma))^2 + \frac{1}{(1-a-b)\alpha^2} c_2 - \frac{2 \log(\sigma)}{(1-a-b)\alpha} c_1 \quad (5.85)$$

(5.79) ve (5.85) denklemlerindeki  $c_1$  ve  $c_2$  değerleri,

$$c_k = \int_a^{1-b} [\log(1-u)]^k du \quad (5.86)$$

işleminin sonucu olup farklı kırılma oranları için  $c_1$  ve  $c_2$  değerleri Çizelge 5.5 ve 5.6'da verilmiştir.

**Çizelge 5.5.**  $F_0$ : Pareto Dağılımı için  $c_1$  değeri

$a$	$b$							
	0.05	0.10	0.15	0.20	0.25	0.49	0.70	0.85
0.05	-0.7989	-0.6685	-0.5642	-0.4768	-0.4022	-0.1592	-0.0491	-0.0106
0.10	-0.7950	-0.6646	-0.5603	-0.4729	-0.3983	-0.1553	-0.0452	-0.0067
0.15	-0.7884	-0.6579	-0.5536	-0.4663	-0.3916	-0.1486	-0.0385	-
0.20	-0.7787	-0.6483	-0.5439	-0.4566	-0.3819	-0.1390	-0.0288	-
0.25	-0.7660	-0.6355	-0.5312	-0.4439	-0.3692	-0.1262	-0.0161	-
0.49	-0.6536	-0.5231	-0.4188	-0.3315	-0.2568	-0.0139	-	-
0.70	-0.4614	-0.3309	-0.2266	-0.1393	-0.0646	-	-	-
0.85	-0.2348	-0.1043	-	-	-	-	-	-

**Çizelge 5.6.**  $F_0$ : Pareto Dağılımı için  $c_2$  değeri

$a$	$b$							
	0.05	0.10	0.15	0.20	0.25	0.49	0.70	0.85
0.05	1.1517	0.8092	0.5910	0.4381	0.3264	0.0715	0.0116	0.0012
0.10	1.1513	0.8089	0.5906	0.4378	0.3260	0.0712	0.0112	0.0009
0.15	1.1504	0.8080	0.5897	0.4369	0.3251	0.0703	0.0103	-
0.20	1.1486	0.8062	0.5879	0.4350	0.3233	0.0684	0.0085	-
0.25	1.1453	0.8029	0.5846	0.4318	0.3200	0.0652	0.0052	-
0.49	1.0897	0.7473	0.5290	0.3762	0.2644	0.0096	-	-
0.70	0.9090	0.5665	0.3483	0.1954	0.0836	-	-	-
0.85	0.5607	0.2183	-	-	-	-	-	-

Tahmin sürecinin son adımında kitle ve örneklem momentleri eşitlendiğinde,

$$\hat{\mu}_1 = \log(\sigma) + \frac{1}{1-a-b}(-1/\alpha)c_1 \quad (5.87)$$

$$\hat{\mu}_2 = (\log(\sigma))^2 + \frac{1}{(1-a-b)\alpha^2}c_2 - \frac{2\log(\sigma)}{(1-a-b)\alpha}c_1 \quad (5.88)$$

denklemlerine ulaşılır. Bu denklem sisteminin çözümü ile Pareto dağılımının MTM tahminçileri elde edilir.

## 6. SİMULASYON ÇALIŞMASI

Bu bölümde en çok olabilirlik, en küçük kareler tahmincileri ile tezde ele alınan robust tahmincilerin performansları bir simülasyon çalışması yardımıyla Weibull ve Pareto dağılımları için değerlendirilmiştir. Weibull ve Pareto dağılımlarının ele alınmasının nedeni pek çok alanda kabul görmüş ve uygulanmış dağılımlar olup, aynı zamanda robust literatüründe de çalışılmış olmasıdır (Boudt ve ark., 2011; He ve Fung, 1999; Brazauskas ve Serfling, 2000; Yavuz, 2012; Zhang ve ark., 2006; Zyl, 2012). Farklı simülasyon kurgularından veri üretilerek, Weibull ve Pareto dağılımlarının parametreleri ele alınan tahminciler yardımıyla tahmin edilmiştir. Simülasyon kurgusu Boudt ve ark. (2011), He ve Fung (1999) ve Zhang ve ark. (2006) çalışmalarına benzer şekilde oluşturulmuştur. Veride aykırı değer olması durumunda, şekil parametresinin tahmini bu değerden daha çok etkilendiğinden dolayı (Zhang ve ark., 2006; Yavuz, 2012; Brazauskas ve Kleefeld, 2009; Brazauskas ve Serfling, 2000) bu çalışmada şekil parametresi için tahminciler değerlendirilmiştir. Her bir simülasyon kurgusu 100000/ $n$  kez yinelenmiş ve tahmincileri karşılaştırmak için hata kareler ortalaması ( $MSE$ ) ve yanlılık ( $Bias$ ) kriterleri kullanılmıştır. Ayrıca LTS ve LMS tahminlerini elde etmek için LIBRA kütüphanesi kullanılmıştır.

$$Bias(\hat{\alpha}) = \left(\frac{1}{100000/n}\sum \hat{\alpha}\right) - \alpha \quad (6.1)$$

$$MSE(\hat{\alpha}) = \frac{1}{100000/n}\sum (\hat{\alpha} - \alpha)^2 \quad (6.2)$$

Weibull ve Pareto dağılımları için tahmincilerin performansları öncelikle aykırı değerlerin bulunmadığı durumlarda incelenmiştir. Daha sonra veri %10 oranında değişik şekillerde kirletilerek tahmincilerin performansları değerlendirilmiştir.

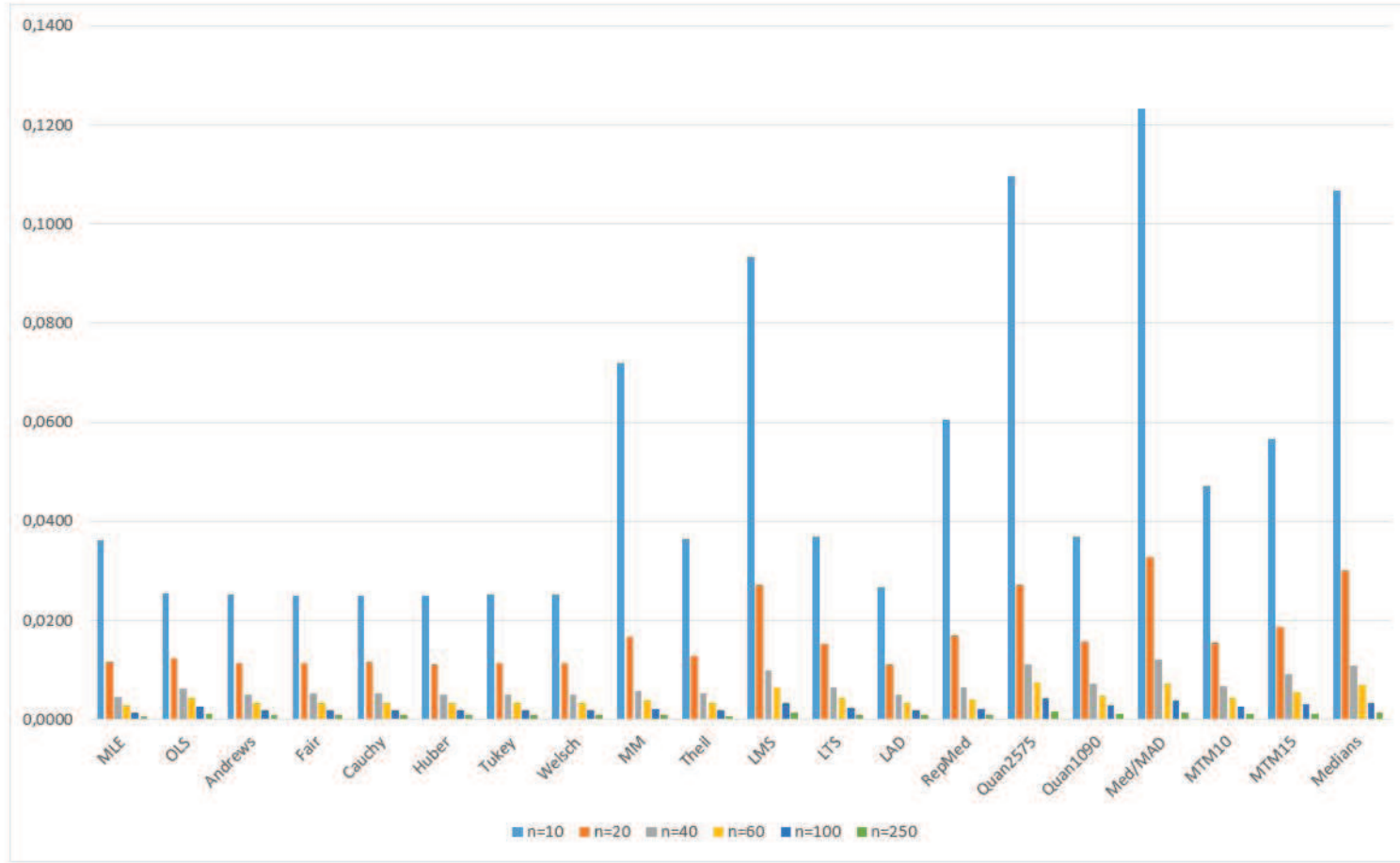
## 6.1. Aykırı Değer Bulunmadığı Durumda Tahmincilerin Performanslarının İncelenmesi

### 6.1.1. Weibull Dağılımı

Weibull dağılımı için simulasyon kurgusu aykırı değerlerin olmadığı durumda örneklem boyutu  $n = 10, 20, 40, 60, 100$  ve  $250$ , şekil parametresi  $\alpha = 0.5, 1, 2, 4$  ve  $8$  ve ölçek parametresi  $\gamma = 1$  şeklinde alınarak incelenmiştir. Kısaca Weib  $(\gamma, \alpha)$ 'dan veri üretilmiştir. Elde edilen sonuçlar Çizelge 6.1-6.5'te ve bu sonuçlar ile oluşturulan grafikler Şekil 6.1-6.5'te verilmiştir.

Çizelge 6.1. Weib ( $\gamma = 1$ ,  $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

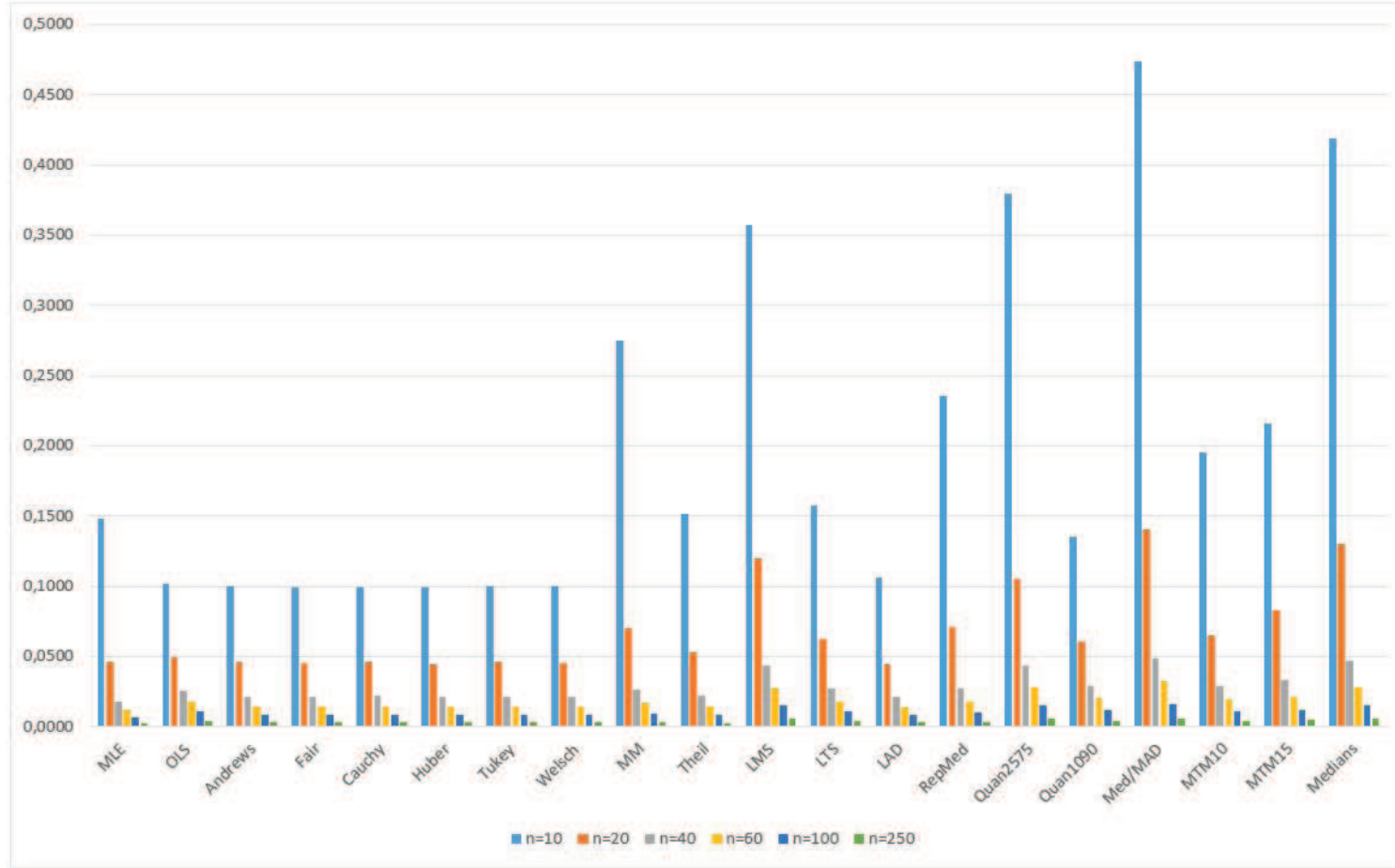
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.0828	0.0361	-0.0372	0.0117	-0.0175	0.0045	-0.0114	0.0028	-0.0063	0.0015	-0.0029	0.0006
LS	0.0170	0.0255	0.0215	0.0123	0.0187	0.0063	0.0163	0.0042	0.0129	0.0027	0.0077	0.0011
Andrews	0.0257	0.0252	0.0138	0.0114	0.0073	0.0051	0.0049	0.0034	0.0037	0.0020	0.0020	0.0009
Fair	0.0221	0.0250	0.0215	0.0113	0.0145	0.0053	0.0108	0.0034	0.0073	0.0020	0.0036	0.0009
Cauchy	0.0218	0.0250	0.0226	0.0115	0.0153	0.0054	0.0116	0.0034	0.0080	0.0020	0.0042	0.0009
Huber	0.0233	0.0249	0.0203	0.0111	0.0131	0.0051	0.0095	0.0033	0.0064	0.0019	0.0031	0.0009
Tukey	0.0256	0.0252	0.0137	0.0114	0.0073	0.0051	0.0049	0.0034	0.0037	0.0019	0.0020	0.0009
Welsch	0.0247	0.0252	0.0140	0.0113	0.0080	0.0051	0.0055	0.0033	0.0039	0.0019	0.0021	0.0009
MM	-0.0472	0.0720	-0.0097	0.0168	-0.0012	0.0059	-0.0002	0.0039	0.0008	0.0021	0.0000	0.0009
Theil	-0.0537	0.0365	-0.0266	0.0129	-0.0131	0.0053	-0.0089	0.0034	-0.0054	0.0019	-0.0023	0.0008
LMS	-0.0681	0.0935	-0.0268	0.0273	-0.0102	0.0098	-0.0083	0.0066	-0.0035	0.0034	-0.0027	0.0014
LTS	-0.0193	0.0370	-0.0090	0.0153	-0.0029	0.0064	-0.0010	0.0042	-0.0005	0.0023	-0.0009	0.0010
LAD	0.0222	0.0266	0.0171	0.0111	0.0106	0.0050	0.0075	0.0033	0.0052	0.0019	0.0025	0.0008
RepMed	-0.0815	0.0604	-0.0383	0.0171	-0.0180	0.0064	-0.0126	0.0041	-0.0070	0.0022	-0.0034	0.0009
Quan2575	-0.1026	0.1098	-0.0432	0.0271	-0.0200	0.0112	-0.0133	0.0075	-0.0074	0.0043	-0.0034	0.0017
Quan1090	-0.0781	0.0369	-0.0384	0.0157	-0.0181	0.0073	-0.0127	0.0048	-0.0074	0.0028	-0.0009	0.0013
Med/MAD	-0.1437	0.1232	-0.0609	0.0328	-0.0279	0.0122	-0.0184	0.0073	-0.0107	0.0039	-0.0040	0.0015
MTM10	-0.0578	0.0471	-0.0207	0.0156	-0.0086	0.0067	-0.0062	0.0044	-0.0038	0.0025	-0.0005	0.0011
MTM15	-0,0308	0,0565	-0,0235	0,0188	-0,0112	0,0091	-0,0083	0,0055	-0,0043	0,0032	-0,0021	0,0013
Medians	-0.1243	0.1067	-0.0537	0.0301	-0.0229	0.0110	-0.0151	0.0070	-0.0051	0.0034	0.0012	0.0013



Şekil 6.1. Weib ( $\gamma = 1$ ,  $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.2. Weib ( $\gamma = 1$ ,  $\alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.1705	0.1483	-0.0719	0.0461	-0.0349	0.0182	-0.0231	0.0115	-0.0141	0.0065	-0.0001	0.0022
LS	0.0329	0.1017	0.0471	0.0492	0.0374	0.0257	0.0319	0.0179	0.0254	0.0109	0.0153	0.0039
Andrews	0.0498	0.1002	0.0314	0.0460	0.0156	0.0216	0.0118	0.0142	0.0061	0.0086	0.0074	0.0029
Fair	0.0428	0.0994	0.0472	0.0454	0.0293	0.0217	0.0217	0.0144	0.0147	0.0086	0.0096	0.0029
Cauchy	0.0421	0.0994	0.0494	0.0464	0.0309	0.0222	0.0232	0.0148	0.0162	0.0087	0.0105	0.0030
Huber	0.0453	0.0992	0.0449	0.0445	0.0264	0.0212	0.0196	0.0141	0.0130	0.0084	0.0090	0.0029
Tukey	0.0496	0.1001	0.0312	0.0460	0.0157	0.0216	0.0118	0.0142	0.0062	0.0086	0.0074	0.0029
Welsch	0.0478	0.0998	0.0321	0.0454	0.0164	0.0214	0.0126	0.0141	0.0070	0.0086	0.0076	0.0029
MM	-0.0915	0.2753	-0.0170	0.0700	-0.0024	0.0263	-0.0011	0.0166	-0.0001	0.0095	0.0001	0.0034
Theil	-0.1127	0.1515	-0.0510	0.0529	-0.0262	0.0225	-0.0187	0.0144	-0.0107	0.0084	0.0003	0.0028
LMS	-0.1314	0.3569	-0.0547	0.1195	-0.0238	0.0432	-0.0166	0.0271	-0.0034	0.0151	-0.0003	0.0056
LTS	-0.0492	0.1573	-0.0175	0.0620	-0.0065	0.0271	-0.0046	0.0178	-0.0022	0.0107	0.0049	0.0037
LAD	0.0415	0.1058	0.0372	0.0447	0.0216	0.0210	0.0156	0.0139	0.0111	0.0083	0.0076	0.0029
RepMed	-0.1637	0.2359	-0.0751	0.0713	-0.0365	0.0272	-0.0260	0.0176	-0.0138	0.0097	-0.0016	0.0033
Quan2575	-0.1991	0.3798	-0.0843	0.1054	-0.0405	0.0437	-0.0271	0.0282	-0.0170	0.0154	-0.0062	0.0055
Quan1090	-0.1307	0.1350	-0.0609	0.0605	-0.0313	0.0292	-0.0230	0.0201	-0.0113	0.0120	-0.0038	0.0041
Med/MAD	-0.2866	0.4737	-0.1255	0.1406	-0.0548	0.0488	-0.0375	0.0321	-0.0214	0.0163	-0.0039	0.0061
MTM10	-0.1223	0.1949	-0.0384	0.0650	-0.0157	0.0287	-0.0099	0.0191	-0.0048	0.0113	0.0017	0.0041
MTM15	-0,0566	0,2155	-0,0535	0,0828	-0,0195	0,0333	-0,0183	0,0216	-0,0086	0,0117	-0,0020	0,0048
Medians	-0.2522	0.4186	-0.1118	0.1303	-0.0466	0.0466	-0.0260	0.0284	-0.0106	0.0148	0.0083	0.0060

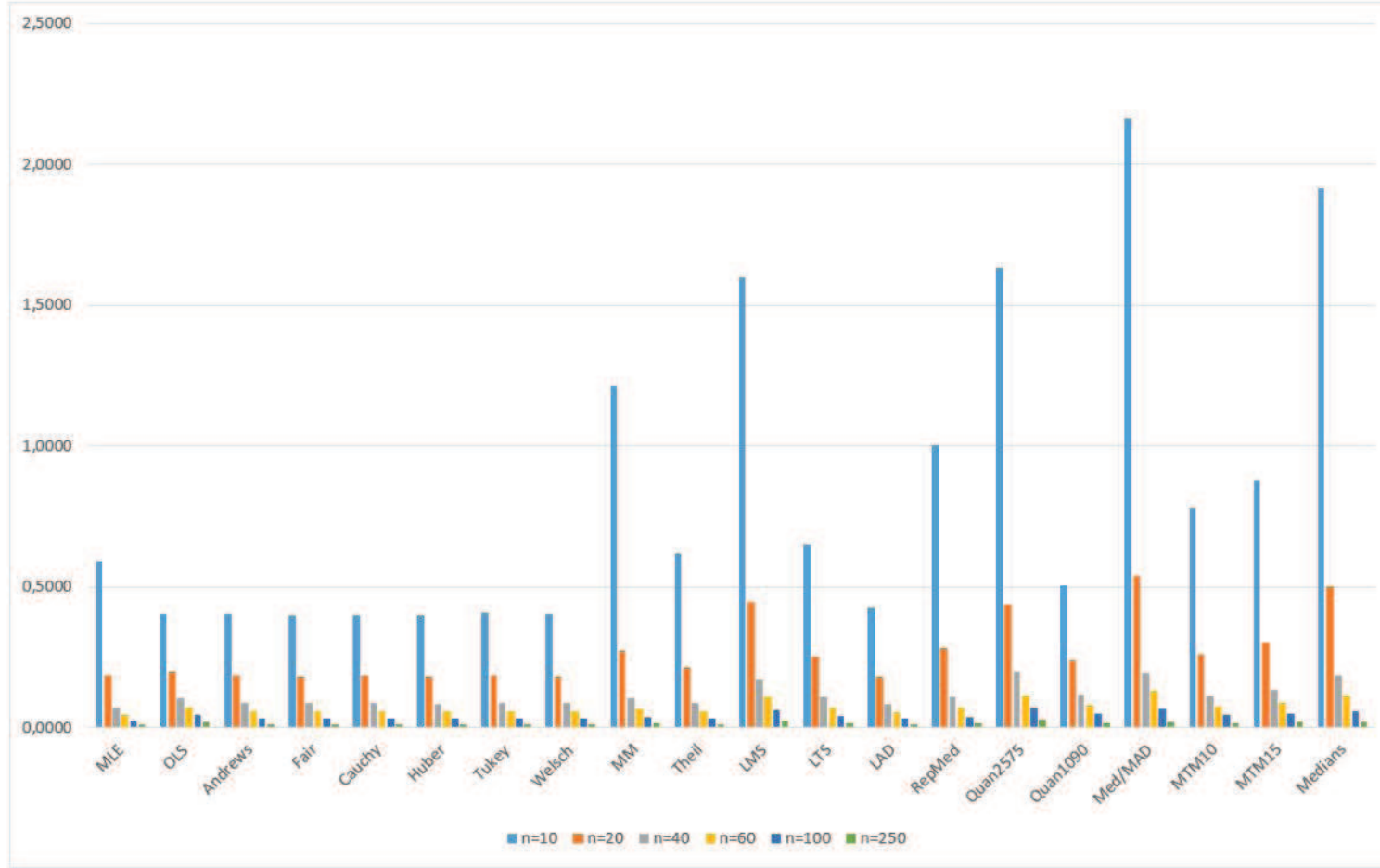


Şekil 6.2. Weib ( $\gamma = 1$ ,  $\alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği



Çizelge 6.3. Weib ( $\gamma = 1$ ,  $\alpha = 2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

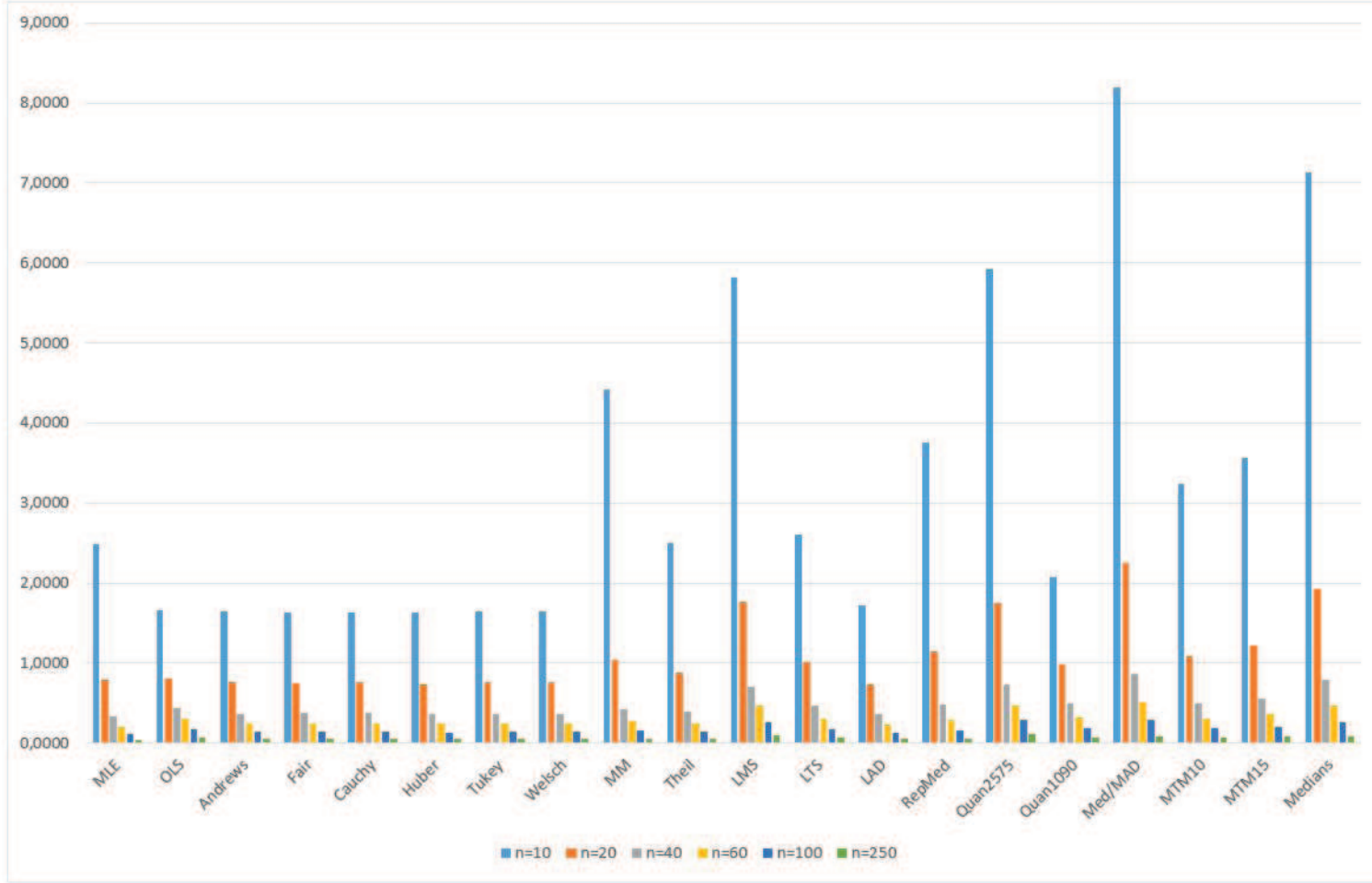
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.3324	0.5902	-0.1484	0.1869	-0.0699	0.0728	-0.0462	0.0458	-0.0282	0.0261	-0.0082	0.0101
LS	0.0711	0.4062	0.0879	0.1980	0.0749	0.1028	0.0638	0.0718	0.0507	0.0437	0.0277	0.0184
Andrews	0.1044	0.4056	0.0572	0.1839	0.0312	0.0865	0.0234	0.0569	0.0121	0.0344	0.0045	0.0132
Fair	0.0914	0.3992	0.0883	0.1824	0.0585	0.0867	0.0434	0.0576	0.0294	0.0343	0.0126	0.0136
Cauchy	0.0895	0.3983	0.0926	0.1863	0.0618	0.0888	0.0465	0.0590	0.0324	0.0350	0.0155	0.0140
Huber	0.0955	0.3983	0.0833	0.1790	0.0528	0.0849	0.0391	0.0564	0.0259	0.0337	0.0108	0.0133
Tukey	0.1037	0.4065	0.0569	0.1839	0.0312	0.0865	0.0235	0.0567	0.0122	0.0344	0.0046	0.0132
Welsch	0.1008	0.4030	0.0590	0.1819	0.0328	0.0858	0.0251	0.0564	0.0140	0.0342	0.0055	0.0131
MM	-0.1946	1.2125	-0.0352	0.2751	-0.0045	0.1052	-0.0021	0.0665	-0.0001	0.0378	-0.0004	0.0143
Theil	-0.2229	0.6186	-0.1053	0.2133	-0.0524	0.0898	-0.0374	0.0578	-0.0215	0.0337	-0.0080	0.0127
LMS	-0.2820	1.5989	-0.1056	0.4483	-0.0476	0.1729	-0.0331	0.1083	-0.0068	0.0605	-0.0022	0.0230
LTS	-0.0892	0.6476	-0.0359	0.2510	-0.0130	0.1086	-0.0092	0.0711	-0.0044	0.0429	-0.0015	0.0170
LAD	0.0915	0.4266	0.0697	0.1796	0.0433	0.0841	0.0313	0.0555	0.0222	0.0334	0.0085	0.0131
RepMed	-0.3353	1.0041	-0.1510	0.2814	-0.0731	0.1089	-0.0520	0.0703	-0.0277	0.0389	-0.0100	0.0150
Quan2575	-0.4180	1.6306	-0.1669	0.4387	-0.0896	0.1972	-0.0578	0.1153	-0.0396	0.0716	-0.0239	0.0275
Quan1090	-0.2072	0.5055	-0.1101	0.2411	-0.0585	0.1166	-0.0440	0.0806	-0.0220	0.0480	-0.0060	0.0174
Med/MAD	-0.5836	2.1647	-0.2439	0.5384	-0.1097	0.1951	-0.0749	0.1285	-0.0428	0.0652	-0.0148	0.0222
MTM10	-0.2377	0.7806	-0.0801	0.2621	-0.0315	0.1148	-0.0199	0.0766	-0.0095	0.0451	-0.0043	0.0167
MTM15	-0.1175	0.8784	-0.1015	0.3051	-0.0418	0.1331	-0.0205	0.0866	-0.0191	0.0517	-0.0017	0.0195
Medians	-0.5097	1.9153	-0.2176	0.5022	-0.0932	0.1863	-0.0520	0.1135	-0.0213	0.0594	0.0065	0.0224



Şekil 6.3. Weib ( $\gamma = 1$ ,  $\alpha = 2$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.4. Weib ( $\gamma = 1$ ,  $\alpha = 4$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

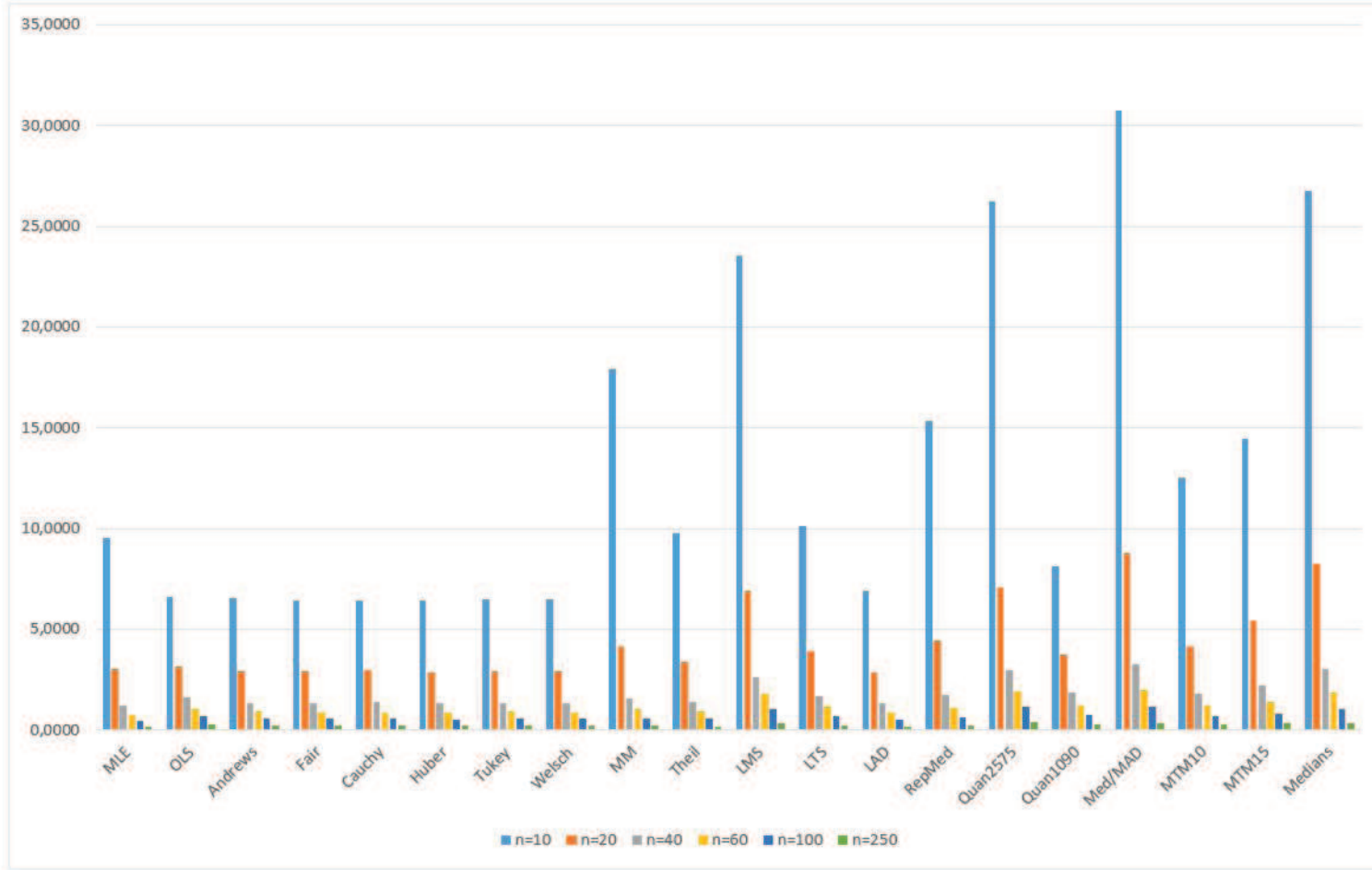
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.6866	2.4931	-0.3062	0.7871	-0.1477	0.3298	-0.0999	0.2065	-0.0587	0.1099	-0.0165	0.0405
LS	0.1271	1.6607	0.1655	0.8094	0.1397	0.4440	0.1196	0.3083	0.1010	0.1755	0.0554	0.0738
Andrews	0.1940	1.6475	0.0998	0.7678	0.0546	0.3709	0.0404	0.2408	0.0275	0.1404	0.0089	0.0529
Fair	0.1683	1.6259	0.1653	0.7487	0.1083	0.3710	0.0813	0.2442	0.0577	0.1362	0.0251	0.0544
Cauchy	0.1640	1.6272	0.1741	0.7645	0.1155	0.3804	0.0876	0.2506	0.0642	0.1389	0.0309	0.0559
Huber	0.1771	1.6273	0.1543	0.7368	0.0977	0.3632	0.0728	0.2390	0.0514	0.1347	0.0216	0.0532
Tukey	0.1929	1.6483	0.0987	0.7678	0.0547	0.3705	0.0408	0.2406	0.0274	0.1400	0.0091	0.0529
Welsch	0.1859	1.6455	0.1031	0.7564	0.0599	0.3667	0.0444	0.2396	0.0310	0.1381	0.0109	0.0525
MM	-0.3782	4.4148	-0.0662	1.0465	-0.0144	0.4230	-0.0083	0.2734	-0.0067	0.1566	-0.0009	0.0572
Theil	-0.4545	2.4988	-0.2174	0.8775	-0.1114	0.3891	-0.0766	0.2428	-0.0491	0.1356	-0.0160	0.0510
LMS	-0.5339	5.8217	-0.2087	1.7660	-0.0937	0.7030	-0.0737	0.4695	-0.0433	0.2568	-0.0044	0.0920
LTS	-0.1935	2.6048	-0.0754	1.0169	-0.0290	0.4671	-0.0269	0.2977	-0.0189	0.1674	-0.0029	0.0679
LAD	0.1654	1.7166	0.1278	0.7280	0.0787	0.3583	0.0584	0.2370	0.0374	0.1329	0.0170	0.0515
RepMed	-0.6585	3.7554	-0.3094	1.1385	-0.1594	0.4810	-0.1108	0.2910	-0.0698	0.1616	-0.0200	0.0600
Quan2575	-0.8273	5.9298	-0.3529	1.7519	-0.1947	0.7353	-0.1288	0.4722	-0.0976	0.2911	-0.0561	0.1193
Quan1090	-0.3721	2.0671	-0.2128	0.9870	-0.1139	0.4893	-0.0683	0.3166	-0.0327	0.1806	-0.0119	0.0695
Med/MAD	-1.1792	8.1900	-0.5151	2.2443	-0.2486	0.8608	-0.1694	0.5111	-0.1097	0.2942	-0.0296	0.0887
MTM10	-0.4981	3.2425	-0.1716	1.0808	-0.0719	0.4928	-0.0443	0.3094	-0.0246	0.1794	-0.0086	0.0670
MTM15	-0.2204	3.5563	-0.1797	1.2260	-0.0721	0.5543	-0.0507	0.3605	-0.0411	0.2079	-0.0179	0.0870
Medians	-1.0297	7.1360	-0.4304	1.9309	-0.2083	0.7879	-0.1209	0.4737	-0.0638	0.2649	0.0131	0.0894



Şekil 6.4. Weib ( $\gamma = 1$ ,  $\alpha = 4$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.5. Weib ( $\gamma = 1$ ,  $\alpha = 8$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-1.3604	9.5319	-0.6122	3.0379	-0.3028	1.2147	-0.2021	0.7343	-0.1300	0.4390	-0.0208	0.1528
LS	0.2655	6.5809	0.3410	3.1559	0.2922	1.6070	0.2546	1.0725	0.1880	0.6961	0.1049	0.2651
Andrews	0.3977	6.5218	0.2103	2.9436	0.0989	1.3575	0.0658	0.9056	0.0341	0.5515	0.0395	0.1954
Fair	0.3443	6.4511	0.3388	2.9034	0.2230	1.3553	0.1671	0.8715	0.1014	0.5488	0.0599	0.2007
Cauchy	0.3379	6.4424	0.3558	2.9678	0.2370	1.3872	0.1806	0.8922	0.1144	0.5597	0.0683	0.2052
Huber	0.3624	6.4397	0.3180	2.8430	0.1993	1.3273	0.1464	0.8589	0.0866	0.5409	0.0552	0.1959
Tukey	0.3959	6.5209	0.2096	2.9420	0.0998	1.3567	0.0669	0.9033	0.0345	0.5547	0.0395	0.1955
Welsch	0.3814	6.5061	0.2197	2.9084	0.1107	1.3426	0.0773	0.8859	0.0424	0.5500	0.0409	0.1953
MM	-0.7469	17.9180	-0.1470	4.1513	-0.0261	1.5745	-0.0355	1.0454	-0.0177	0.6043	0.0125	0.2166
Theil	-0.9082	9.7817	-0.4366	3.3876	-0.2287	1.4094	-0.1660	0.9115	-0.1103	0.5476	-0.0155	0.1891
LMS	-1.0801	23.5619	-0.4341	6.9187	-0.2004	2.5985	-0.1666	1.8064	-0.1248	1.0379	-0.0287	0.3605
LTS	-0.3993	10.1447	-0.1460	3.8884	-0.0620	1.6708	-0.0646	1.1590	-0.0480	0.6654	0.0128	0.2367
LAD	0.3320	6.8826	0.2558	2.8488	0.1580	1.3191	0.1076	0.8489	0.0580	0.5371	0.0465	0.1940
RepMed	-1.3191	15.3269	-0.6226	4.4693	-0.3279	1.7520	-0.2369	1.1304	-0.1537	0.6388	-0.0305	0.2184
Quan2575	-1.7239	26.2472	-0.6947	7.0958	-0.3497	2.9585	-0.2563	1.8938	-0.1842	1.1643	-0.0962	0.4191
Quan1090	-0.6757	8.1596	-0.3897	3.7207	-0.2166	1.8395	-0.1524	1.2258	-0.1131	0.7647	-0.0344	0.2943
Med/MAD	-2.3071	30.7297	-1.0254	8.7597	-0.5252	3.2656	-0.3545	1.9718	-0.2511	1.1434	-0.0332	0.3640
MTM10	-0.9877	12.5308	-0.3567	4.1246	-0.1593	1.7798	-0.1141	1.1968	-0.0829	0.7112	0.0135	0.2630
MTM15	-0.5184	14.4834	-0.4275	5.4189	-0.1444	2.2240	-0.0799	1.3783	-0.0339	0.8294	-0.0137	0.3453
Medians	-2.0177	26.7624	-0.9200	8.2305	-0.4382	3.0489	-0.2529	1.8429	-0.1462	1.0205	0.0714	0.3694



Şekil 6.5. Weib ( $\gamma = 1$ ,  $\alpha = 8$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.1-6.5'teki tablolar ve Şekil 6.1-6.5'deki grafikler incelendiğinde aşağıdaki sonuçlar çıkarılabilir:

- (i)  $n$  örneklem boyutu arttıkça tüm tahmincilerin MSE ve Bias değerleri genellikle azalmaktadır.
- (ii)  $n = 10$  ve  $n = 20$  için M-tahminciler ve LAD genellikle ML ve LS tahmincilerinden daha düşük MSE değerleri ortaya çıkarmaktadırlar.
- (iii)  $n = 40, 60, 100$  ve  $250$  için M-tahminciler, Theil ve LAD tahmincileri ML tahmincisinden sonra MSE'ye göre en iyi performansı sergilemektedir.
- (iv) MM tahmincisinin Bias değeri  $n = 10$  durumu dışında tüm tahminciler arasında en küçüktür.
- (v)  $n = 10$  için robust tahminciler arasında Med/MAD'ın performansı en kötüdür.

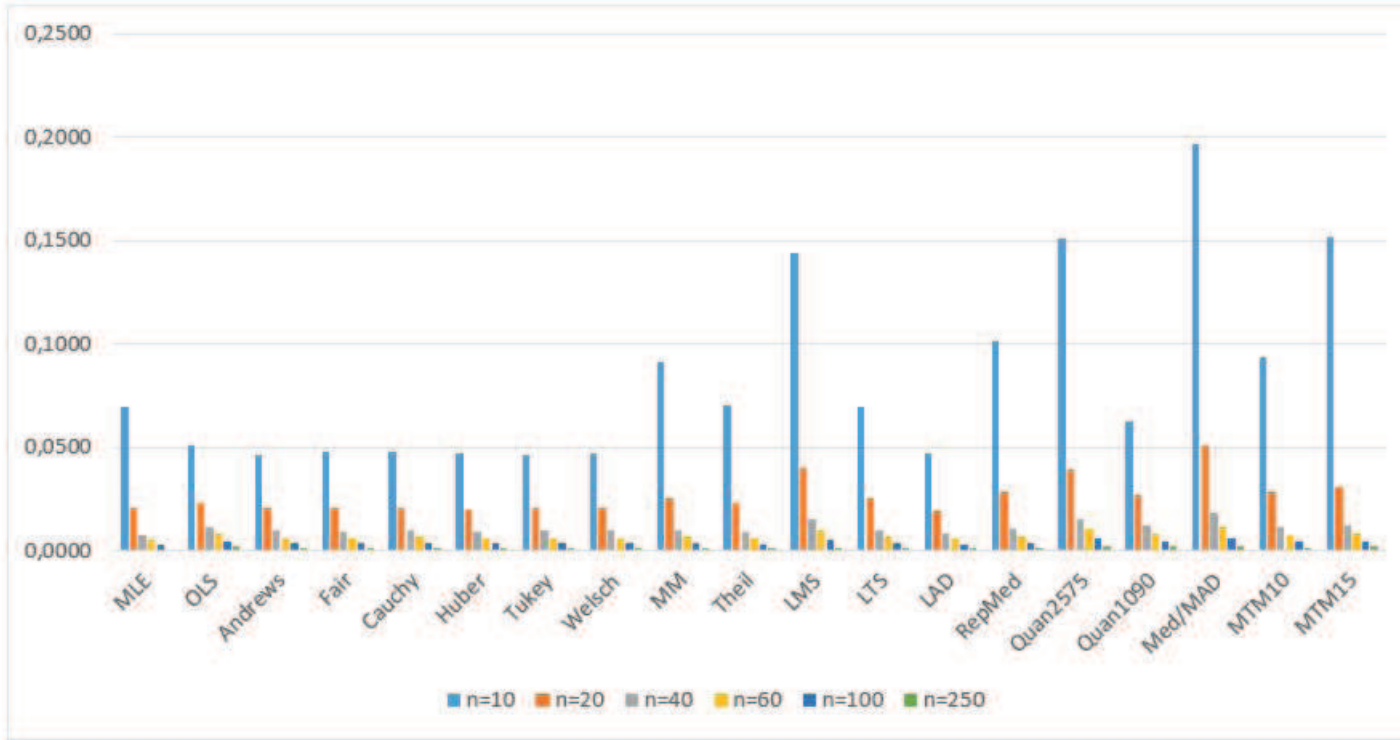
### 6.1.2. Pareto Dağılımı

Pareto dağılımında da temiz veri için, Weibull dağılımında olduğu gibi örneklem boyutu  $n = 10, 20, 40, 60, 100$  ve  $250$ , şekil parametresi  $\alpha = 0.5, 1, 2, 4$  ve  $8$  ve ölçek parametresi  $\sigma = 1$  şeklinde alınarak incelenmiştir. Temiz veriler ile elde edilen tahmincilerin MSE ve Bias değerleri Çizelge 6.6-6.10'da ve bu sonuçlara ilişkin grafikler Şekil 6.6-6.10'da verilmiştir.

Çizelge 6.6. Pareto ( $\sigma = 1$ ,  $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
ML	-0.1251	0.0700	-0.0548	0.0210	-0.0256	0.0079	-0.0167	0.0050	-0.0097	0.0026	-0.0037	0.0010
LS	-0.0072	0.0508	0.0125	0.0231	0.0166	0.0115	0.0158	0.0080	0.0142	0.0048	0.0096	0.0020
Andrews	0.0124	0.0466	0.0134	0.0205	0.0049	0.0096	0.0036	0.0064	0.0029	0.0037	0.0013	0.0014
Fair	0.0021	0.0477	0.0171	0.0206	0.0137	0.0095	0.0105	0.0063	0.0077	0.0037	0.0037	0.0015
Cauchy	0.0008	0.0481	0.0176	0.0211	0.0147	0.0098	0.0116	0.0065	0.0089	0.0038	0.0041	0.0015
Huber	0.0043	0.0473	0.0166	0.0203	0.0120	0.0093	0.0091	0.0062	0.0068	0.0036	0.0034	0.0015
Tukey	0.0122	0.0466	0.0134	0.0205	0.0049	0.0096	0.0037	0.0064	0.0029	0.0037	0.0015	0.0015
Welsch	0.0101	0.0468	0.0121	0.0205	0.0048	0.0096	0.0036	0.0064	0.0028	0.0037	0.0016	0.0015
MM	-0.0781	0.0915	-0.0262	0.0253	-0.0098	0.0103	-0.0078	0.0069	-0.0042	0.0039	-0.0024	0.0015
Theil	-0.0896	0.0705	-0.0415	0.0228	-0.0205	0.0092	-0.0147	0.0060	-0.0086	0.0033	-0.0037	0.0012
LMS	-0.1032	0.1437	-0.0417	0.0402	-0.0181	0.0153	-0.0134	0.0096	-0.0077	0.0052	-0.0049	0.0018
LTS	-0.0469	0.0700	-0.0175	0.0251	-0.0067	0.0102	-0.0053	0.0070	-0.0026	0.0040	-0.0017	0.0015
LAD	0.0109	0.0473	0.0128	0.0189	0.0090	0.0087	0.0063	0.0057	0.0046	0.0033	0.0023	0.0013
RepMed	-0.1186	0.1015	-0.0536	0.0284	-0.0260	0.0107	-0.0186	0.0069	-0.0102	0.0036	-0.0049	0.0013
Quan2575	-0.1387	0.1512	-0.0447	0.0391	-0.0228	0.0157	-0.0163	0.0107	-0.0095	0.0059	-0.0045	0.0023
Quan1090	-0.0222	0.0623	-0.0199	0.0273	-0.0147	0.0125	-0.0124	0.0087	-0.0070	0.0049	-0.0015	0.0021
Med/MAD	-0.1901	0.1970	-0.0813	0.0509	-0.0378	0.0182	-0.0254	0.0115	-0.0154	0.0059	-0.0067	0.0020
MTM10	-0.0970	0.0933	-0.0334	0.0282	-0.0137	0.0118	-0.0088	0.0078	-0.0044	0.0045	-0.0012	0.0018
MTM15	-0.1763	0.1519	-0.0373	0.0311	-0.0152	0.0121	-0.0107	0.0084	-0.0063	0.0049	-0.0061	0.0020

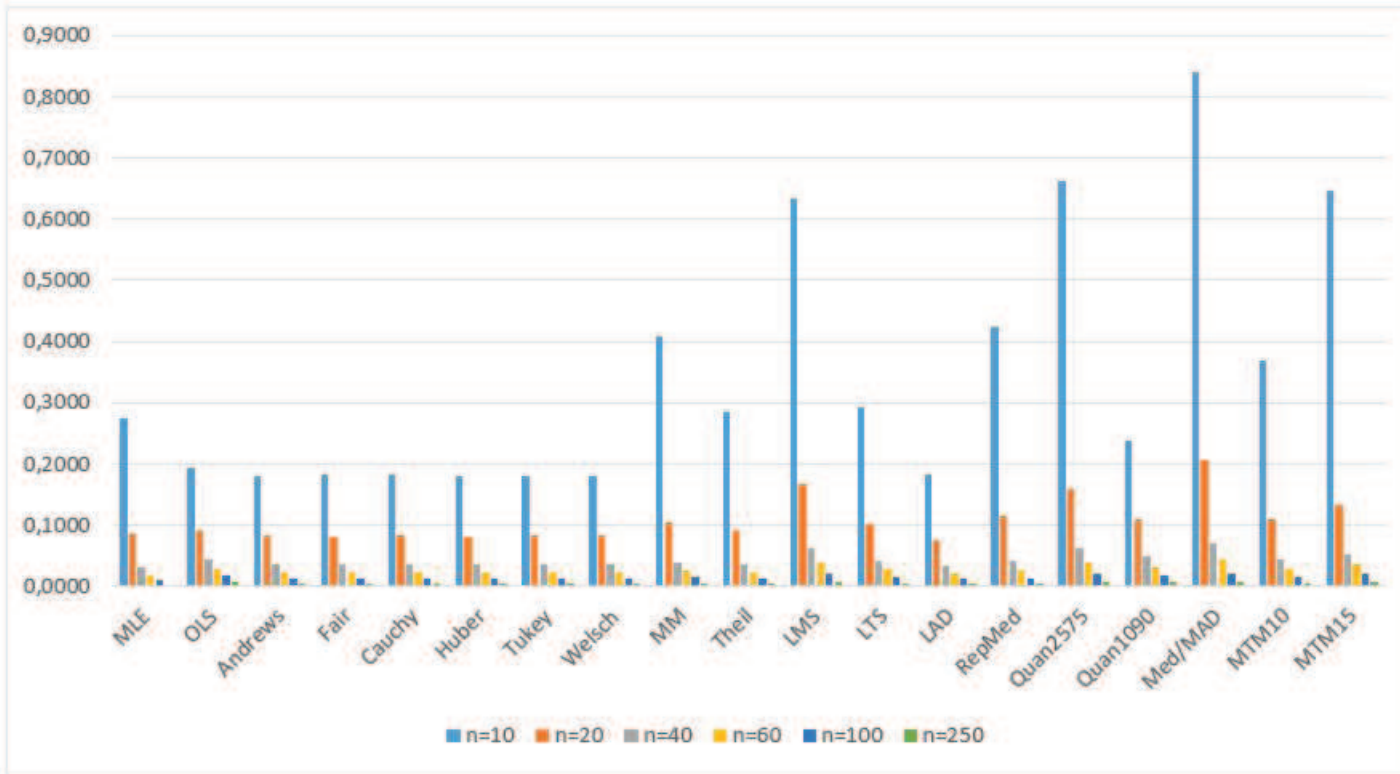




Şekil 6.6. Pareto ( $\sigma = 1$ ,  $\alpha = 0.5$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.7. Pareto ( $\sigma = 1$ ,  $\alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

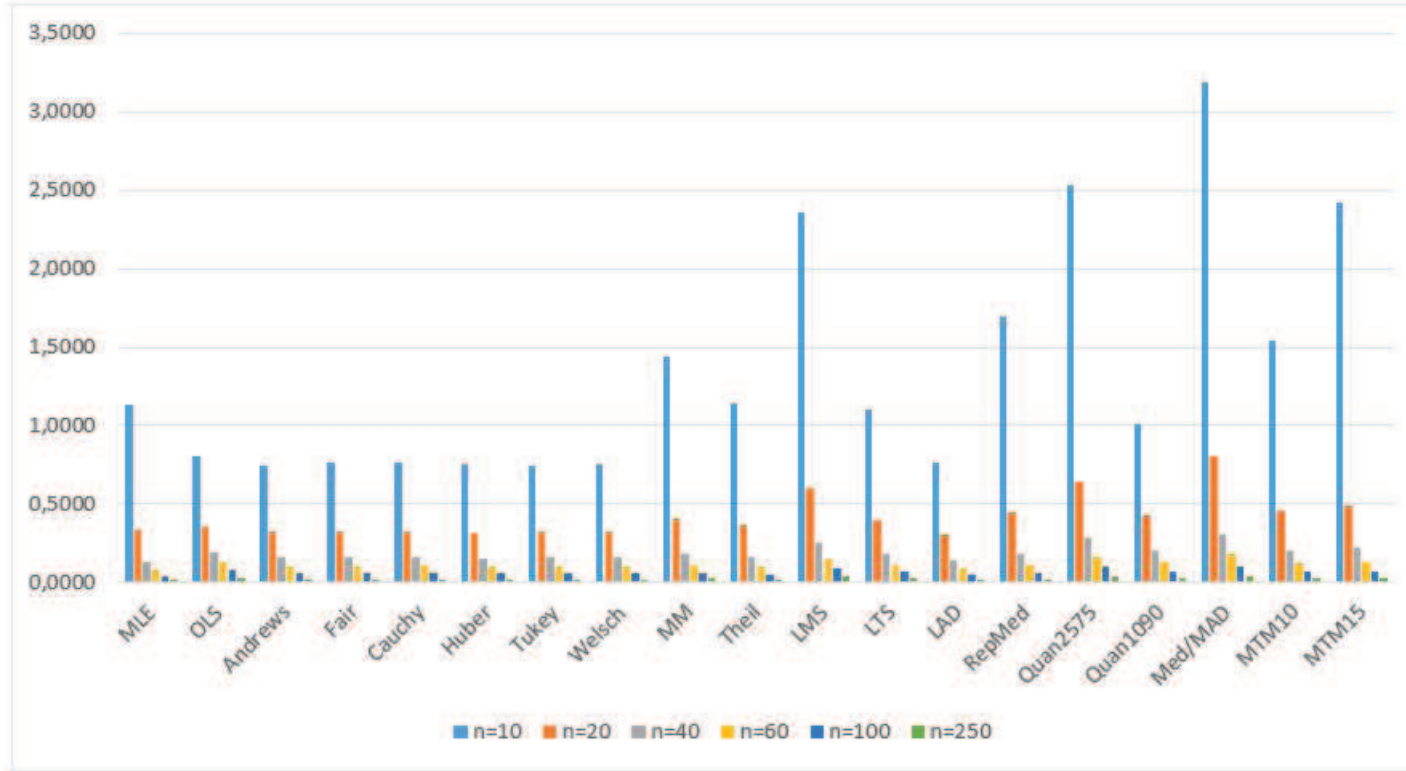
n	10		20		40		60		100		250	
Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
ML	-0.2494	0.2759	-0.1104	0.0858	-0.0514	0.0316	-0.0329	0.0189	-0.0183	0.0100	-0.0063	0.0035
LS	-0.0162	0.1926	0.0243	0.0910	0.0340	0.0443	0.0342	0.0292	0.0295	0.0183	0.0197	0.0079
Andrews	0.0233	0.1794	0.0246	0.0822	0.0096	0.0358	0.0056	0.0229	0.0046	0.0136	0.0028	0.0050
Fair	0.0023	0.1821	0.0330	0.0817	0.0276	0.0362	0.0225	0.0226	0.0155	0.0133	0.0078	0.0052
Cauchy	-0.0003	0.1835	0.0340	0.0835	0.0295	0.0375	0.0244	0.0234	0.0176	0.0136	0.0105	0.0054
Huber	0.0065	0.1813	0.0313	0.0805	0.0245	0.0352	0.0195	0.0221	0.0135	0.0130	0.0066	0.0050
Tukey	0.0231	0.1794	0.0242	0.0824	0.0096	0.0359	0.0056	0.0229	0.0046	0.0136	0.0028	0.0050
Welsch	0.0189	0.1793	0.0209	0.0824	0.0093	0.0361	0.0059	0.0228	0.0055	0.0136	0.0022	0.0051
MM	-0.1651	0.4080	-0.0567	0.1047	-0.0211	0.0394	-0.0126	0.0261	-0.0073	0.0140	-0.0025	0.0051
Theil	-0.1813	0.2847	-0.0849	0.0913	-0.0418	0.0354	-0.0280	0.0220	-0.0170	0.0119	-0.0064	0.0041
LMS	-0.2103	0.6343	-0.0868	0.1683	-0.0387	0.0615	-0.0268	0.0385	-0.0127	0.0195	-0.0040	0.0071
LTS	-0.0951	0.2931	-0.0380	0.1022	-0.0133	0.0402	-0.0076	0.0271	-0.0037	0.0143	-0.0008	0.0049
LAD	0.0187	0.1819	0.0240	0.0759	0.0172	0.0329	0.0130	0.0205	0.0091	0.0120	0.0045	0.0045
RepMed	-0.2442	0.4241	-0.1087	0.1139	-0.0520	0.0412	-0.0358	0.0256	-0.0210	0.0132	-0.0076	0.0043
Quan2575	-0.2856	0.6612	-0.1007	0.1594	-0.0498	0.0630	-0.0335	0.0400	-0.0212	0.0213	-0.0078	0.0073
Quan1090	-0.0734	0.2373	-0.0565	0.1086	-0.0360	0.0501	-0.0281	0.0321	-0.0174	0.0183	-0.0092	0.0078
Med/MAD	-0.3887	0.8412	-0.1611	0.2063	-0.0731	0.0715	-0.0488	0.0435	-0.0265	0.0213	-0.0091	0.0074
MTM10	-0.1925	0.3697	-0.0684	0.1087	-0.0283	0.0441	-0.0198	0.0281	-0.0117	0.0159	-0.0050	0.0059
MTM15	-0.3546	0.6472	-0.0821	0.1335	-0.0379	0.0532	-0.0247	0.0355	-0.0161	0.0200	-0.0087	0.0069



Şekil 6.7. Pareto ( $\sigma = 1, \alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.8. Pareto ( $\sigma = 1$ ,  $\alpha = 2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

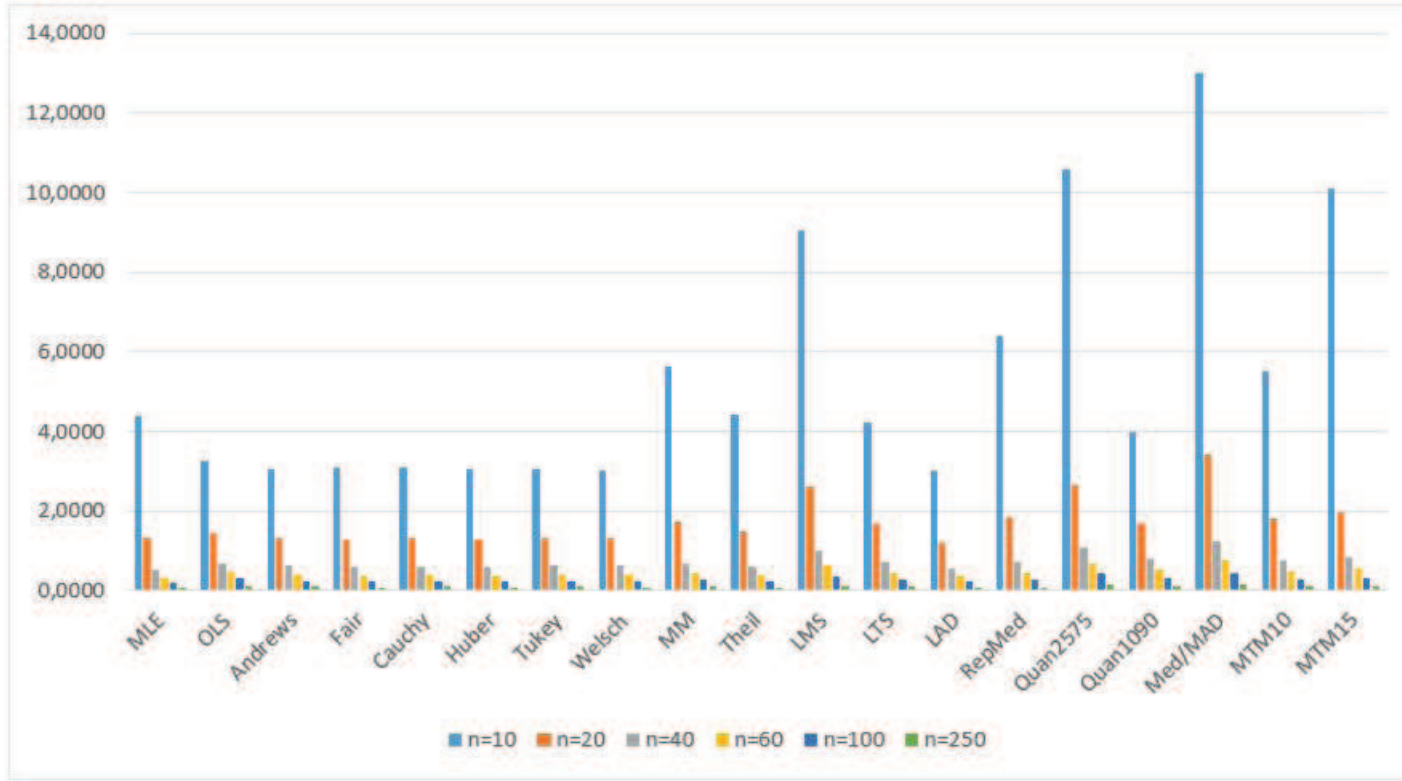
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.4926	1.1276	-0.2188	0.3315	-0.1027	0.1341	-0.0670	0.0819	-0.0380	0.0436	-0.0141	0.0170
LS	-0.0308	0.8086	0.0494	0.3562	0.0618	0.1937	0.0594	0.1331	0.0574	0.0762	0.0390	0.0323
Andrews	0.0493	0.7468	0.0482	0.3267	0.0140	0.1617	0.0126	0.1042	0.0104	0.0594	0.0098	0.0230
Fair	0.0077	0.7609	0.0668	0.3213	0.0510	0.1596	0.0419	0.1026	0.0317	0.0577	0.0183	0.0228
Cauchy	0.0017	0.7686	0.0688	0.3281	0.0554	0.1651	0.0461	0.1072	0.0363	0.0592	0.0232	0.0233
Huber	0.0155	0.7577	0.0628	0.3165	0.0435	0.1557	0.0363	0.1004	0.0277	0.0567	0.0170	0.0224
Tukey	0.0487	0.7473	0.0466	0.3278	0.0138	0.1614	0.0127	0.1041	0.0104	0.0594	0.0098	0.0230
Welsch	0.0397	0.7515	0.0398	0.3275	0.0153	0.1612	0.0135	0.1045	0.0125	0.0596	0.0103	0.0230
MM	-0.3029	1.4368	-0.1104	0.4082	-0.0457	0.1775	-0.0246	0.1120	-0.0205	0.0643	-0.0072	0.0252
Theil	-0.3540	1.1395	-0.1685	0.3620	-0.0856	0.1568	-0.0575	0.0956	-0.0369	0.0528	-0.0143	0.0204
LMS	-0.4033	2.3588	-0.1626	0.6049	-0.0775	0.2514	-0.0516	0.1477	-0.0368	0.0913	-0.0159	0.0349
LTS	-0.1684	1.1012	-0.0720	0.3929	-0.0308	0.1800	-0.0209	0.1121	-0.0198	0.0664	-0.0089	0.0265
LAD	0.0436	0.7607	0.0480	0.2998	0.0323	0.1445	0.0271	0.0932	0.0164	0.0531	0.0109	0.0208
RepMed	-0.4771	1.6943	-0.2167	0.4466	-0.1082	0.1805	-0.0742	0.1091	-0.0474	0.0609	-0.0201	0.0229
Quan2575	-0.5400	2.5306	-0.2080	0.6434	-0.1112	0.2844	-0.0735	0.1653	-0.0466	0.1003	-0.0285	0.0398
Quan1090	-0.1918	1.0058	-0.1275	0.4285	-0.0746	0.2061	-0.0445	0.1296	-0.0218	0.0722	-0.0021	0.0301
Med/MAD	-0.7527	3.1883	-0.3286	0.8065	-0.1628	0.3081	-0.1140	0.1867	-0.0701	0.1041	-0.0310	0.0419
MTM10	-0.3780	1.5421	-0.1381	0.4589	-0.0577	0.1995	-0.0340	0.1230	-0.0178	0.0698	-0.0039	0.0259
MTM15	-0.7029	2.4223	-0.1362	0.4910	-0.0639	0.2184	-0.0469	0.1339	-0.0275	0.0743	-0.0139	0.0310



Şekil 6.8. Pareto ( $\sigma = 1$ ,  $\alpha = 2$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.9. Pareto ( $\sigma = 1, \alpha = 4$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
ML	-0.9545	4.4005	-0.4442	1.3382	-0.2152	0.5174	-0.1425	0.3115	-0.0901	0.1853	-0.0294	0.0619
LS	-0.0442	3.2631	0.1045	1.4367	0.1362	0.6979	0.1335	0.4639	0.1074	0.2974	0.0818	0.1287
Andrews	0.1124	3.0365	0.0928	1.3258	0.0179	0.6213	0.0143	0.3956	0.0094	0.2422	0.0184	0.0937
Fair	0.0307	3.0782	0.1360	1.2899	0.1057	0.5869	0.0859	0.3727	0.0535	0.2307	0.0372	0.0911
Cauchy	0.0196	3.1052	0.1412	1.3185	0.1155	0.6040	0.0954	0.3836	0.0638	0.2359	0.0464	0.0936
Huber	0.0476	3.0608	0.1279	1.2669	0.0897	0.5796	0.0734	0.3673	0.0450	0.2278	0.0331	0.0897
Tukey	0.1112	3.0355	0.0900	1.3339	0.0175	0.6215	0.0146	0.3957	0.0091	0.2421	0.0184	0.0937
Welsch	0.0965	3.0324	0.0827	1.3211	0.0195	0.6249	0.0158	0.3931	0.0103	0.2410	0.0171	0.0935
MM	-0.5805	5.6416	-0.2339	1.7147	-0.1166	0.6973	-0.0736	0.4368	-0.0485	0.2755	-0.0053	0.0975
Theil	-0.6636	4.4085	-0.3494	1.4971	-0.1828	0.6007	-0.1304	0.3780	-0.0855	0.2281	-0.0258	0.0772
LMS	-0.7932	9.0495	-0.3497	2.6055	-0.1793	0.9835	-0.1311	0.6247	-0.0937	0.3650	-0.0316	0.1229
LTS	-0.3076	4.2378	-0.1662	1.6691	-0.0911	0.7125	-0.0630	0.4502	-0.0474	0.2758	-0.0062	0.0981
LAD	0.0966	3.0328	0.0894	1.2035	0.0597	0.5487	0.0441	0.3454	0.0240	0.2188	0.0208	0.0830
RepMed	-0.9163	6.4062	-0.4479	1.8565	-0.2359	0.7169	-0.1694	0.4465	-0.1086	0.2604	-0.0364	0.0858
Quan2575	-1.0700	10.5794	-0.4562	2.6707	-0.2506	1.0901	-0.1687	0.6818	-0.1240	0.4288	-0.0221	0.1398
Quan1090	-0.4076	3.9661	-0.2669	1.7043	-0.1557	0.7994	-0.1102	0.5081	-0.0808	0.3153	-0.0175	0.1146
Med/MAD	-1.4753	12.9942	-0.6841	3.4092	-0.3537	1.2385	-0.2454	0.7504	-0.1761	0.4402	-0.0532	0.1450
MTM10	-0.7072	5.5122	-0.2888	1.8258	-0.1288	0.7592	-0.0882	0.4874	-0.0602	0.2921	-0.0065	0.1057
MTM15	-1.4395	10.1099	-0.2972	1.9568	-0.1177	0.8510	-0.0702	0.5738	-0.0556	0.3101	-0.0381	0.1233

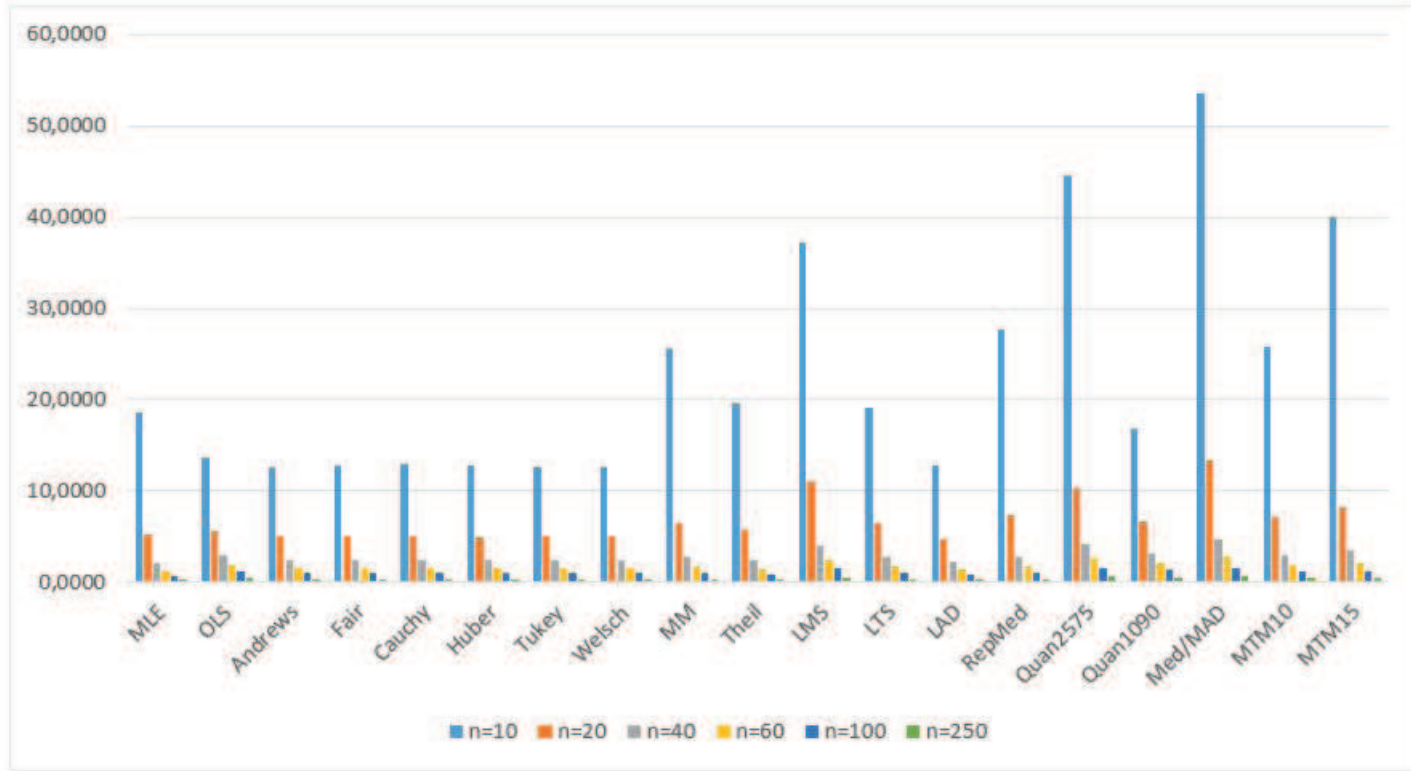


Şekil 6.9. Pareto ( $\sigma = 1$ ,  $\alpha = 4$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.10. Pareto ( $\sigma = 1$ ,  $\alpha = 8$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-2.0237	18.5533	-0.8850	5.2738	-0.4252	2.0915	-0.2783	1.2348	-0.1675	0.7370	-0.0771	0.2521
LS	-0.1800	13.7301	0.1498	5.5611	0.1951	2.8827	0.1855	1.9817	0.1502	1.2531	0.1450	0.4948
Andrews	0.1421	12.6034	0.1698	5.0298	0.0491	2.4117	0.0429	1.5717	0.0035	0.9933	0.0092	0.3528
Fair	-0.0262	12.8800	0.2269	4.9883	0.1616	2.3895	0.1256	1.5797	0.0745	0.9536	0.0543	0.3501
Cauchy	-0.0499	12.9930	0.2335	5.0869	0.1769	2.4739	0.1350	1.6250	0.0842	0.9732	0.0705	0.3631
Huber	0.0061	12.7835	0.2177	4.8961	0.1382	2.3417	0.1116	1.5531	0.0618	0.9387	0.0457	0.3453
Tukey	0.1398	12.6100	0.1654	5.0343	0.0496	2.4116	0.0435	1.5724	0.0037	0.9934	0.0092	0.3528
Welsch	0.1043	12.6118	0.1452	5.0221	0.0427	2.4138	0.0419	1.5678	0.0077	0.9788	0.0073	0.3550
MM	-1.2993	25.5956	-0.4470	6.5352	-0.1811	2.7065	-0.1107	1.6866	-0.0594	1.0309	-0.0474	0.3676
Theil	-1.4663	19.6118	-0.6832	5.8032	-0.3462	2.3759	-0.2298	1.4401	-0.1517	0.8905	-0.0789	0.3176
LMS	-1.6538	37.1428	-0.7155	11.0948	-0.2803	3.9406	-0.1660	2.3924	-0.1376	1.6030	-0.1194	0.5806
LTS	-0.7286	19.1794	-0.3026	6.4587	-0.1244	2.7587	-0.0882	1.7172	-0.0589	1.0972	-0.0592	0.3977
LAD	0.1183	12.7504	0.1719	4.7012	0.1054	2.2047	0.0843	1.4257	0.0397	0.8879	0.0168	0.3233
RepMed	-1.9401	27.7050	-0.8900	7.3923	-0.4322	2.7637	-0.2813	1.6479	-0.1798	1.0021	-0.1103	0.3616
Quan2575	-2.2887	44.6668	-0.8890	10.3438	-0.4433	4.2330	-0.3143	2.6745	-0.2244	1.6321	-0.1322	0.6706
Quan1090	-0.9928	16.9238	-0.5478	6.6654	-0.3045	3.1667	-0.1922	2.0334	-0.1473	1.2981	-0.0564	0.4401
Med/MAD	-3.0500	53.5848	-1.3259	13.3819	-0.6232	4.7469	-0.4058	2.7171	-0.2521	1.6007	-0.2081	0.6586
MTM10	-1.5823	25.8414	-0.5562	7.1233	-0.2476	3.0331	-0.1550	1.9011	-0.1058	1.1788	-0.0518	0.4316
MTM15	-2.8977	39.9465	-0.6126	8.2892	-0.2227	3.4085	-0.1647	2.1316	-0.0949	1.2297	-0.0911	0.4759





Şekil 6.10. Pareto ( $\sigma = 1$ ,  $\alpha = 8$ ) dağılımı için tahmincilerin MSE grafiği

Pareto dağılımı için oluşturulan Çizelge 6.6-6.10 ve Şekil 6.6-6.10 birlikte yorumlanırsa aşağıdaki sonuçlar çıkartılabilir:

- (i) Örneklem boyutu  $n$  arttıkça tüm tahmincilerin MSE değeri azalmaktadır.
- (ii) Örneklem boyutu  $n$  arttıkça tüm tahmincilerin yanlışlık değeri genellikle azalmaktadır.
- (iii) Weibull dağılımında olduğu gibi,  $n = 10$  ve  $20$  iken M-tahmincileri (Andrews, Fair, Cauchy, Huber, Tukey ve Welsch) ve LAD tahmincisi ML tahmincisinden daha küçük MSE değeri ortaya çıkarmışlardır.
- (iv)  $n = 40, 60, 100$  ve  $250$  için Theil ve LAD tahmincileri, ML tahmincisinden sonra MSE'ye göre en iyidir.
- (v) Med/MAD,  $n = 10$  olduğunda en yüksek MSE değerlerine sahiptir.

## 6.2. Aykırı Değer Bulunduğu Durumda Tahmincilerin Performanslarının İncelenmesi

### 6.2.1. Weibull Dağılımı

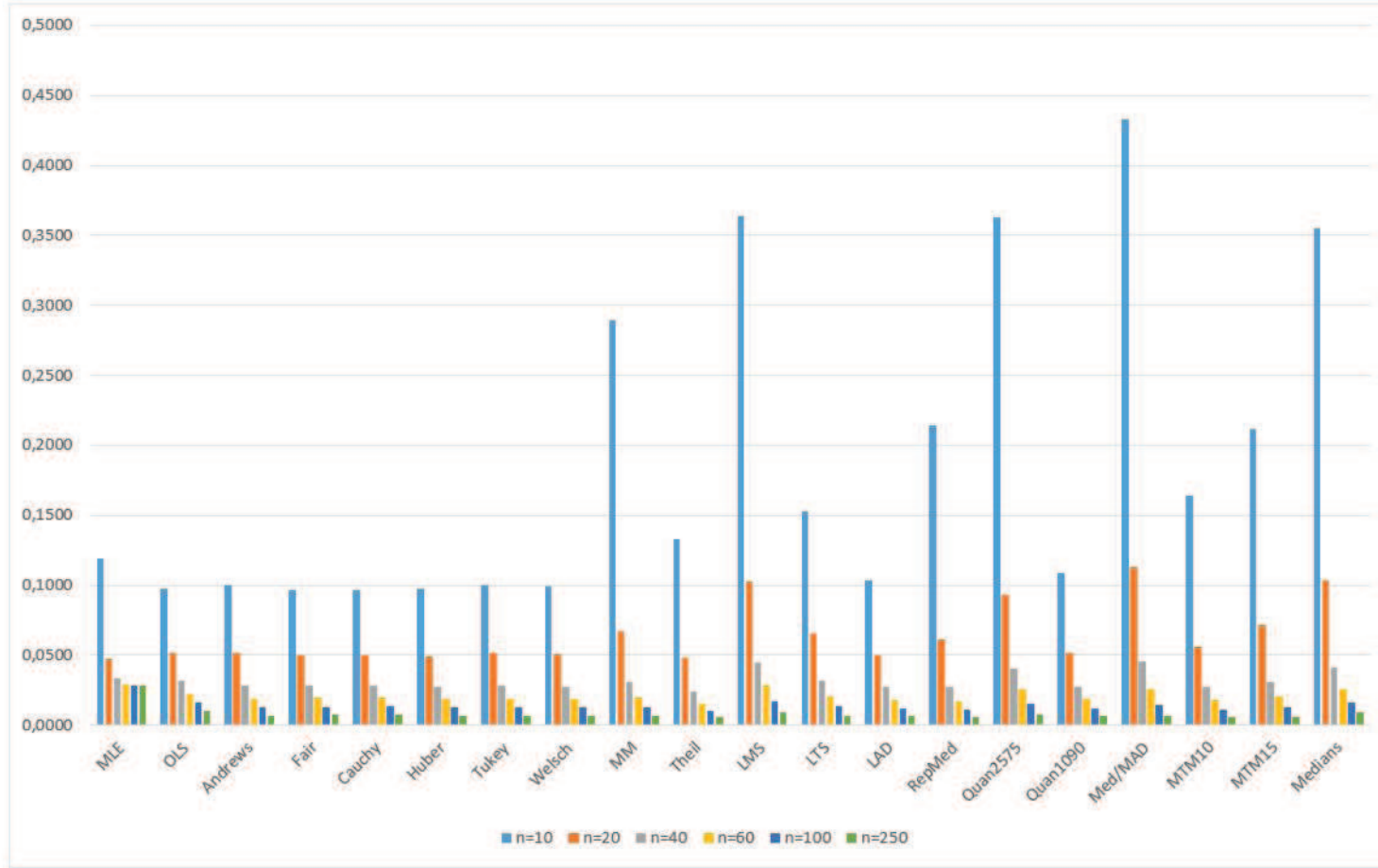
Aykırı değerlerin bulunduğu durumlarda tahmincilerin performanlarının incelenmesi amacıyla, farklı simülasyon kurgularıyla aykırı değerler içeren veriler üretilmiştir. Weibull dağılımı için,

- (i) Weib ( $\gamma = 1, \alpha = 1$ )' den üretilen veri %10 oranında Weib ( $\gamma = 1, \alpha = 0.2$ )'den aykırı değer atılarak,
- (ii) Weib ( $\gamma = 1, \alpha = 1$ )' den üretilen veri %10 oranında Weib ( $\gamma = 0.2, \alpha = 1$ )'den aykırı değer atılarak,
- (iii) Weib ( $\gamma = 1, \alpha = 1$ )' den üretilen veri %10 oranında  $U(0,20)$ 'den aykırı değer atılarak ,
- (iv) Weib ( $\gamma = 1, \alpha = 1$ )' den üretilen veri %10 oranında 100 ile aykırı değer atılarak,

oluřturulan simulasyon kurguları ile aykırı deęer ieren veri seti iin tahminci-lerin performansları deęerlendirilmiřtir. Elde edilen sonular izelgelerle ve bu sonular ile oluřturulan grafikler de řekiller olarak verilmiřtir.

**Çizelge 6.11.** 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 0.2, \alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

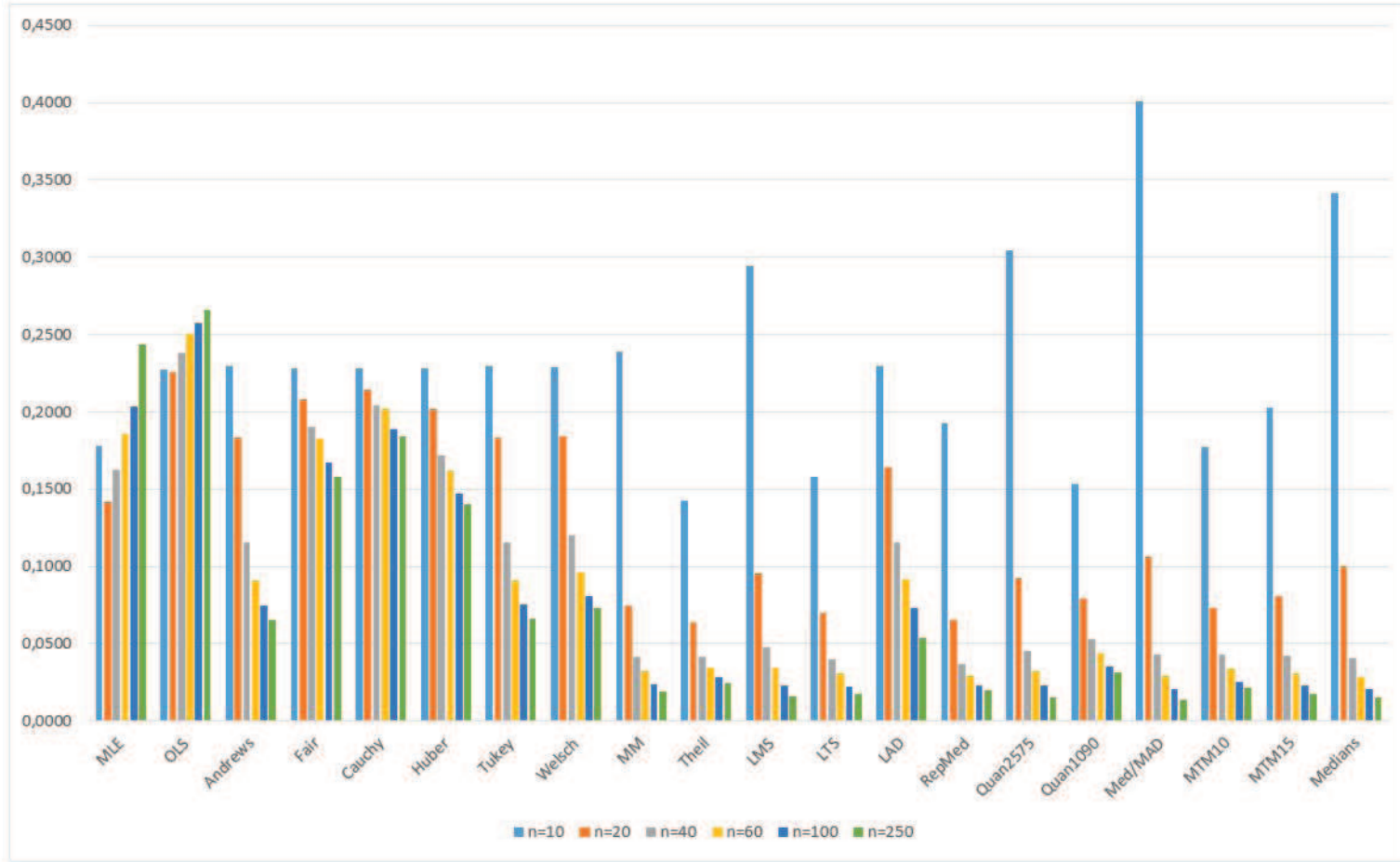
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	-0.0467	0.1193	0.0666	0.0476	0.1142	0.0334	0.1275	0.0289	0.1490	0.0282	0.1591	0.0280
LS	0.1020	0.0972	0.1149	0.0513	0.1059	0.0317	0.0957	0.0227	0.0925	0.0161	0.0819	0.0100
Andrews	0.1163	0.0999	0.0993	0.0518	0.0801	0.0281	0.0720	0.0190	0.0674	0.0127	0.0563	0.0067
Fair	0.1111	0.0970	0.1128	0.0496	0.0950	0.0283	0.0839	0.0195	0.0790	0.0132	0.0664	0.0075
Cauchy	0.1110	0.0968	0.1146	0.0502	0.0968	0.0287	0.0854	0.0198	0.0807	0.0135	0.0685	0.0077
Huber	0.1132	0.0973	0.1107	0.0492	0.0921	0.0278	0.0815	0.0192	0.0766	0.0130	0.0641	0.0072
Tukey	0.1160	0.0999	0.0991	0.0517	0.0802	0.0280	0.0720	0.0190	0.0675	0.0127	0.0563	0.0067
Welsch	0.1143	0.0991	0.1000	0.0508	0.0819	0.0277	0.0730	0.0189	0.0685	0.0127	0.0568	0.0067
MM	-0.0318	0.2892	0.0507	0.0673	0.0585	0.0307	0.0588	0.0195	0.0566	0.0129	0.0510	0.0064
Theil	-0.0406	0.1328	0.0201	0.0485	0.0412	0.0240	0.0466	0.0157	0.0537	0.0104	0.0548	0.0062
LMS	-0.0686	0.3635	0.0190	0.1025	0.0375	0.0447	0.0456	0.0286	0.0445	0.0168	0.0463	0.0090
LTS	0.0202	0.1525	0.0484	0.0655	0.0563	0.0317	0.0541	0.0205	0.0533	0.0134	0.0481	0.0067
LAD	0.1100	0.1039	0.1032	0.0501	0.0844	0.0271	0.0751	0.0184	0.0702	0.0123	0.0583	0.0066
RepMed	-0.0968	0.2142	-0.0071	0.0612	0.0249	0.0274	0.0352	0.0172	0.0432	0.0110	0.0470	0.0061
Quan2575	-0.1506	0.3628	-0.0364	0.0936	0.0063	0.0405	0.0156	0.0254	0.0264	0.0158	0.0360	0.0074
Quan1090	-0.0368	0.1084	0.0110	0.0518	0.0297	0.0278	0.0359	0.0185	0.0467	0.0118	0.0514	0.0067
Med/MAD	-0.2203	0.4327	-0.0635	0.1134	-0.0072	0.0454	0.0123	0.0259	0.0257	0.0145	0.0343	0.0070
MTM10	-0.0646	0.1644	0.0108	0.0563	0.0295	0.0275	0.0309	0.0176	0.0384	0.0109	0.0390	0.0055
MTM15	-0.0193	0.2115	-0.0033	0.0715	0.0271	0.0308	0.0304	0.0205	0.0346	0.0124	0.0395	0.0060
Medians	-0.1597	0.3547	-0.0302	0.1035	0.0214	0.0417	0.0395	0.0255	0.0530	0.0165	0.0639	0.0091



Şekil 6.11.  $0.90 \text{ Weib}(\gamma = 1, \alpha = 1) + 0.10 \text{ Weib}(\gamma = 0.2, \alpha = 1)$  dağılımı için tahmincilerin MSE grafiği

**Çizelge 6.12.** 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	0.1174	0.1779	0.2493	0.1422	0.3388	0.1624	0.3913	0.1855	0.4253	0.2032	0.4842	0.2439
LS	0.2967	0.2273	0.3776	0.2255	0.4342	0.2382	0.4635	0.2505	0.4851	0.2573	0.5054	0.2656
Andrews	0.3050	0.2296	0.3041	0.1833	0.2569	0.1152	0.2440	0.0905	0.2383	0.0747	0.2424	0.0658
Fair	0.3032	0.2282	0.3567	0.2082	0.3734	0.1900	0.3818	0.1825	0.3788	0.1673	0.3828	0.1576
Cauchy	0.3031	0.2280	0.3645	0.2145	0.3896	0.2044	0.4036	0.2016	0.4047	0.1889	0.4158	0.1846
Huber	0.3041	0.2278	0.3472	0.2021	0.3510	0.1720	0.3566	0.1617	0.3536	0.1472	0.3609	0.1402
Tukey	0.3048	0.2295	0.3040	0.1832	0.2571	0.1152	0.2448	0.0909	0.2392	0.0752	0.2436	0.0664
Welsch	0.3037	0.2289	0.3067	0.1842	0.2659	0.1200	0.2546	0.0963	0.2497	0.0810	0.2559	0.0731
MM	0.0368	0.2387	0.1056	0.0747	0.1242	0.0418	0.1260	0.0321	0.1217	0.0240	0.1227	0.0192
Theil	0.0618	0.1424	0.1107	0.0635	0.1343	0.0413	0.1415	0.0344	0.1417	0.0283	0.1449	0.0246
LMS	-0.0092	0.2941	0.0563	0.0957	0.0907	0.0474	0.0926	0.0344	0.0959	0.0233	0.0982	0.0157
LTS	0.0880	0.1580	0.0994	0.0699	0.1157	0.0396	0.1169	0.0305	0.1136	0.0222	0.1153	0.0173
LAD	0.2935	0.2295	0.2977	0.1638	0.2737	0.1156	0.2570	0.0919	0.2392	0.0728	0.2162	0.0536
RepMed	-0.0067	0.1923	0.0687	0.0652	0.1054	0.0371	0.1145	0.0292	0.1191	0.0233	0.1251	0.0196
Quan2575	-0.0666	0.3048	0.0318	0.0927	0.0701	0.0454	0.0786	0.0319	0.0907	0.0228	0.1003	0.0156
Quan1090	0.1050	0.1532	0.1302	0.0792	0.1413	0.0529	0.1497	0.0436	0.1498	0.0350	0.1611	0.0313
Med/MAD	-0.1559	0.4008	-0.0154	0.1061	0.0482	0.0430	0.0672	0.0289	0.0769	0.0206	0.0901	0.0141
MTM10	0.0358	0.1776	0.1033	0.0729	0.1195	0.0429	0.1240	0.0337	0.1228	0.0255	0.1291	0.0214
MTM15	0.0917	0.2027	0.0744	0.0809	0.0983	0.0426	0.1033	0.0307	0.1077	0.0233	0.1151	0.0175
Medians	-0.1216	0.3419	-0.0042	0.0999	0.0626	0.0409	0.0738	0.0286	0.0850	0.0209	0.1000	0.0153

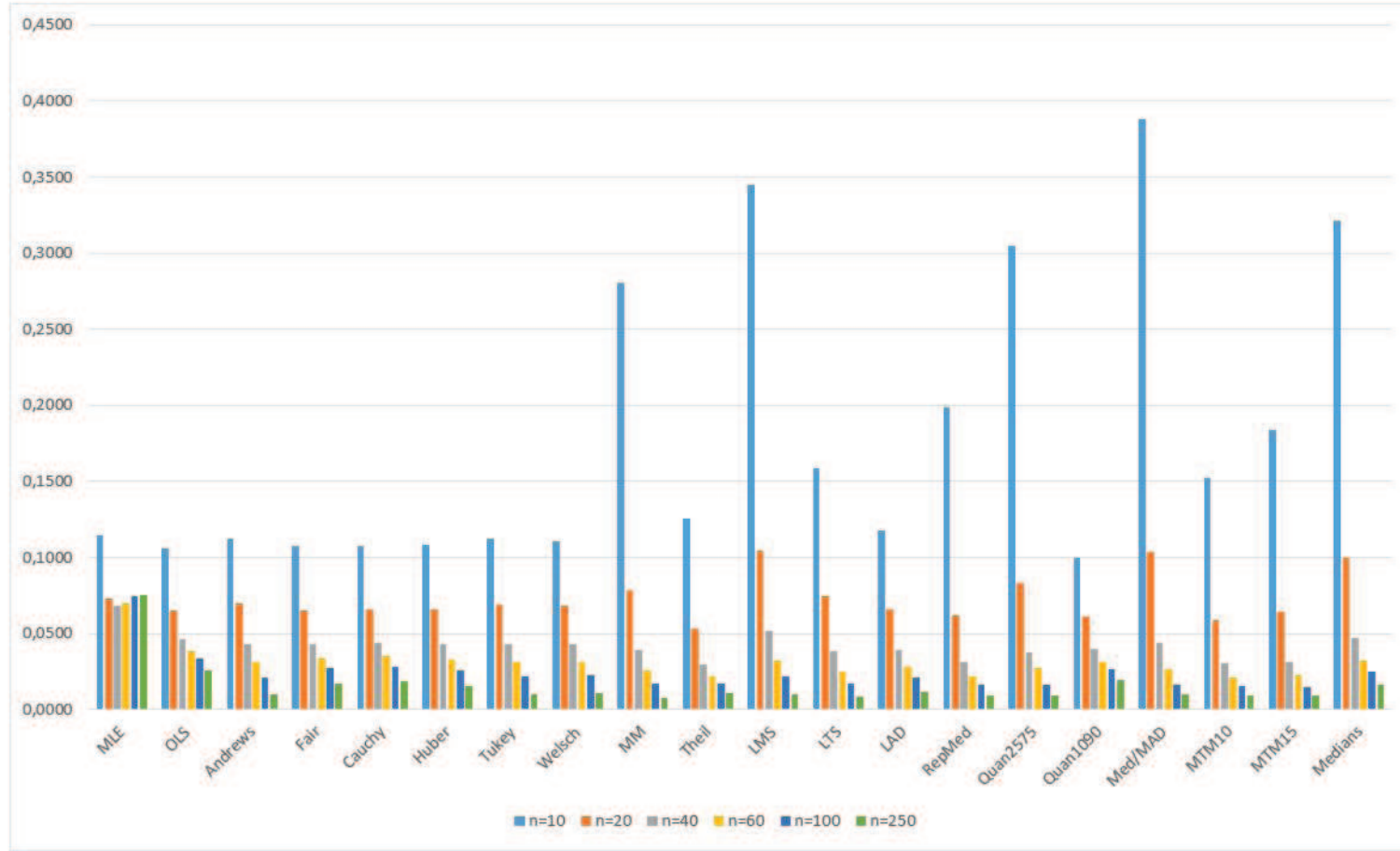


Şekil 6.12. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 Weib ( $\gamma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.13. 0.90 Weib ( $\gamma = 1, \alpha = 1$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE ve Bias değerleri

n Tahminciler	10		20		40		60		100		250	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
ML	0.0789	0.1147	0.1903	0.0727	0.2371	0.0684	0.2524	0.0700	0.2672	0.0744	0.2719	0.0750
LS	0.1791	0.1063	0.1871	0.0651	0.1770	0.0461	0.1699	0.0384	0.1674	0.0335	0.1541	0.0259
Andrews	0.1874	0.1121	0.1656	0.0695	0.1411	0.0430	0.1225	0.0311	0.1067	0.0214	0.0780	0.0102
Fair	0.1863	0.1078	0.1813	0.0651	0.1641	0.0434	0.1525	0.0340	0.1439	0.0273	0.1218	0.0175
Cauchy	0.1872	0.1077	0.1837	0.0657	0.1668	0.0442	0.1557	0.0349	0.1477	0.0283	0.1269	0.0187
Huber	0.1879	0.1083	0.1793	0.0656	0.1601	0.0430	0.1474	0.0331	0.1371	0.0259	0.1132	0.0156
Tukey	0.1869	0.1120	0.1655	0.0693	0.1414	0.0431	0.1230	0.0311	0.1072	0.0216	0.0784	0.0103
Welsch	0.1853	0.1105	0.1675	0.0683	0.1451	0.0430	0.1279	0.0316	0.1124	0.0225	0.0830	0.0109
MM	0.0259	0.2805	0.0973	0.0787	0.0961	0.0394	0.0873	0.0257	0.0779	0.0169	0.0615	0.0080
Theil	0.0262	0.1260	0.0759	0.0532	0.0934	0.0301	0.0943	0.0221	0.0975	0.0175	0.0877	0.0107
LMS	-0.0100	0.3447	0.0616	0.1043	0.0678	0.0518	0.0681	0.0323	0.0709	0.0218	0.0612	0.0104
LTS	0.0794	0.1585	0.0922	0.0749	0.0896	0.0385	0.0800	0.0252	0.0750	0.0170	0.0618	0.0082
LAD	0.1791	0.1177	0.1648	0.0657	0.1427	0.0394	0.1268	0.0284	0.1169	0.0212	0.0924	0.0116
RepMed	-0.0359	0.1984	0.0410	0.0623	0.0684	0.0312	0.0741	0.0216	0.0794	0.0162	0.0739	0.0090
Quan2575	-0.0930	0.3045	-0.0041	0.0835	0.0360	0.0374	0.0468	0.0273	0.0540	0.0163	0.0593	0.0090
Quan1090	0.0623	0.0994	0.0968	0.0609	0.1128	0.0400	0.1143	0.0315	0.1217	0.0265	0.1174	0.0192
Med/MAD	-0.1594	0.3878	-0.0219	0.1040	0.0321	0.0438	0.0482	0.0262	0.0577	0.0166	0.0661	0.0099
MTM10	-0.0043	0.1523	0.0603	0.0588	0.0759	0.0309	0.0740	0.0210	0.0793	0.0157	0.0743	0.0092
MTM15	0.0180	0.1835	0.0345	0.0644	0.0610	0.0315	0.0650	0.0227	0.0662	0.0147	0.0703	0.0091
Medians	-0.0779	0.3217	0.0296	0.1001	0.0765	0.0468	0.0915	0.0321	0.1021	0.0247	0.1057	0.0167

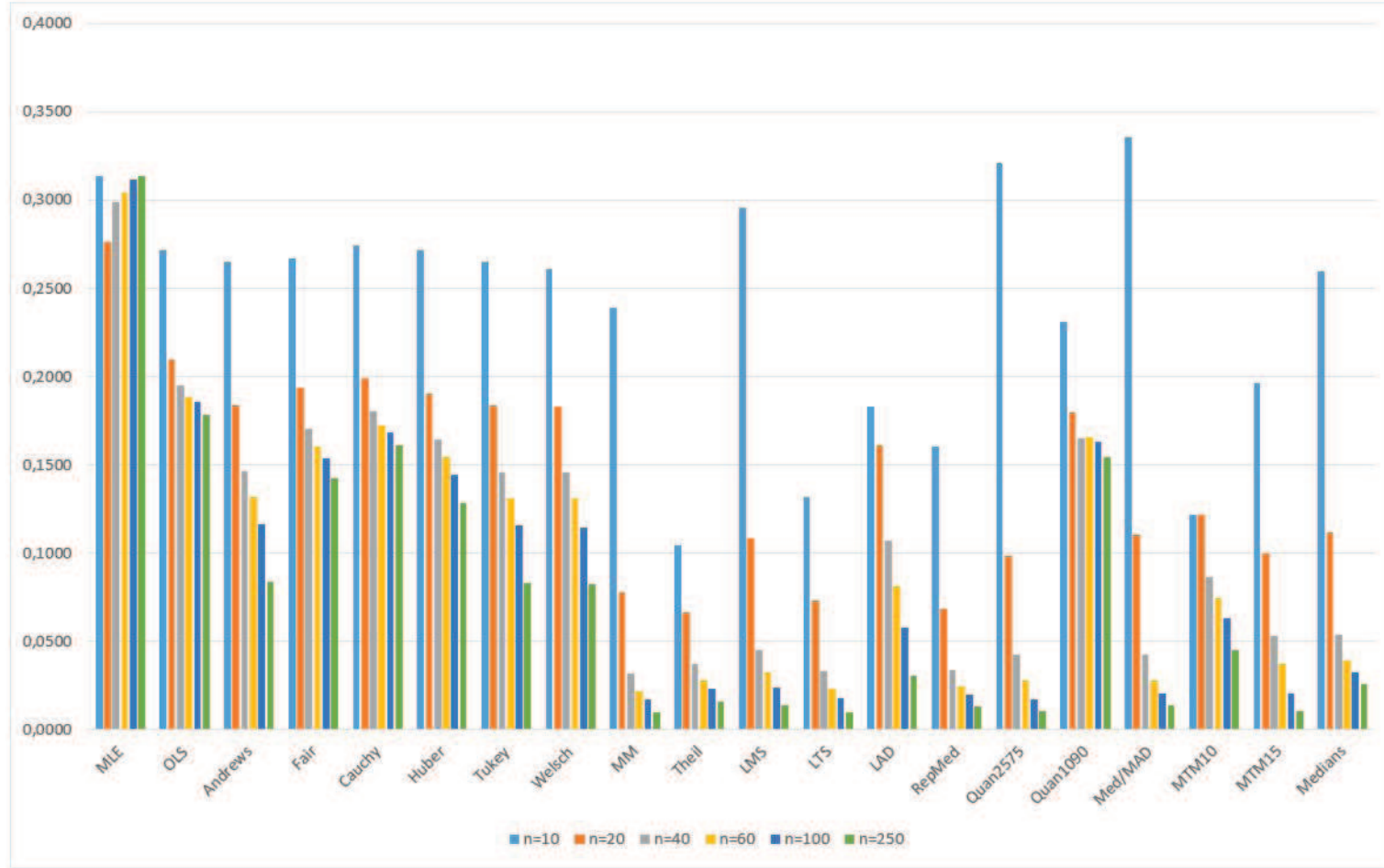




Şekil 6.13. 0.90 Weib ( $\gamma = 1$ ,  $\alpha = 1$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE grafiği

**Çizelge 6.14.** Weib ( $\gamma = 1$ ,  $\alpha = 1$ ) dağılımı için tahmincilerin MSE ve Bias değerleri (%10 100 değeri aykırı değer olarak atılmış)

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	0.5586	0.3134	0.4736	0.2761	0.5409	0.2991	0.5499	0.3045	0.5574	0.3113	0.5599	0.3138
LS	0.5190	0.2716	0.4192	0.2094	0.4304	0.1951	0.4269	0.1883	0.4277	0.1858	0.4208	0.1783
Andrews	0.5064	0.2653	0.3350	0.1840	0.3025	0.1461	0.2857	0.1320	0.2679	0.1168	0.2131	0.0836
Fair	0.5143	0.2671	0.3899	0.1939	0.3862	0.1702	0.3797	0.1606	0.3793	0.1538	0.3702	0.1422
Cauchy	0.5214	0.2741	0.3996	0.1990	0.4013	0.1801	0.3975	0.1724	0.4007	0.1687	0.3969	0.1612
Huber	0.5183	0.2716	0.3793	0.1905	0.3720	0.1642	0.3650	0.1542	0.3615	0.1448	0.3463	0.1282
Tukey	0.5059	0.2648	0.3344	0.1836	0.3021	0.1459	0.2844	0.1314	0.2665	0.1158	0.2121	0.0831
Welsch	0.5026	0.2612	0.3359	0.1831	0.3042	0.1455	0.2891	0.1314	0.2701	0.1145	0.2188	0.0823
MM	0.0075	0.2391	0.0670	0.0776	0.0782	0.0318	0.0770	0.0220	0.0821	0.0171	0.0737	0.0098
Theil	0.1005	0.1046	0.1166	0.0664	0.1210	0.0371	0.1170	0.0278	0.1198	0.0232	0.1104	0.0156
LMS	-0.0295	0.2956	0.0340	0.1082	0.0657	0.0451	0.0675	0.0326	0.0751	0.0240	0.0728	0.0137
LTS	0.0696	0.1321	0.0843	0.0733	0.0788	0.0330	0.0763	0.0232	0.0829	0.0178	0.0724	0.0100
LAD	0.3867	0.1831	0.3200	0.1610	0.2719	0.1074	0.2395	0.0813	0.2082	0.0578	0.1614	0.0306
RepMed	0.0149	0.1607	0.0634	0.0684	0.0863	0.0341	0.0874	0.0246	0.0964	0.0200	0.0924	0.0129
Quan2575	-0.0564	0.3210	0.0358	0.0986	0.0444	0.0426	0.0548	0.0277	0.0624	0.0172	0.0718	0.0108
Quan1090	0.4757	0.2310	0.3019	0.1798	0.3144	0.1653	0.3271	0.1657	0.3371	0.1632	0.3363	0.1546
Med/MAD	-0.1414	0.3354	-0.0010	0.1105	0.0464	0.0425	0.0564	0.0279	0.0793	0.0208	0.0906	0.0138
MTM10	-0.0035	0.1221	0.1762	0.1216	0.1865	0.0867	0.1895	0.0743	0.1881	0.0635	0.1713	0.0452
MTM15	0.0484	0.1965	0.1008	0.0997	0.0979	0.0530	0.0961	0.0374	0.0867	0.0204	0.0808	0.0106
Medians	-0.0466	0.2597	0.0706	0.1115	0.1078	0.0538	0.1186	0.0390	0.1349	0.0327	0.1423	0.0257



Şekil 6.14. Weib ( $\gamma = 1$ ,  $\alpha = 1$ ) dağılımı için tahmincilerin MSE grafiği (%10 100 değeri aykırı değer olarak atılmış)

Aykırı değerlerin bulunduğu Weibull dağılımı için elde edilen Çizelge 6.11-6.14'deki tablolar ve Şekil 6.11-6.14'de bulunan grafikler incelendiğinde aşağıdaki sonuçlara ulaşılmaktadır:

- (i) Aykırı değer konfigürasyonuna bağlı olarak tahmincilerin ortaya çıkardığı sonuçların farklı olacağı açıktır.
- (ii) Tüm simulasyon kurguları için  $n = 100$  ve  $250$ 'de ele alınan tüm robust tahminciler daha küçük MSE değeri ortaya çıkarmaktadır.
- (iii) Ayrıca görülmektedir ki  $n$  arttıkça ML tahmincisinin Bias ve MSE değerleri genellikle artış göstermektedir.
- (iv) Çizelge 6.14'de tüm örneklem boyutları için tüm robust tahminciler MSE ve Bias kriterlerine göre ML tahmincisinden daha iyi sonuç vermektedir.
- (v) Robust tahminciler arasında Çizelge 6.14'deki simulasyon kurgusu için M-tahmincilerin performansının diğer tahmincilere göre daha kötü olduğu görülmektedir. M-tahmincilerin x-yönlü aykırı değerlere karşı duyarlı olmalarıdır.

### 6.2.2. Pareto Dağılımı

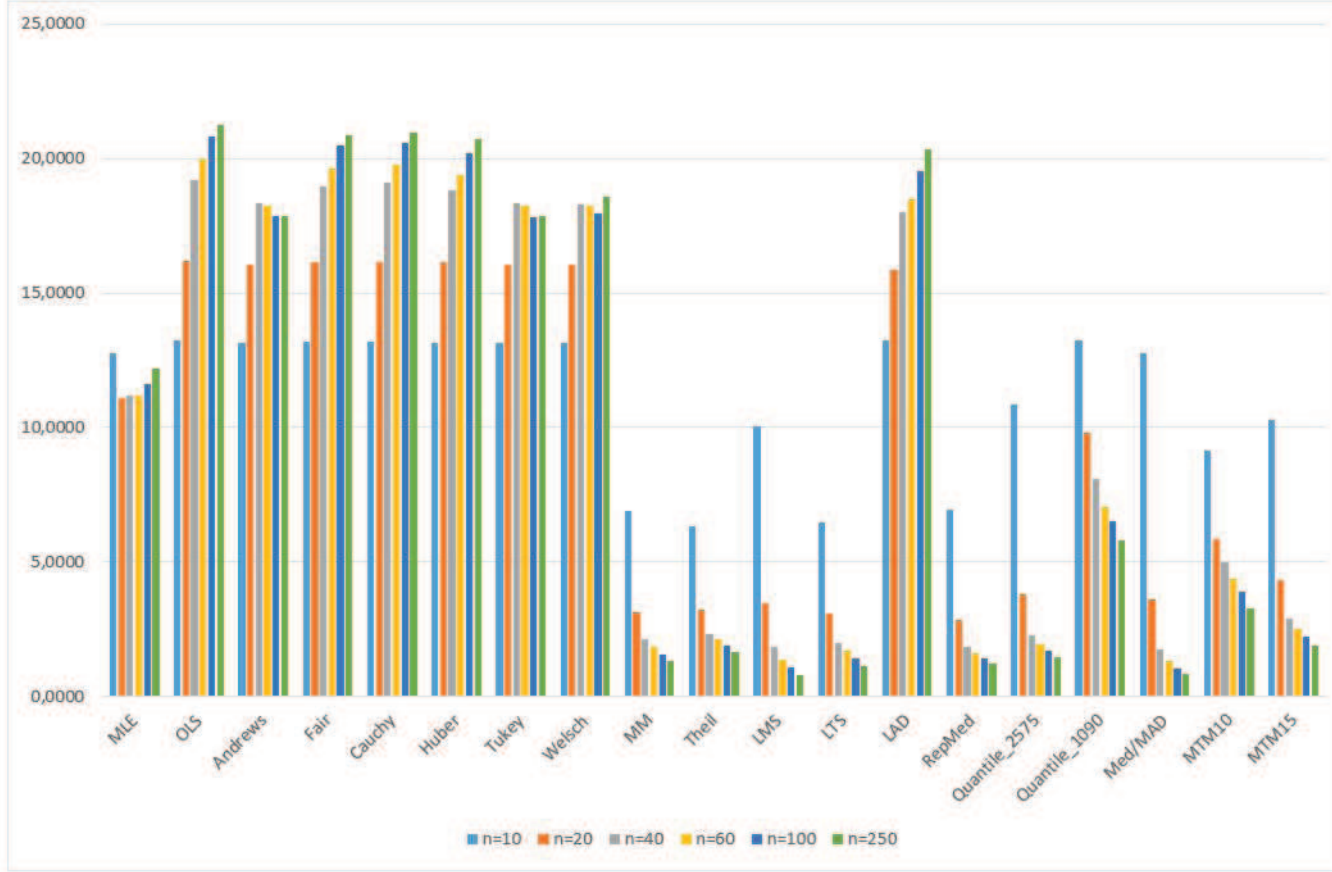
Pareto dağılımı için,

- (i) Pareto ( $\sigma_0 = 1, \alpha_0 = 5$ )'den üretilen veri %10 oranında Pareto ( $\sigma = 1, \alpha = 0.5$ )'den aykırı değer atılarak,
- (ii) Pareto ( $\sigma_0 = 1, \alpha_0 = 5$ )'den üretilen veri %10 oranında Pareto ( $\sigma = 0.2, \alpha = 5$ )'den aykırı değer atılarak,
- (iii) Pareto ( $\sigma = 1, \alpha = 5$ )'den üretilen veri %10 oranında  $U(0, 20)$ 'den aykırı değer atılarak,
- (iv) Pareto ( $\sigma = 1, \alpha = 5$ )'den üretilen veri %10 oranında 100 aykırı değer atılarak,

oluřturulan simulasyon kurguları ile aykırı deęerlerin bulunduęu veri seti zerinde tahmincilerin performansları deęerlendirilmiř olup elde edilen MSE ve Bias deęerleri izelge 6.15-6.18’de ve bu sonulara iliřkin MSE grafikleri Őekil 6.15-6.18’de verilmiřtir.

Çizelge 6.15. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	1.8179	12.7403	2.7126	11.0756	3.1484	11.2057	3.2377	11.1962	3.3660	11.6088	3.4791	12.1915
OLS	2.5810	13.2530	3.6373	16.1886	4.2890	19.1863	4.4361	19.9623	4.5588	20.8253	4.6070	21.2355
Andrews	2.6608	13.1380	3.6081	16.0730	4.1094	18.3384	4.1374	18.2311	4.1113	17.8571	4.1310	17.8523
Fair	2.6219	13.1729	3.6391	16.1495	4.2530	18.9851	4.3880	19.6453	4.5182	20.4745	4.5663	20.8661
Cauchy	2.6146	13.1776	3.6429	16.1671	4.2695	19.0828	4.4086	19.7841	4.5336	20.6061	4.5752	20.9459
Huber	2.6269	13.1416	3.6341	16.1251	4.2258	18.8373	4.3529	19.4108	4.4849	20.2035	4.5509	20.7279
Tukey	2.6602	13.1390	3.6081	16.0732	4.1088	18.3346	4.1376	18.2317	4.1078	17.8337	4.1308	17.8462
Welsch	2.6517	13.1509	3.6024	16.0570	4.1062	18.3108	4.1390	18.2379	4.1295	17.9660	4.2431	18.5775
MM	0.5115	6.9153	0.8973	3.1568	1.0118	2.1281	1.0777	1.8275	1.0877	1.5837	1.0629	1.3055
Theil	0.7529	6.3167	1.0690	3.2221	1.1900	2.3381	1.2568	2.1262	1.2542	1.8924	1.2323	1.6547
LMS	0.0385	10.0280	0.4853	3.4498	0.6593	1.8399	0.7365	1.3684	0.7614	1.0710	0.7730	0.7924
LTS	0.8173	6.4864	0.8921	3.1000	0.9732	2.0103	1.0231	1.6862	1.0184	1.4150	0.9949	1.1433
LAD	2.6663	13.2581	3.5691	15.8399	4.0856	18.0225	4.2065	18.4993	4.3941	19.5341	4.5079	20.3470
RepMed	0.2859	6.9307	0.7396	2.8372	0.9302	1.8660	1.0216	1.6301	1.0419	1.4178	1.0369	1.2140
Quan2575	0.1220	10.8441	0.7848	3.8100	0.9350	2.2569	1.0237	1.9301	1.0991	1.7111	1.1219	1.4635
Quan1090	2.3353	13.2344	2.3452	9.7936	2.3485	8.0975	2.2924	7.0405	2.3061	6.5096	2.2976	5.7939
Med/MAD	-0.6242	12.7639	0.1953	3.6157	0.5256	1.7330	0.6830	1.3012	0.7439	1.0138	0.8260	0.8493
MTM10	0.9181	9.1187	1.5289	5.8694	1.7525	5.0158	1.7685	4.3791	1.7640	3.9047	1.7206	3.2831
MTM15	-0.2544	10.2618	1.1125	4.3355	1.2529	2.9144	1.3063	2.5081	1.3197	2.2086	1.2972	1.8787

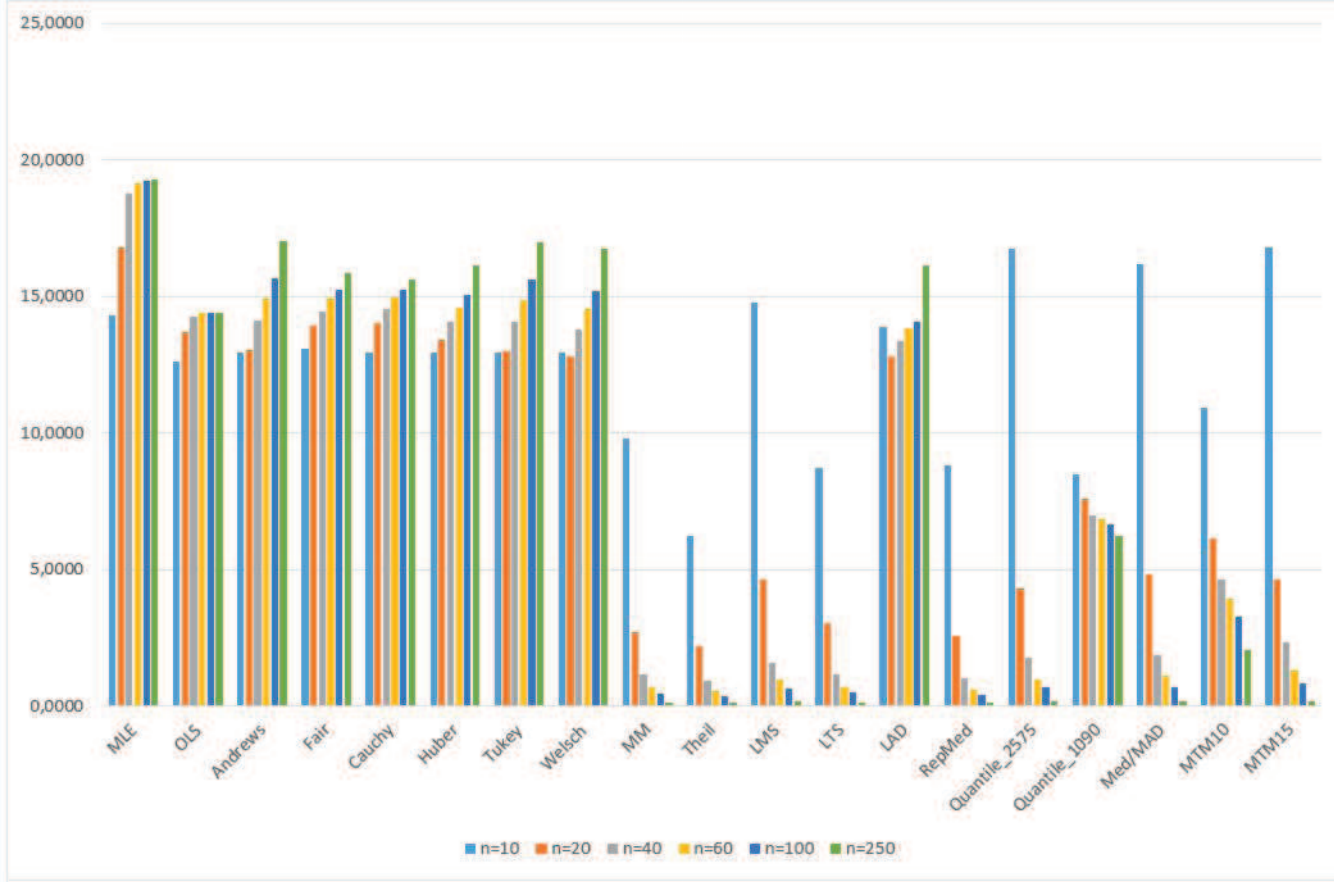


Şekil 6.15. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 1, \alpha = 0.2$ ) dağılımı için tahmincilerin MSE grafiği

**Çizelge 6.16.** 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 0.2, \alpha = 5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri

n	10		20		40		60		100		250	
Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
MLE	2.3637	14.2963	3.7438	16.7825	4.2918	18.7679	4.3702	19.1257	4.3842	19.2214	4.3902	19.2744
OLS	2.6585	12.6352	3.4438	13.6805	3.7332	14.2617	3.7834	14.4051	3.7872	14.3800	3.7914	14.3842
Andrews	2.7821	12.9321	3.1885	13.0353	3.4395	14.1089	3.6392	14.9066	3.8131	15.6488	4.1224	17.0174
Fair	2.7568	13.0838	3.4667	13.9447	3.6968	14.4473	3.8105	14.9183	3.8829	15.2621	3.9796	15.8614
Cauchy	2.7338	12.9439	3.4886	14.0421	3.7366	14.5565	3.8363	14.9831	3.8897	15.2400	3.9527	15.6405
Huber	2.7513	12.9498	3.3756	13.4129	3.6067	14.0540	3.7386	14.6104	3.8375	15.0794	4.0163	16.1592
Tukey	2.7814	12.9345	3.1809	13.0054	3.4278	14.0549	3.6342	14.8750	3.8106	15.6287	4.1197	16.9952
Welsch	2.7763	12.9611	3.1324	12.8057	3.3741	13.7903	3.5729	14.5453	3.7297	15.2077	4.0870	16.7626
MM	-0.6281	9.8012	-0.2047	2.7184	-0.1054	1.1463	-0.0355	0.6800	-0.0415	0.4800	-0.0059	0.1406
Theil	0.0557	6.2099	0.1595	2.2065	0.1710	0.9201	0.1744	0.5629	0.1430	0.3842	0.1199	0.1210
LMS	-0.8092	14.7751	-0.4261	4.6203	-0.1968	1.5964	-0.1218	0.9848	-0.1064	0.6372	-0.0364	0.1719
LTS	0.1035	8.7082	-0.0366	3.0298	-0.0669	1.1645	-0.0377	0.6847	-0.0440	0.4815	-0.0090	0.1365
LAD	2.8629	13.8760	3.0639	12.8259	3.2579	13.3818	3.3660	13.8578	3.3949	14.0521	3.7957	16.1557
RepMed	-0.6063	8.8029	-0.2059	2.5780	-0.0509	1.0152	0.0133	0.6178	0.0206	0.4077	0.0511	0.1150
Quan2575	-1.0062	16.7720	-0.3893	4.3023	-0.2998	1.7709	-0.1883	0.9937	-0.1440	0.6687	-0.0371	0.1937
Quan1090	1.8149	8.4885	1.7889	7.5800	1.8140	6.9999	1.8112	6.8258	1.7831	6.6355	1.7440	6.2139
Med/MAD	-1.1947	16.2043	-0.5399	4.8416	-0.3297	1.8864	-0.2105	1.0916	-0.1748	0.6785	-0.0690	0.1878
MTM10	0.3620	10.9184	0.9155	6.1617	1.0788	4.6380	1.0899	3.9377	1.0575	3.2856	0.8537	2.0541
MTM15	-0.2091	16.8180	0.1498	4.6581	0.1408	2.3183	0.0819	1.3068	0.0233	0.8204	-0.0478	0.1611

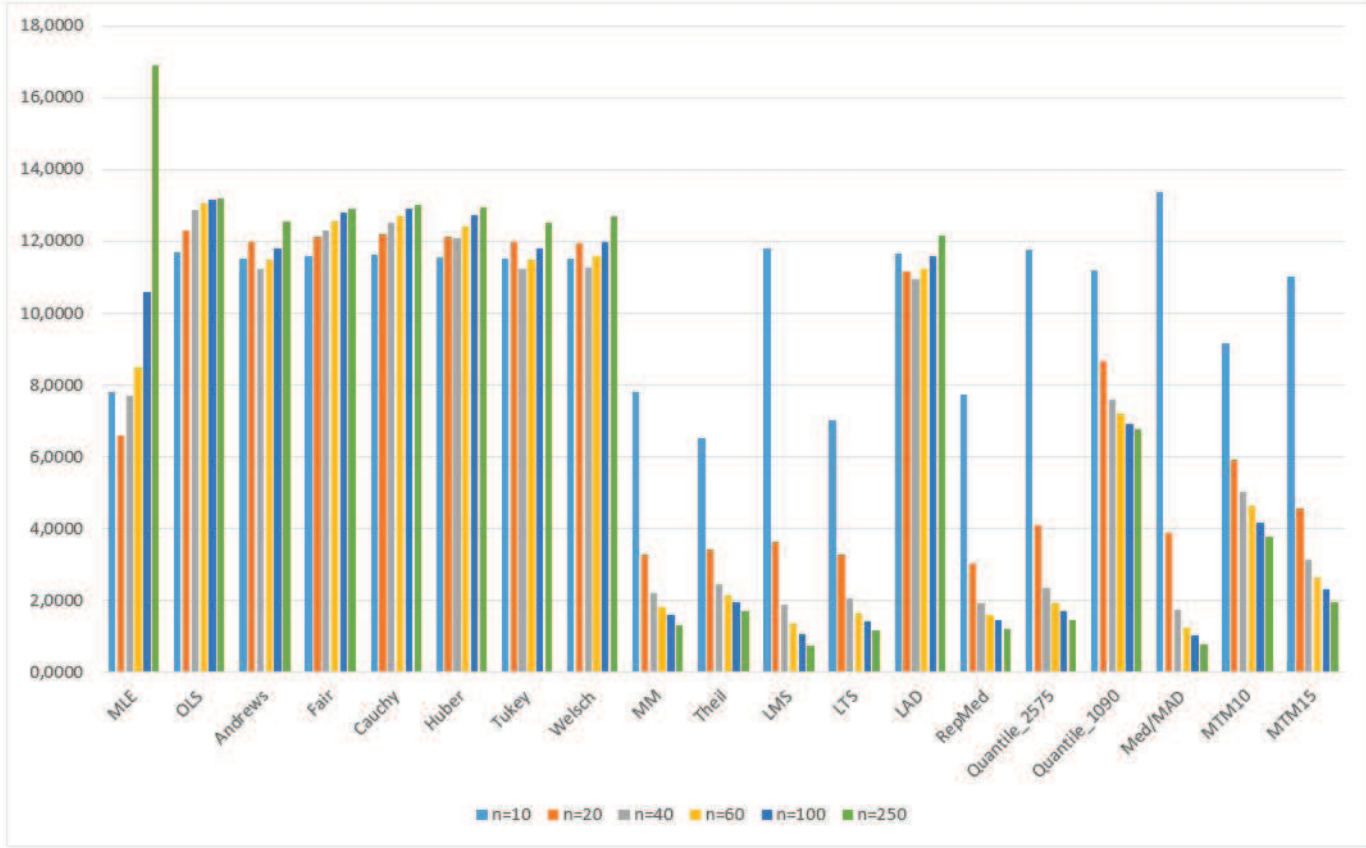




Şekil 6.16. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 Pareto ( $\sigma = 0.2, \alpha = 5$ ) dağılımı için tahmincilerin MSE grafiği

Çizelge 6.17. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 U (0, 20) dağılımı için tahmincilerin MSE ve Bias değerleri

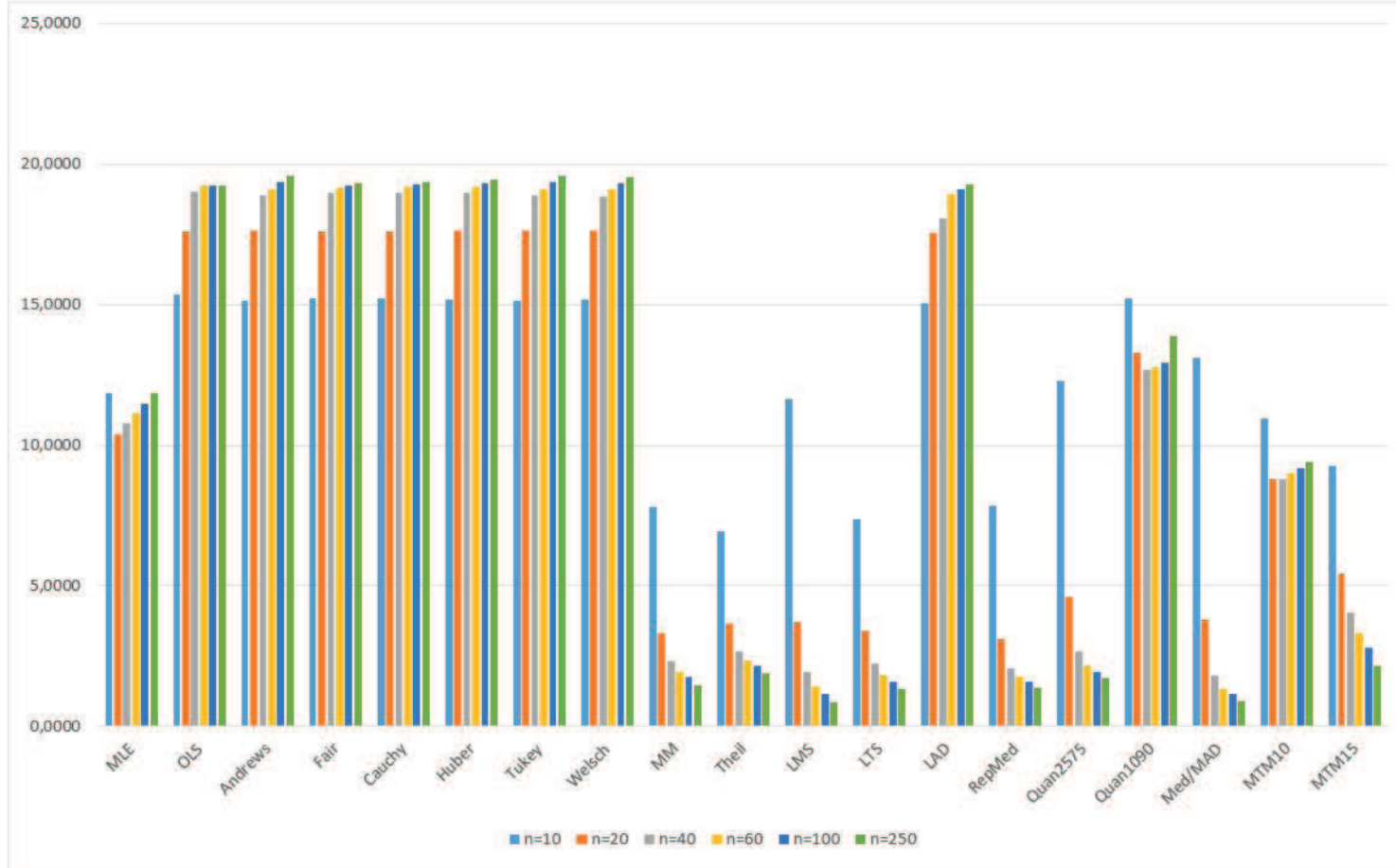
n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	1.2995	7.8009	2.0862	6.5889	2.5156	7.7028	2.7090	8.4718	3.0582	10.5880	3.9662	16.9119
OLS	2.4441	11.6949	3.2022	12.2960	3.5218	12.8697	3.5886	13.0413	3.6219	13.1733	3.6280	13.1824
Andrews	2.5194	11.5178	3.1285	11.9758	3.1337	11.2349	3.2423	11.4667	3.3432	11.8068	3.5123	12.5500
Fair	2.4839	11.6052	3.1828	12.1336	3.4191	12.3058	3.5056	12.5562	3.5645	12.7929	3.5907	12.9189
Cauchy	2.4776	11.6113	3.1955	12.2107	3.4588	12.5185	3.5337	12.7109	3.5827	12.9069	3.6023	13.0000
Huber	2.4881	11.5619	3.1778	12.1337	3.3676	12.0797	3.4718	12.4002	3.5487	12.7177	3.5953	12.9572
Tukey	2.5185	11.5196	3.1282	11.9766	3.1331	11.2300	3.2428	11.4682	3.3439	11.8096	3.5024	12.5028
Welsch	2.5104	11.5338	3.1204	11.9437	3.1418	11.2845	3.2700	11.5968	3.3848	11.9982	3.5427	12.6953
MM	0.4661	7.7881	0.9344	3.2909	1.0414	2.1947	1.0554	1.7929	1.0832	1.6026	1.0579	1.3044
Theil	0.7376	6.5253	1.1188	3.4276	1.2258	2.4529	1.2550	2.1474	1.2782	1.9675	1.2438	1.6899
LMS	-0.0138	11.8063	0.5051	3.6439	0.6294	1.8732	0.7097	1.3522	0.7482	1.0619	0.7365	0.7524
LTS	0.7962	7.0383	0.9264	3.2855	0.9993	2.0590	0.9967	1.6487	1.0211	1.4343	0.9995	1.1601
LAD	2.5199	11.6633	2.9906	11.1559	3.1255	10.9384	3.2438	11.2422	3.3437	11.5760	3.4710	12.1454
RepMed	0.2592	7.7515	0.7850	3.0279	0.9460	1.9064	1.0046	1.6036	1.0469	1.4527	1.0327	1.2105
Quan2575	0.0526	11.7601	0.7837	4.1055	0.9573	2.3542	1.0301	1.9014	1.0904	1.6862	1.1231	1.4617
Quan1090	2.1372	11.2107	2.2286	8.6622	2.3120	7.5808	2.3514	7.1991	2.4067	6.9173	2.4775	6.7832
Med/MAD	-0.6712	13.3787	0.1868	3.8946	0.5116	1.7542	0.6503	1.2412	0.7304	1.0299	0.7896	0.7915
MTM10	0.8836	9.1624	1.5479	5.9143	1.7636	5.0236	1.8228	4.6273	1.8343	4.1657	1.8339	3.7681
MTM15	-0.3364	11.0241	1.1330	4.5468	1.2835	3.1433	1.3190	2.6363	1.3449	2.3208	1.3158	1.9401



Şekil 6.17. 0.90 Pareto ( $\sigma = 1, \alpha = 5$ )+0.10 U (0,20) dağılımı için tahmincilerin MSE grafiği

**Çizelge 6.18.** Pareto ( $\sigma = 1, \alpha = 5$ ) dağılımı için tahmincilerin MSE ve Bias değerleri (%10 100 değeri aykırı değer olarak atılmış)

n	10		20		40		60		100		250	
	Tahminciler	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
ML	2.0088	11.8347	2.8611	10.3935	3.1876	10.7745	3.2940	11.1397	3.3646	11.4535	3.4367	11.8553
LS	2.9361	15.3351	3.9070	17.6115	4.3256	19.0164	4.3789	19.2162	4.3845	19.2299	4.3871	19.2489
Andrews	3.0083	15.1493	3.9085	17.6396	4.2986	18.9071	4.3480	19.1091	4.3965	19.3677	4.4229	19.5645
Fair	2.9709	15.2184	3.9132	17.5886	4.3165	18.9584	4.3717	19.1649	4.3858	19.2430	4.3958	19.3258
Cauchy	2.9655	15.2315	3.9149	17.6014	4.3214	18.9954	4.3774	19.2089	4.3885	19.2667	4.3980	19.3449
Huber	2.9780	15.1991	3.9157	17.6201	4.3153	18.9687	4.3709	19.1896	4.3967	19.3408	4.4099	19.4495
Tukey	3.0074	15.1512	3.9084	17.6401	4.2986	18.9072	4.3463	19.1002	4.3942	19.3531	4.4227	19.5630
Welsch	2.9993	15.1651	3.9051	17.6358	4.2885	18.8451	4.3496	19.1128	4.3920	19.3282	4.4176	19.5182
MM	0.5184	7.8210	0.9736	3.3142	1.1103	2.3218	1.1222	1.9399	1.1556	1.7636	1.1318	1.4674
Theil	0.8617	6.9592	1.1989	3.6399	1.3053	2.6649	1.3231	2.3306	1.3540	2.1692	1.3217	1.8896
LMS	0.0077	11.6577	0.5106	3.7153	0.6774	1.9253	0.7535	1.4057	0.8013	1.1388	0.7971	0.8455
LTS	0.8901	7.3815	1.0103	3.4066	1.0760	2.2307	1.0623	1.7909	1.0920	1.5950	1.0739	1.3137
LAD	2.9868	15.0661	3.9019	17.5516	4.1670	18.0649	4.3290	18.9152	4.3674	19.1106	4.3909	19.2831
RepMed	0.3479	7.8507	0.8447	3.1169	1.0199	2.0539	1.0703	1.7459	1.1176	1.6097	1.1044	1.3651
Quan2575	0.2253	12.2743	0.9413	4.6026	1.0791	2.6581	1.1336	2.1522	1.1925	1.9331	1.2407	1.7364
Quan1090	2.7092	15.2286	2.9099	13.2617	3.0725	12.6608	3.1829	12.7832	3.3067	12.9496	3.5427	13.8705
Med/MAD	-0.5981	13.1132	0.2562	3.8029	0.5803	1.8107	0.7157	1.3310	0.8005	1.1378	0.8662	0.9120
MTM10	1.2248	10.9596	2.0241	8.7915	2.3652	8.7756	2.5233	9.0140	2.6700	9.1991	2.8039	9.3915
MTM15	-0.0938	9.2574	1.3647	5.4072	1.4864	4.0263	1.4452	3.2928	1.4466	2.8110	1.3767	2.1394



Şekil 6.18. Pareto ( $\sigma = 1$ ,  $\alpha = 5$ ) dağılımı için tahmincilerin MSE grafiği (%10 100 değeri aykırı değer olarak atılmış)

Aykırı değerlerin bulunduğu Pareto dağılımı için elde edilen Çizelge 6.15-6.18'deki tablolar ve Şekil 6.15-6.18'de bulunan grafikler incelendiğinde aşağıdaki sonuçlara ulaşılmaktadır:

- (i) Aykırı değer konfigürasyonuna bağlı olarak tahmincilerin ortaya çıkardığı sonuçlar Weibull dağılımında olduğu gibi Pareto dağılımında da farklılık göstermektedir.
- (ii) Örneklem büyüklüğü arttığında MM, LTS, LMS, Theil, Quan2575, RepMed, Med/MAD ve MTM15 robust tahmincilerine ait MSE değerleri azalmaktadır.
- (iii) Tüm simülasyon kurguları için  $n = 20, 40, 60, 100$  ve  $250$ 'de, ele alınan MM, Theil, LMS, LTS, RepMed, Quan2575, Med/MAD ve MTM15 tahmincileri daha küçük MSE değerine sahiptir.
- (iv)  $n = 10$  olduğunda Theil ve RepMed robust tahmincileri diğer tahmincilere göre daha iyi performans ortaya çıkarmaktadırlar.
- (v) Örneklem boyutu  $n$  arttıkça ML, LS ve M-tahmincilerinin Bias ve MSE değerleri artış göstermektedir.
- (vi) Robust tahminciler arasındaki M-tahmincilerin genellikle tüm simülasyon kurguları için performansının, diğer robust tahmincilere göre daha kötü olduğu görülmektedir. Bunun sebebi M-tahmincilerin x-yönlü aykırı değerlere karşı duyarlı olmalarıdır.

## 7. UYGULAMALAR

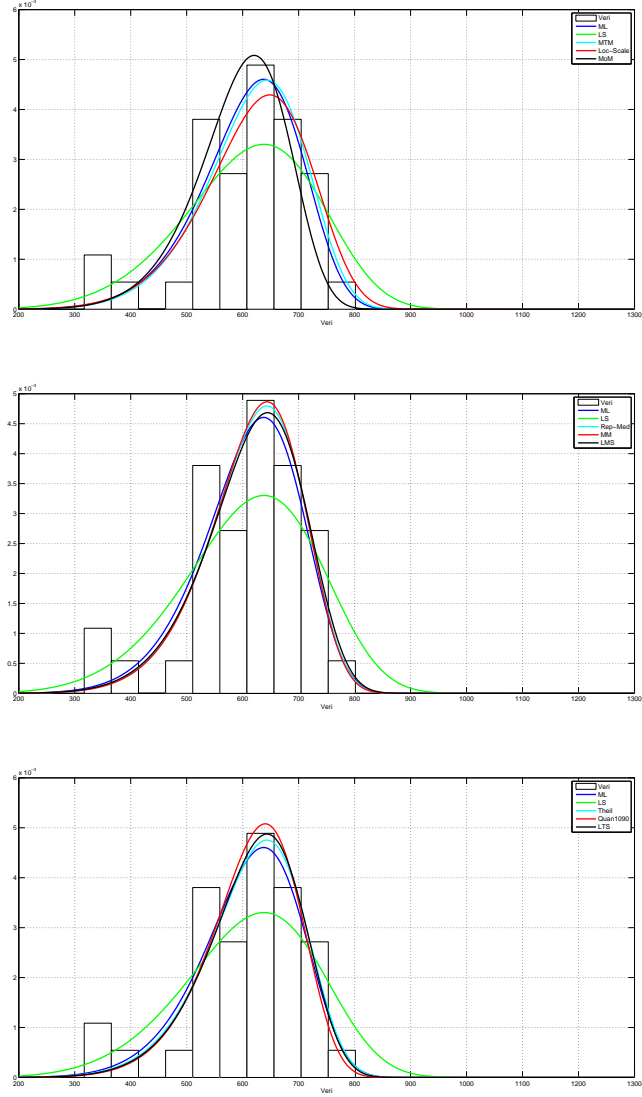
Bu bölümde, tezde ele alınan robust tahmincilerin performanslarını incelemek amacıyla iki ayrı gerçek veri seti üzerinde Weibull ve Pareto dağılımlarının uygulaması yapılmıştır.

Uygulamanın ilk kısmında ilk olarak Hoel (1972)'in çalışmasında kullanılan daha sonra Boudt ve ark. (2011) tarafından Weibull dağılımı ile de çalışılan veri seti kullanılmıştır. Bu veriler 5-6 hafta boyunca 300 radlık radyasyona maruz kalmış 38 tane farenin yaşam sürelerinden oluşmaktadır ve verilerin orjinal hali Çizelge 7.1'de verilmiştir.

**Çizelge 7.1.** Radyasyona maruz kalan farelerin yaşam süresi verisi

Gözlem No	Yaşam Süresi (Gün)	Gözlem No	Yaşam Süresi (Gün)	Gözlem No	Yaşam Süresi (Gün)	Gözlem No	Yaşam Süresi (Gün)
1	317	11	558	21	636	31	697
2	318	12	571	22	643	32	700
3	399	13	586	23	647	33	705
4	495	14	594	24	648	34	712
5	525	15	596	25	649	35	713
6	536	16	605	26	661	36	738
7	549	17	612	27	663	37	748
8	552	18	621	28	666	38	753
9	554	19	628	29	670		
10	557	20	631	30	695		

Öncelikle, simulasyon bölümünde aykırı değerlere karşı dayanıklı olduğu gözlenen MM, Theil, LTS, LMS, kırılmış momentler (MTM), *Med/MAD* (Loc-Sc), tekrarlı medyanlar (Rep-Med) ve kantillere dayalı (Quan) tahminciler ile ML ve LS tahmincileri kullanılarak Weibull dağılımının parametreleri tahmin edilmiştir. Bu tahminlerle elde edilen Weibull olasılık yoğunluk fonksiyonları Şekil 7.1'de verilmiştir.



Şekil 7.1. Yaşam süresi verisi için Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği

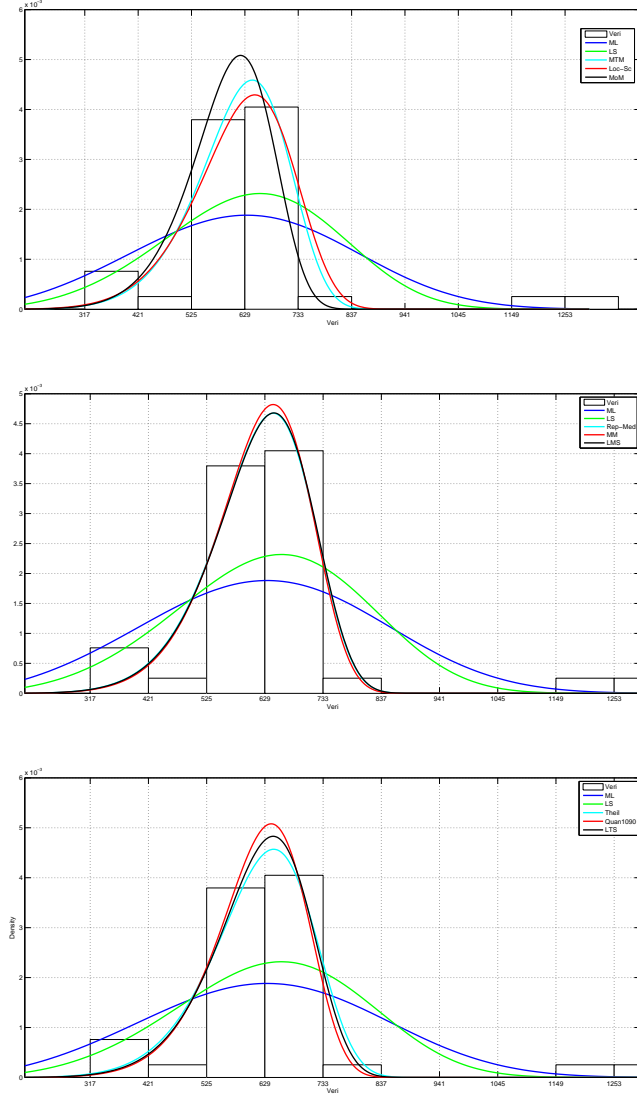


Daha sonra Çizelge 7.1'deki yaşam süresi verilerinin 37 ve 38'inci gözlem değerleri 1248 ve 1253 ile değiştirilerek, veri setine yapay olarak aykırı değerler atılmıştır. Bu veri seti için aynı tahminçiler kullanılarak veri tekrar modellenmiş ve ortaya çıkan sonuçlar grafikler üzerinde Şekil 7.2'de ve elde edilen tahmin değerleri Çizelge 7.2'de verilmiştir.

**Çizelge 7.2.** Yaşam süresi verisi için tahminler

Tahminçiler	Değiştirilmiş veri	Orjinal veri
ML	3,4175	8,0475
LS	4,2842	5,8243
Med/MAD	7,6380	7,6380
MTM10	8,0944	8,0944
MED	8,6380	8,6380
MM	8,4989	8,5742
RepMed	8,2483	8,4520
LMS	8,2704	8,2704
Theil	8,0758	8,3765
LTS	8,5148	8,5883
Quant	8,9020	8,9020

Grafiklerden de görüldüğü üzere ML ve LS tahminçileri aykırı değerlerden oldukça etkilenmiştir ve bu durum olasılık yoğunluk fonksiyonu grafiklerine yansımıştır. Robust tahminçiler ise beklenildiği üzere aykırı değerlerden etkilenmemiştir. Bu durum Çizelge 7.2'deki tahmin değerlerinde de görülmektedir.



Şekil 7.2. Aykırı değer içeren yaşam süresi verisi için Weibull dağılımının olasılık yoğunluk fonksiyonları grafiği

Uygulamanın ikinci kısmında ise Marr (1920)'in çalışmasında yer alan 14 birimlik sulama verisi Pareto dağılımı ile modellenip, tahmincilerin gerçek veri üzerindeki performansını görmek amacıyla kullanılmıştır. Veriler Çizelge 7.3'de verilmiştir.

**Çizelge 7.3.** Sulama verisi

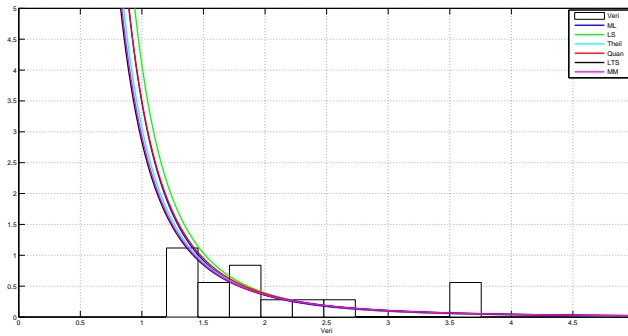
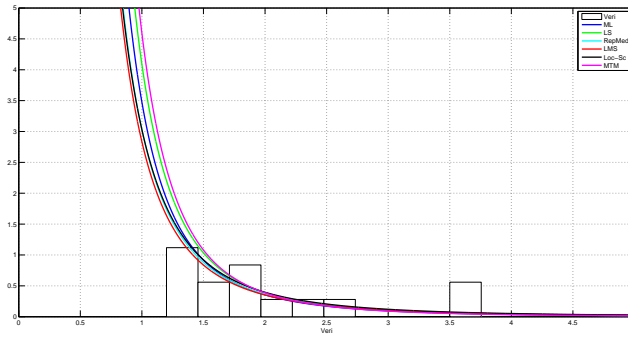
Gözlem No	Sulama Verisi	Gözlem No	Sulama Verisi
1	1.8	8	1.5
2	1.9	9	1.5
3	2.5	10	1.2
4	1.4	11	1.3
5	1.3	12	1.8
6	2.1	13	3.5
7	2.3	14	3.5

İlk uygulamada olduğu gibi öncelikle orjinal veriler kullanılarak ML, LS ve robust tahmincileri yardımıyla tahminleme yapılmış, daha sonra verinin 13. ve 14. değerleri 35 ile değiştirilmiş yani veri setine yapay olarak aykırı değerler atılmış ve tekrar tahminleme yapılmıştır. Bu sonuçlara ilişkin grafikler Şekil 7.3-7.4'te ve elde edilen tahmin değerleri Çizelge 7.4'de verilmiştir.

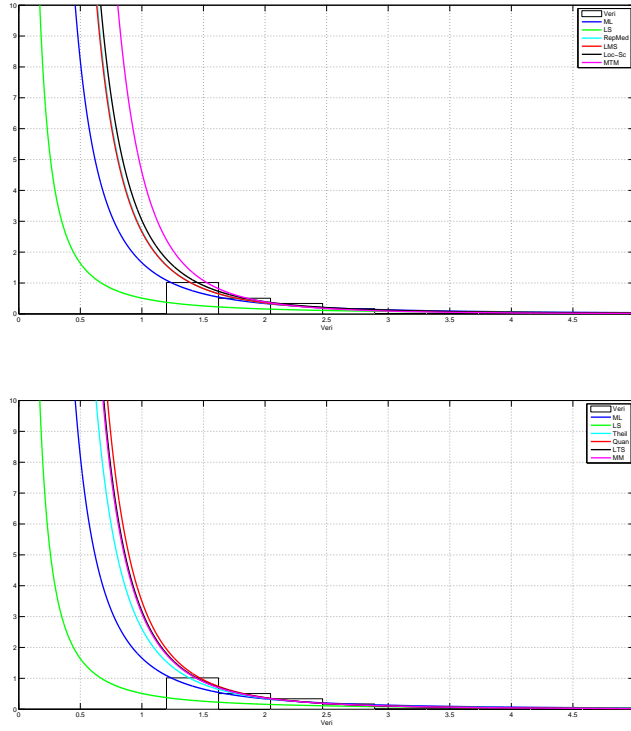
Çizelge 7.4'deki sonuçlar incelendiğinde ML ve LS tahmincilerinin aykırı değerlerden etkilendikleri açık bir şekilde görülmektedir. Robust tahminciler ise her iki durumda da birbirine yakın sonuçlar vermiştir. Diğer bir ifade ile veri setindeki aykırı değerlerden etkilenmemişlerdir. Bu sonuç simülasyon sonuçlarıyla paraleldir.

Çizelge 7.4. Sulama verisi için tahminler

Tahminciler	Orjinal veri	Değiştirilmiş veri
ML	2,2891	1,3058
LS	2,3903	0,6846
Med/MAD	1,9387	1,9387
MTM15	2,5698	2,5698
MM	2,0207	2,0861
RepMed	2,0487	1,9300
LMS	2,0077	1,9017
Theil	2,0624	1,8950
LTS	1,9928	2,1054
Quan2575	2,2130	2,2130



Şekil 7.3. Sulama verisi için Pareto dağılımının olasılık yoğunluk fonksiyonları grafiği



Şekil 7.4. Aykırı değer içeren sulama verisi için Pareto dağılımının olasılık yoğunluk fonksiyonları grafiği

## 8. SONUÇLAR

Bu tez çalışmasının amacı, veride aykırı değer olması durumunda, dağılımsal parametrelerin tahmininde kullanılan çeşitli sağlam (robust) tahmincilerin incelenmesidir. Ele alınan tahminciler; doğrusal regresyon modeli için robust tahminciler (M-tahminciler (Andrews, Tukey, Welsch, Cauchy, Huber ve Fair), MM-tahmincisi, LAD, LMS, LTS, tekrarlı medyanlar tahmincisi ve Theil tahmincisi), kantillere dayalı robust tahminci, Konum-ölçek ailesine dönüştürülebilen dağılımlar için robust tahminciler, medyanlar metodu, kırılmış momentler metodu şeklinde sıralanabilir. Bu tahminciler farklı açılardan incelenmiş, Weibull ve Pareto dağılımları için performansları araştırılmıştır.

Yapılan Monte-Carlo simülasyonu çalışmasının sonuçları aşağıdaki gibi özetlenebilir:

- (i) Aykırı değerlerin olmadığı veride, örneklem boyutu arttıkça ele alınan robust tahmincilerin Bias ve MSE değerleri azalmaktadır.
- (ii) Aykırı değerlerin olmadığı veride, örneklem boyutu  $n = 10$  ve  $n = 20$  için MSE açısından M-tahminciler ve LAD tahmincisi, hem Weibull hem de Pareto dağılımları için ML tahmincisinden daha iyi sonuçlar üretmektedir. Diğer örneklem boyutları için M-tahmincisi, LAD ve Theil tahmincileri ML tahmincisinden sonra en küçük MSE değerleri ortaya çıkartmaktadırlar.
- (iii) Aykırı değerlerin bulunduğu durumlarda MSE kriterine göre MM, Theil, LMS, LTS, RepMed, Quan2575, Med/MAD ve MTM15 tahmincileri hem Weibull hem de Pareto dağılımları için ML ve LS tahmincilerinden daha iyi performans göstermektedirler.
- (iv) M-tahminciler veride aykırı değer olduğu çoğu durumda ML tahmincisinden bile daha kötü sonuçlar ortaya çıkartmaktadır.

Ayrıca tezde ele alınan robust tahmincilerin aykırı değerlere olan dayanıklılığı iki ayrı gerçek veri seti üzerinde de gözlemlenmiştir.

Gelecek çalışma olarak tezde ele alınan tahmincilerin sansürlenmiş veriler için hesaplanması ve performanslarının incelenmesi hedeflenmektedir.

## KAYNAKLAR

- Akdi, Y. (2010), *Matematiksel İstatistiğe Giriş*, Gazi, Türkiye.
- Andersen, R. (2008), *Modern Methods for Robust Regression*, SAGE Publications.
- Bajorunaite, R. & Brazauskas, V. (2008), “Method of trimmed moments for robust fitting of parametric failure time models”, *International Journal of Statistics*, **3**, 341-360.
- Bickel, P. J. (1975), “One-step Huber estimates in the linear model”, *J. Amer. Statist. Assoc.*, **70**, 428-434.
- Boudt, K., Çalışkan, D. & Croux, C. (2011), “Robust explicit estimators of Weibull parameters”, *Metrika*, **73**, 187-209.
- Brazauskas, V. & Serfling, R. (2000), “Robust and efficient estimation of the tail index of a single-parameter Pareto distribution”, *North American Actuarial Journal*, **4**(4), 12-27.
- Brazauskas, V. & Serfling, R. (2003), “Favorable estimators for fitting pareto models: A study using goodness-of-fit measures with actual data”, *ASTIN Bulletin*, **33**(2), 365-381.
- Brazauskas, V. & Kleefeld, A. (2009), “Robust and efficient fitting of the generalized Pareto distribution with actuarial applications in view”, *Insurance Mathematics & Economics*, **45**(3), 424-435.
- Croux, C. & Rousseuw, P. J. (1992), “Time-Efficient Algorithms for Two Highly Robust Estimators of Scale”, *Computational Statistics*, **1**, 411-428.
- Dodge, Y. & Jureckova, J. (2002), *Adaptive Regression*, Springer, New York.
- Ergül, B. , *Robust Regresyon ve Uygulamaları*, Yüksek Lisans Tezi, Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir, 2006.
- Gather, U. & Schultze, V. (1999), “Robust estimation of scale of an exponential distribution”, *Statistica Neerlandica*, **53**(3), 327-341.



- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.
- He, X.& Fung, W. K. (1999), “Method of medians for lifetime data with Weibull models”, *Statistics in Medicine*, **18**, 1993-2009.
- Holland, P.W. & Welsch, R.E. (1977), “Robust regression using iteratively reweighted least-squares”, *Communications in Statistics - Theory and Methods*, **A6**(9), 813-827.
- Huber, P. J. (1964), “Robust estimation of a location parameter”, *Annals of Mathematical Statistics*, **35**, 73-101.
- Huber, P.J. (1973), “Robust regression: asymptotics, conjectures and Monte Carlo” *Annals of Statistics*, **1**(5), 799-821.
- Huber, P.J. (1981), *Robust Statistics*, Wiley.
- Jureckova, J. & Picek, J. (2006), *Robust Statistical Methods with R*, Chapman and Hall.
- Kantar, Y. M. (2006), “The Other Derivation of Maximum Spacing Method Second part. A comparison of Maximum Likelihood and Maximum Spacing Method for the 3-parameters Weibull distribution”, *Pakistan Journal of Statistics*, **22**(3).
- Kantar, Y. M. & Şenoğlu, B. (2008), “A comparative study for the location and scale parameters of the Weibull distribution with a given shape parameter”, *Computers & Geosciences*, **34**(12), 1900-1909.
- Kantar, Y. M. & Arık, İ. (2014), “M-Estimation of log-logistic distribution parameters with outliers”, *International Journal of Agricultural and Statistical Sciences*, (Basımda).
- Kantar, Y. M. & Yıldırım, V. (2014), “Robust estimation for parameters of the Extended Burr Type III distribution”, *Communications in Statistics-*

*Simulation and Computation*, (Basımda DOI#10.1080/03610918.2013.-839032).

Karagöz, D. , *Sağlam Varyans Çözümlemesi*, Doktora Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2012.

Kayhan, Y. , *Regresyon Çözümlemesinde Sağlam Konum Kestiricilerinin Karşılaştırılması*, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2006.

Marks, N. B. (2005), “Estimation of Weibull Parameters from Common Percentiles”, *Journal of Applied Statistics*, **32**, 1, 17-24.

Maronna, R. A., Martin, D. R. & Yohai, V. J. (2006), *Robust Statistics Theory & Methods*, Wiley, İngiltere.

Marr, J. C. (1927), *The use and duty of water in the Salt River Valley*, University of Arizona.

Montgomery, D.C., Peck, E.A. & Vining, G.G. (2001), *Introduction to Linear Regression Analysis*, John Wiley & Sons, New York.

Mutan, O. C. , *Comparison Of Regression Techniques Via Monte Carlo Simulation*, Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2004.

Mutan, O. C. & Şenoğlu, B. (2009), “A Monte Carlo comparison of regression estimators when the error distribution is long-tailed symmetric.”, *Journal of Modern Applied Statistical Methods*, **8**, 1, 161-172.

Olive, D. (2006), “Robust estimators for transformed location scale families”, *Mimeo*.

Öztürk, F., Akdi, Y., Aydoğdu, H. & Karabulut, İ. (2006), *Parametre Tahmini & Hipotez Testleri*, Bıçaklar, Türkiye.

Rousseuw, P.J., Leroy, A. M. & Daniels, B. (1984), “Resistant Line Fitting in Actuarial Science”, *Advanced Science Institutes Series C*, **88**, 315-332.

- Rousseau, P.J.& Leroy, A. M. (1987), *Robust Regression & Outlier Detection*, John Wiley & Sons, Amerika.
- Rousseau, P. J. & Croux, C. (1993), “Alternatives to the Median Absolute Deviation”, *Journal of the American Statistical Association*, **88**, 424, 1273-1283.
- Shier, D. R. & Lawrence K. D. (1984), “A comparison of robust regression techniques for the estimation of Weibull parameters”, *Communications in Statistics-Simulation and Computation*, **3**(6), 743-750.
- Siegel A. F. (1982), “Robust regression using repeated medians”, *Biometrika*, **69**, 242-244.
- Skinner, K. R., Keats, J. B. & Zimmer, W. J. (2001), “A comparison of three estimators of the Weibull parameters”, *Quality and Reliability Engineering International*, **17**, 249-256.
- Stuart, C. (2011), *Robust Regression*.
- Tiryakioğlu, M. & Hudak, D. (2007), “On estimating Weibull modulus by the linear regression method”, *Journal of Materials Science*, **42**, 10173-10179.
- Wang, F. K. & Lee, C. W (2010), “An m-estimator for estimating the extended Burr Type III parameters with outliers”, *Communications in Statistics - Theory and Methods*, **40**(2), 304-322.
- Wang, F. K. & Cheng, Y. (2010), “Robust regression for estimating the Burr XII parameters with outliers”, *Journal of Applied Statistics*, **37**(5), 807-819.
- Wu L.L. (1985), “Robust M-Estimation of Location and Regression”, *Sociological Methodology*, **15**, 316-388.
- Yavuz A.A. (2012), “Estimation of the shape parameter of the Weibull distribution using linear regression methods: non-censored samples”, *Quality and Reliability Engineering International*, **29**(8), 1207-1219.

- Yıldırım, V. & Kantar, Y. M. (2014), “Comprehensive evaluation of m-estimators for linear regression models”, *International Journal of Mathematics and Statistics*, **151**.
- Yıldırım, V. , *Doğrusal Regresyonda M-Tahmincilerin İncelenmesi*, Yüksek Lisans Tezi, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir, 2012.
- Yohai, V. J. (1987), “High breakdown point and high efficiency robust estimates for regression”, *The Annals of Statistics*, **15**, 20, 642-656.
- Zhang, L. F., Xie, M. & Tang, L. C. (2006), “Robust regression using probability plots for estimating the Weibull shape parameter”, *Quality and Reliability Engineering International*, **22**, 905-917.
- Zyl, J. M. V (2012), “A median regression model to estimate the parameters of the three-parameter generalized Pareto distribution”, *Communication in Statistics-Simulation & Computation*, **41**, 544-553.
- Zyl, J. M. V & Schall, R. (2012), “Parameter estimation through weighted least-squares rank regression with specific reference to the Weibull and Gumbel distributions”, *Communication in Statistics-Simulation & Computation*, **41**(9), 1654-1666.