



**KARARLI REGRESYON ANALİZİNDE SAĞLAM TAHMİN
YÖNTEMLERİNİN PERFORMANSLARININ İNCELENMESİ**

Enise KORALTAN

**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANABİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

OCAK 2015

Enise KORALTAN tarafından hazırlanan “KARARLI REGRESYON ANALİZİNDE SAĞLAM TAHMİN YÖNTEMLERİNİN PERFORMANSLARININ İNCELENMESİ” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi İstatistik Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Yrd. Doç. Dr. Filiz KARDİYEN

İstatistik Anabilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum

.....

Başkan : Prof. Dr. Hasan Hüseyin TATLIDİL

İstatistik Anabilim Dalı, Hacettepe Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum

.....

Üye : Prof. Dr. Mehmet Akif BAKIR

İstatistik Anabilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum

.....

Tez Savunma Tarihi: 30/01/2015

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

.....

Prof. Dr. Şeref SAĞIROĞLU

Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Enise Koraltan

30.01.2015

KARARLI REGRESYON ANALİZİNDE SAĞLAM TAHMİN YÖNTEMLERİNİN PERFORMANSLARININ İNCELENMESİ

(Yüksek Lisans Tezi)

Enise KORALTAN

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Ocak 2015

ÖZET

Regresyon analizinde yaygın olarak kullanılan En Küçük Kareler yöntemi hata terimleri normal dağılmadığında veya veri seti aykırı değer içerdiğinde iyi bir tahmin yöntemi olmaktan uzaklaşacaktır. Klasik doğrusal regresyon modeli varsayımlarının sağlanamaması durumunda En Küçük Kareler yöntemi minimum varyanslı ve yansız tahmin olma özelliğini kaybedecektir. Söz konusu olumsuzlukların üstesinden gelebilmek amacıyla geliştirilen sağlam tahmin yöntemleri regresyon analizinin varsayımları geçerli olmadığında dahi güvenilir sonuçlar vermektedir. Yapılan çalışmalar gerçek hayatta verilerin normal dağılma olasılığının oldukça düşük olduğunu göstermektedir. Ekonomiden fiziğe birçok alanda karşımıza çıkan alfa kararlı dağılımlar çarpık ve kalın kuyruklu yapısı ile oldukça ilgi gören bir dağılımdır. Veride aykırı değerler bulunduğunda hata terimlerinin dağılımı normalden uzaklaşır ve kalın kuyruklu bir yapı gösterir. Bu tezde hata terimlerinin alfa kararlı dağıldığı basit doğrusal regresyon modeline bazı sağlam tahmin yöntemleri dört farklı örnek çapı ve alfa kararlı dağılımın farklı parametre değerleri için uygulanmış ve bu yöntemlerin performansları karşılaştırılmıştır.

Bilim Kodu : 205.1.066

Anahtar Kelimeler : Aykırı değer, sağlam yöntemler, kararlı dağılım, kararlı regresyon, kalın kuyruk

Sayfa Adedi : 107

Danışman : Yrd. Doç. Dr. Filiz KARDİYEN

PERFORMANCE EVALUATION OF ROBUST ESTIMATORS UNDER THE STABLE
REGRESSION MODEL

(M. Sc. Thesis)

Enise KORALTAN

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

January 2015

ABSTRACT

The Ordinary Least Squares estimation is the best method when the assumptions of classical linear regression are met. When some of these assumptions are violated, least squares regression can perform poorly. Robust regression is an alternative to ordinary least squares that can be appropriately used when the errors are nonnormal and/or a data set contains outliers. Especially in finance, the typical empirical distributions have more observations around the mean and in the extreme tails than the normal. Among the most frequently candidates for describing these distributions is the family of probability distributions known as stable Paretian. In this theses, the performances of various robust regression estimators are investigated where the errors follow a member of the class of Stable Paretian Distributions.

Science Code : 205.1.066

Key Words : Outlier, robust methods, stable distributions, stable regression, fat tail

Page Number : 107

Supervisor : Assist. Prof. Dr. Filiz KARDİYEN

TEŞEKKÜR

Çalışmalarım boyunca değerli bilgi ve yardımlarını esirgemedi beni yönlendiren hocam Yrd. Doç. Dr. Filiz KARDİYEN'e, manevi destekleri ile her zaman yanımda olan biricik annem Ayşe KORALTAN ve sevgili kardeşlerim başta olmak üzere büyük ailemin her ferdine teşekkürü bir borç bilirim.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	xii
SİMGELER VE KISALTMALAR.....	xv
1. GİRİŞ.....	1
2. REGRESYON ANALİZİ.....	9
2.1. En Küçük Kareler Yöntemi.....	12
3. AYKIRI DEĞERLER.....	15
4. SAĞLAM YÖNTEMLER.....	19
4.1. Sağlam Yöntemlere Neden İhtiyaç Duyulur?.....	19
4.2. Sağlam Regresyon.....	21
4.3. Sağlam Yöntemin Özellikleri.....	21
4.3.1. Kırılma noktası (Breakdown point).....	21
4.3.2. Asimptotik etkinlik.....	23
4.4. L Tipi Tahmin Ediciler.....	24
4.4.1. En küçük mutlak sapmalar yöntemi (Least absolute deviation)	24
4.5. M Tipi Tahmin Ediciler.....	27
4.6. S Tipi Tahmin Ediciler.....	35
4.7. MM Tahmini.....	36
4.8. En Küçük Budanmış Kareler Yöntemi (Least Trimmed Squares).....	37

	Sayfa
4.9. En Küçük Medyan Kareler Yöntemi (Least Median Squares).....	40
4.10. Kantil Regresyon.....	41
4.11. Bootstrap (Yeniden Örnekleme)	45
5. KARARLI DAĞILIMLAR.....	51
5.1. Kararlı Dağılımların Parametrelendirilmesi.....	56
5.2. Kararlı Dağılımların Parametre Tahmini.....	58
6. KARARLI REGRESYON.....	63
7.UYGULAMA.....	67
8. SONUÇ ve ÖNERİLER.....	95
KAYNAKLAR	97
ÖZGEÇMİŞ	105
DİZİN.....	107

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 7.1. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=15$ için simülasyon sonuçları.....	71
Çizelge 7.2. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=20$ için simülasyon sonuçları.....	71
Çizelge 7.3. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=30$ için simülasyon sonuçları.....	71
Çizelge 7.4. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=40$ için simülasyon sonuçları	72
Çizelge 7.5. Senaryo 2 ($S_1(0,1,0)$) $n=15$ için simülasyon sonuçları.....	72
Çizelge 7.6. Senaryo 2 ($S_1(0,1,0)$) $n=20$ için simülasyon sonuçları.....	72
Çizelge 7.7. Senaryo 2 ($S_1(0,1,0)$) $n=30$ için simülasyon sonuçları.....	73
Çizelge 7.8. Senaryo 2 ($S_1(0,1,0)$) $n=40$ için simülasyon sonuçları.....	73
Çizelge 7.9. Senaryo 3 ($S_{1,3}(0,1,0)$) $n=15$ için simülasyon sonuçları.....	73
Çizelge 7.10. Senaryo 3 ($S_{1,3}(0,1,0)$) $n=20$ için simülasyon sonuçları	74
Çizelge 7.11. Senaryo 3 ($S_{1,3}(0,1,0)$) $n=30$ için simülasyon sonuçları.....	74
Çizelge 7.12. Senaryo 3 ($S_{1,3}(0,1,0)$) $n=40$ için simülasyon sonuçları.....	74
Çizelge 7.13. Senaryo 4 ($S_{1,7}(0,1,0)$) $n=15$ için simülasyon sonuçları.....	75
Çizelge 7.14. Senaryo 4 ($S_{1,7}(0,1,0)$) $n=20$ için simülasyon sonuçları.....	75
Çizelge 7.15. Senaryo 4 ($S_{1,7}(0,1,0)$) $n=30$ için simülasyon sonuçları.....	75
Çizelge 7.16. Senaryo 4 ($S_{1,7}(0,1,0)$) $n=40$ için simülasyon sonuçları.....	76
Çizelge 7.17. Senaryo 5 $S_{0,7}(1,1,0)$, $n=15$ için simülasyon sonuçları.....	76
Çizelge 7.18. Senaryo 5 ($S_{0,7}(1,1,0)$) $n=20$ için simülasyon sonuçları.....	76
Çizelge 7.19. Senaryo 5 ($S_{0,7}(1,1,0)$) $n=30$ için simülasyon sonuçları.....	77
Çizelge 7.20. Senaryo 5 ($S_{0,7}(1,1,0)$) $n=40$ için simülasyon sonuçları.....	77
Çizelge 7.21. Senaryo 6 $S_{0,7}(-1,1,0)$ $n=15$ için simülasyon sonuçları.....	77
Çizelge 7.22. Senaryo 6 $S_{0,7}(-1,1,0)$ $n=20$ için simülasyon sonuçları.....	78
Çizelge 7.23. Senaryo 6 $S_{0,7}(-1,1,0)$ $n=30$ için simülasyon sonuçları.....	78

Çizelge	Sayfa
Çizelge 7.24. Senaryo 6 $S_{0,7}(-1,1,0)$ $n=40$ için simülasyon sonuçları.....	78
Çizelge 7.25. Senaryo 7 $S_1(1,1,0)$ $n=15$ için simülasyon sonuçları.....	79
Çizelge 7.26. Senaryo 7 $S_1(1,1,0)$ $n=20$ için simülasyon sonuçları.....	79
Çizelge 7.27. Senaryo 7 $S_1(1,1,0)$ $n=30$ için simülasyon sonuçları.....	79
Çizelge 7.28. Senaryo 7 $S_1(1,1,0)$ $n=40$ için simülasyon sonuçları	80
Çizelge 7.29. Senaryo 8 $S_1(-1,1,0)$ $n=15$ için simülasyon sonuçları.....	80
Çizelge 7.30. Senaryo 8 $S_1(-1,1,0)$ $n=20$ için simülasyon sonuçları.....	80
Çizelge 7.31. Senaryo 8 $S_1(-1,1,0)$ $n=30$ için simülasyon sonuçları.....	81
Çizelge 7.32. Senaryo 8 $S_1(-1,1,0)$ $n=40$ için simülasyon sonuçları.....	81
Çizelge 7.33. Senaryo 9 $S_{1,3}(1,1,0)$ $n=15$ için simülasyon sonuçları.....	81
Çizelge 7.34. Senaryo 9 $S_{1,3}(1,1,0)$ $n=20$ için simülasyon sonuçları	82
Çizelge 7.35. Senaryo 9 $S_{1,3}(1,1,0)$ $n=30$ için simülasyon sonuçları.....	82
Çizelge 7.36. Senaryo 9 $S_{1,3}(1,1,0)$ $n=40$ için simülasyon sonuçları.....	82
Çizelge 7.37. Senaryo 10 $S_{1,3}(-1,1,0)$ $n=15$ için simülasyon sonuçları.....	83
Çizelge 7.38. Senaryo 10 $S_{1,3}(-1,1,0)$ $n=20$ için simülasyon sonuçları.....	83
Çizelge 7.39. Senaryo 10 $S_{1,3}(-1,1,0)$ $n=30$ için simülasyon sonuçları.....	83
Çizelge 7.40. Senaryo 10 $S_{1,3}(-1,1,0)$ $n=40$ için simülasyon sonuçları.....	84
Çizelge 7.41. Senaryo 11 $S_{1,7}(1,1,0)$ $n=15$ için simülasyon sonuçları.....	84
Çizelge 7.42. Senaryo 11 $S_{1,7}(1,1,0)$ $n=20$ için simülasyon sonuçları.....	84
Çizelge 7.43. Senaryo 11 $S_{1,7}(1,1,0)$ $n=30$ için simülasyon sonuçları.....	85
Çizelge 7.44. Senaryo 11 $S_{1,7}(1,1,0)$ $n=40$ için simülasyon sonuçları.....	85
Çizelge 7.45. Senaryo 12 $S_{1,7}(-1,1,0)$ $n=15$ için simülasyon sonuçları.....	85
Çizelge 7.46. Senaryo 12 $S_{1,7}(-1,1,0)$ $n=20$ için simülasyon sonuçları.....	86
Çizelge 7.47. Senaryo 12 $S_{1,7}(-1,1,0)$ $n=30$ için simülasyon sonuçları.....	86

Çizelge	Sayfa
Çizelge 7.48. Senaryo 12 $S_{1,7}(-1,1,0)$ $n=40$ için simülasyon sonuçları.....	86

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Doğrusal regresyon modelinin grafik gösterimi	9
Şekil 2.2. Tahmin edilen regresyon doğrusu	13
Şekil 3.1. Beş noktalı orijinal veri kümesi ve EKK regresyon doğrusu	15
Şekil 3.2. Beş noktalı y-yönünde aykırı gözlem içeren veri kümesi	15
Şekil 3.3. Beş noktalı orijinal veri kümesi ve EKK regresyon doğrusu	16
Şekil 3.4. Beş noktalı x-yönünde aykırı gözlem içeren veri kümesi	16
Şekil 3.5. (x_i, y_i) ; iyi kaldıraç noktası.	17
Şekil 4.1. Bazı M tahminlerinin ρ , ψ ve w fonksiyonlarının grafik gösterimi.....	34
Şekil 4.2. Kantil regresyonun ρ fonksiyonu.....	43
Şekil 5.1. Standart Normal $N(0,1)$, Cauchy $(1,0)$, Levy $(1,0)$ yoğunluklarının şekilleri	55
Şekil 5.2. $S_{\alpha}(0,5;1;0)$ parametrelemesinde, $\alpha=0,5; 0,75; 1; 1,25$ ve $1,50$ için kararlı yoğunluklar.....	56
Şekil 7.1. $S_{0,7}(-1;1;0)$, $S_{0,7}(0;1;0)$, $S_{0,7}(1;1;0)$ için kararlı yoğunluklar.....	69
Şekil 7.2. $S_1(-1;1;0)$, $S_1(0;1;0)$, $S_1(1;1;0)$ için kararlı yoğunluklar.....	69
Şekil 7.3. $S_{1,3}(-1;1;0)$, $S_{1,3}(0;1;0)$, $S_{1,3}(1;1;0)$ için kararlı yoğunluklar.....	70
Şekil 7.4. $S_{1,7}(-1;1;0)$, $S_{1,7}(0;1;0)$, $S_{1,7}(1;1;0)$ kararlı yoğunluklar.....	70
Şekil 7.5. Senaryo1 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	87
Şekil 7.6. Senaryo1 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	87
Şekil 7.7. Senaryo2 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	87
Şekil 7.8. Senaryo2 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	88
Şekil 7.9. Senaryo 3 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	88
Şekil 7.10. Senaryo 3 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	88
Şekil 7.11. Senaryo 4 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	89

Şekil	Sayfa
Şekil 7.12. Senaryo 4 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	89
Şekil 7.13. Senaryo 5 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	89
Şekil 7.14. Senaryo 5 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	90
Şekil 7.15. Senaryo 6 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	90
Şekil 7.16. Senaryo 6 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	90
Şekil 7.17. Senaryo 7 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	91
Şekil 7.18. Senaryo 7 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	91
Şekil 7.19. Senaryo 8 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	91
Şekil 7.20. Senaryo 8 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	92
Şekil 7.21. Senaryo 9 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	92
Şekil 7.22. Senaryo 9 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	92
Şekil 7.23. Senaryo 10 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	93
Şekil 7.24. Senaryo 10 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	93
Şekil 7.25. Senaryo 11 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	93
Şekil 7.26. Senaryo 11 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	94
Şekil 7.27. Senaryo 12 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi.....	94
Şekil 7.28. Senaryo 12 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi.....	94

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar

Açıklamalar

EKK

En Küçük Kareler

HKO

Hata Kareler Ortalaması

LAD

En Küçük Mutlak Sapmalar

LMS

En Küçük Medyan Kareler

LTS

En Küçük Budanmış Kareler

MAD

Mutlak Sapmaların Medyanı

1. GİRİŞ

Regresyon analizinde yaygın olarak kullanılan En Küçük Kareler yöntemi Klasik Doğrusal Regresyon Modelinin varsayımları sağlandığında en iyi sonuçları veren tahmin yöntemidir. Bu analizde hata terimlerine ilişkin birtakım varsayımlar söz konusudur. Bunlar; hata terimlerinin 0 ortalama ve σ^2 sabit varyansına sahip olması ve hem birbirleriyle hem de bağımsız değişkenler ile ilişkisiz olmasıdır. Veride aykırı değerlerin bulunması, örnek çapı artırılrsa bile değişen varyanslılığa neden olacaktır. Böylece varsayımlardan sapmalar meydana gelecek ve En Küçük Kareler güvenilir bir yöntem olmaktan uzaklaşacaktır.

Geniş uygulama alanlarında, toplanan veriler sıklıkla aykırı değer olarak adlandırılan olağan olmayan değerler içerir. Bunlar verinin tamamından veya büyük çoğunluğundan kolayca ayrılırlar. Örnek ortalaması, örnek varyansı, örnek kovaryansı ve korelasyonu gibi klasik tahminler sadece bir aykırı değerden bile olumsuz etkilenirler ve veriye uygun modeli sağlamada başarısız olurlar. Bu nedenle veri aykırı değer içerdiğinde iyi bir uyum sağlayan sağlam parametre tahminleri önerilmiştir (Maronna, Martin ve Yohai, 2006: 1).

Sağlam yöntemler, verinin karakteristik yapısından kaynaklanan veya uygulama sırasında meydana gelen problemlere karşı geliştirilmiştir. Bu problemler varsayımlardan sapmalara neden olacağından sağlam yöntemler tahminlerin olumsuz olarak etkilenmesini engeller ve uygulayıcıya daha sağlıklı sonuçlar sunar. Bu tahminler verideki aykırı değerlerin varlığına dirençlidirler ve geçerli bir tahmin doğrusu elde ederler. Aynı zamanda uygulamada normal dağılan verilerle karşılaşma olasılığı düşük olduğundan normallik varsayımının sağlanamaması durumuna karşı birçok alanda güven duyulan yöntemlerdir.

Başta finans olmak üzere birçok alanda yapılan çalışmalar, gerçek hayatta verilerin sıklıkla normal dağılmadığını göstermektedir. Çarpık ve kalın kuyruklu yapısıyla kararlı dağılımlar son zamanlarda oldukça ilgi görmektedir. İktisatta, servet dağılımını göstermek için Pareto dağılımı olarak, fizikte Lorenz dağılımı olarak karşımıza çıkan kararlı dağılımların aykırı değerlerin varlığında iyi performansla sahip olduğu birçok çalışmada gösterilmiştir. Gözlemde aykırı değerlerin varlığı hataların kalın kuyruklu olmasına neden olur ve En Küçük Kareler Yöntemi bu durumdan bir hayli etkilenir. Bu nedenle normallik varsayımının sağlanamaması durumunda kararlı dağılımları kullanmak geçerli bir

alternatiftir. Hata terimleri, alfa kararlı dağıldığında karşılaşılabilecek problem, kararlı dağılımların Normal, Cauchy ve Levy dağılımları dışında yoğunluklarının kapalı bir formunun olmayışdır. Ancak günümüzde mevcut bilgisayar programları ile bu sorunun üstesinden rahatça gelinebilmektedir.

Regresyon terimi ilk kez Francis Galton (1886) tarafından çok uzun ya da kısa ebeveynlerin çocuklarının çok uzun ya da kısa olmadığını, nüfusun ortalama boyuna yakın olma eğiliminin olduğunu gösterdiği çalışmasında kullanılmıştır (Gujarati, 1995/1999: 15).

Regresyon analizinde en yaygın kullanılan metot En Küçük Kareler (EKK) metodudur. 1775'te Almanya'da Carl Friedric Gauss ve 1805'te Fransa'da Adrien Marie Legendre tarafından birbirlerinden etkilenmeden keşfedilmiştir. Metodun ilk uygulamaları astronomik ve jeodezik veriler üzerine olmuştur (Birkes ve Dodge, 1993: 29).

Regresyon için mevcut olan varsayımlar sağlandığında EKK yöntemi yansız, minimum varyanslı en iyi tahminleri verir. EKK yöntemi hata kareleri kullandığından büyük hata değerine sahip gözlem değerinin hata karesi daha büyük olacak ve regresyon doğrusu üzerinde çok daha büyük bir etkiye sahip olacaktır.

Sağlamlık, istatistik alanında şüphe edilemez bir değişim olmuştur. Bu tanımın kullanımı, Box'un 1953'te Biometrika'da yayınlanan "Non-Normality and Tests on Variance" makalesine dayanır. Burada Box, normal olmayan durumlarda sağlamlığın cazip özelliklere sahip olduğunu yazmıştır. Fakat bu terimi bize Box tanıtırsa bazıları istatistikte bir alt disiplin olan sağlamlığı Tukey ve Huber'in bir buluşu olarak kabul eder (Stigler, 2010).

Sağlam yöntemler 19. yy'ın sonuna kadar uzanan uzun bir tarihe sahiptir. Fakat ilk büyük adımlar 1960'lı ve 1970'li yılların başında J.Tukey (1960, 1962), P.Huber (1964, 1967) ve F.Hampel (1971, 1974)'in temel çalışmalarıyla atılmıştır. Huber (1981), Hampel, Ronchetti, Rousseeuw ve Shantel (1986), Rousseeuw ve Leroy (1987) ve Staudte ve Sheather (1990) tarafından güçlü eserler yazılmıştır (Maronna ve diğerleri, 2006: xvii).

Aslında sağlam yöntemler üzerine çalışmalar 1757'de R. Joseph Boscovich tarafından En Küçük Mutlak Sapmalar (LAD) yönteminin bulunmasıyla başlamış fakat uzun

hesaplamalar gerektirmesinden dolayı En Küçük Kareler yöntemi popüler olmuştur. Bilgisayar programcılığının gelişmesi üzerine sağlam yöntemler kaybettiği ilgiyi tekrar kazanmış, Boscovich'in temelini attığı L_1 regresyon olarak da geçen LAD metodu Laplace ve Edgeworth tarafından tekrar ele alınarak günümüzde tercih edilen bir yöntem olmuştur. L_1 tipi regresyon daha detaylı olarak Sielkan ve Hartley (1973), Book ve diğerleri (1980), Gertle Kennedy ve Spotso (1977), Bloomfield ve Steiger (1983) ve Dodge (1987) tarafından ele alınmıştır.

Rousseeuw (1984), En Küçük Medyan Kareler (LMS) ve En Küçük Budanmış Kareler (LTS) yöntemlerini içeren birkaç sağlam metot sunmuştur. Rousseeuw ve Van Driessen (1999), LTS için hızlı bir algoritma sunmuş ve son olarak Rousseeuw ve Van Driessen (2006) yayınlamışlardır. Rousseeuw (1984), ağırlıkları LMS tahminine dayanan yeniden ağırlıklandırılmış en küçük kareler yöntemini sunmuştur. Rousseeuw ve Hubert (1997), buna eşdeğer yeni bir tanımda bulunurken Rousseeuw ve Leroy (1987), yöntemi daha da geliştirmişlerdir (Doornik, 2011).

Son yirmi yıldan daha fazla süredir sağlam tekniklere ilgi oldukça büyümüştür. Sağlam tahminlerin temel amacı büyük hatalarla ilişkili ölçümlerin etkisini azaltmaktır. Aykırı değerlere karşı koymak için kurulan sağlam yöntemlerden biri de M tahminidir. Bu yöntem, iyi istatistiksel özellikleri ve hesaplamada daha az çaba gerektirmesinden dolayı literatürde önemli bir role sahiptir. Huber'in M tahmini en popüler olanıdır (Guo, 2003).

Huber (1964), hata terimi normal dağılımdan başka keyfi dağılımlarla kirletildiğinde optimal çözüm için M tahminini sunmuştur. Huber'in optimallik kriteri sonsuz büyük örnekler için maksimum olabilirlik varyansının minimizasyonudur (Birkes ve Dodge, 1993: 99).

Tukey'in Biweight fonksiyonu tahmin yönteminde farklı bir fonksiyon kullanmış ve Huber'in yönteminden daha yaygın hale gelmiştir. Ancak Huber M tahmini, aykırı değerlerin belirlenmesinde daha başarılı olabilmektedir (Hund, Massart ve Verbeke, 2002).

Andrew (1974), M tahminine dalga fonksiyonu olarak da bilinen etki fonksiyonu ile farklı bir yaklaşımda bulunmuştur. Hampel 'in M tahmininde kullandığı fonksiyon ise üç parçalı ve 2-4-8 fonksiyonu adıyla bilinen bir fonksiyondur.

Yohai (1987), yüksek kırılma noktası ve yüksek etkinliğe sahip MM tahminini önermiştir. Yohai bu çalışmasında, MM tahmininin aykırı değerlerin varlığından çok fazla etkilenmediğini göstermiştir. Barrera ve Yohai (2008), sansürlü veride bu tahmini uygulamışlardır. Yüksek kırılma noktasını garanti eden, başlangıç S tahmini ile hesaplanan MM tahmini normal etkinliğe izin verir. Tukey'in fonksiyonu hem başlangıç hem de son tahminde yaygın olarak kullanılır (Verardi ve Croux, 2009).

Koç (2007), bitkilerde amonyağın nitrik aside oksidasyonu işlemi için LMS, LTS, LAD, S, M ve Genelleştirilmiş M metodlarını karşılaştırmış ve en iyi kırılma noktasına S tahmininin sahip olduğunu göstermiştir.

Teoride; LTS tahmini x ve y yönündeki aykırı değerlere karşı yüksek kırılma noktasına sahip olmasına rağmen Lin (2006), çalışmasında bu tahminin M tahmininden daha düşük bir performans sergilediğini göstermiştir.

Gürünlü Alma (2011), dört sağlam regresyon metodunu EKK regresyon metoduna karşı değerlendirmiş, simülasyon çalışmasının sonuçları S ve M tahminlerinin, LTS ve MM tahmininden daha iyi performansa sahip olduğunu göstermiştir.

Verardi ve Croux (2009), yapmış oldukları simülasyon çalışmasında EKK, LAD, M ve MM tahminlerini yanlışlık ve hata kareler ortalaması kriterine göre karşılaştırmış, MM tahmininin kirlenmelere karşı daha sağlam olduğunu göstermişlerdir.

Croux ve diğerleri (1994), En küçük Budanmış Çeyrek Farklar olarak bilinen yöntemde artık çiftlerinin farklarının bir fonksiyonunu minimize etmeye çalışmışlardır. Bu yöntemde artık çiftlerinin mutlak değerleri sıralanır en küçük çeyreği minimize eden tahmin sonucu verir. Bu yöntem için modelde kesişim parametresi bulunmaz, hesaplanması isteniyorsa diğer klasik yöntemler kullanılır.

En küçük Budanmış Çeyrek Farklar yöntemine benzer olarak Stromberg ve diğerleri (1995), tarafından sunulan En Küçük Budanmış Farklar, artık çiftlerinin farklarının karelerinin en küçüğünü minimum etmeye çalışmıştır.

Koenker ve Bassett (1978), doğrusal regresyon modelinin bir uzantısı olarak Kantil Regresyonu sunmuşlardır. Bu regresyonda değişkenler arasındaki ilişki kantil fonksiyonları kullanılarak açıklanmaya çalışılır.

Akar (2013), finansal piyasa istikrarını ölçmeye yönelik ampirik bir çalışmada kantil regresyonu kullanmıştır. Kantil regresyon yardımıyla farklı piyasa şoklarının hisse senedi piyasa getirilerine etkisini araştırmıştır.

Efron tarafından 1979 yılında önerilen bootstrap yöntemi, dağılım varsayımlarına gerek duymadığından (hata terimlerinin normal dağılması gibi) örnek çapı küçük olduğunda veya verinin davranışı kötü olduğunda daha doğru çıkarımlar sağlayan bir yeniden örnekleme yöntemidir.

Basawa ve diğerleri (1991), Beran ve Ducharme (1991), Efron ve Tibshirani (1993), Davison ve Hinkley (1997), Horowitz (1998), Horowitz (2003), Davidson ve MacKinnon (2006), Chernick (2008), Godfrey (2009), çalışmaları ile bootstrap yöntemine katkıda bulunmuşlardır. Sufahani ve Ahmad (2012), normal ve normal olmayan veriye uyguladıkları bootstrap sonunda dağılımın neredeyse normal dağıldığını göstermişlerdir.

Aykırı değerlerin varlığında bir tahminin sağlamlığının derecesi Hampel tarafından sunulan kırılma noktası kavramı ile ölçülebilir. Donoho (1982) ve Donoho ve Huber (1983), konunun bu çalışmada kullanılabilecek sınırlı basit bir versiyonunu sunmuşlardır. Bu sınırlı örnekte kırılma noktası ölçümü, tahminin tamamı bozulmaksızın örneğin içerebildiği aykırı değerlerin maksimum kısmıdır.

EKK, finansal veri analizinde sıklıkla kullanılan bir metoddur. Fakat ticari dönemler ve finansal politikalardaki değişimlerden dolayı tüm veri içinde aykırı değerlerin artması kaçınılmazdır. Bundan dolayı finansal veri analizinde sonucun yansızlığını koruyan ve aykırı değerlere karşı zayıf olmayan bir regresyon modelinin kullanılması gereklidir (Momeni, Nayeri, Ghayoumi, Ghorbani, 2010).

Finans ve ekonomide alfa kararlı dağılımların kullanımı için ilk destek 60'lı ve 70'li yıllarda Mandelbrot ve Fama'nın çalışmalarından gelmiştir. Çoğu varlık getirilerinin bu süreçler için tipik özelliklere sahip olduğunu bulmuşlardır. İlk araştırma döneminden sonra

alfa kararlı süreçlere ilginin azaldığı görülmüştür. Bunun temel nedeni hesaplama sorunları iken, bir başka faktör de o dönemde finans dünyasında büyük buluşların olmasıdır. Bunlar büyük ölçüde hataların normal dağıldığı varsayımına dayanan çalışmalar olmuştur. Normal dağılım kullanan bu çalışmaların başarısı Markovich, Millar, Sharpe, Merton ve Scholes 'e portföy dağılımı, sermaye varlıklarını fiyatlama modeli, opsiyon fiyatlama ve yatırım teorisine katkılarından dolayı Nobel ödülünü kazandırmasıdır. Bu çalışmalar kantitatif finans gündemindeki araştırmaların başlıca bileşenleri olmuştur. O zamanlar sadece bu çalışmalar için değil tüm ekonometrik çalışmalarda alfa kararlı dağılımlarının geçersiz olacağı gibi yanlış bir düşünce mevcuttu. Son zamanlarda bilgisayar hesaplamalarındaki gelişmeler α kararlı dağılımlara ilgiyi artırmıştır ve bu trendin devam etmesi muhtemeldir (Frain, 2006).

Finans ve ekonomi alanındaki örnekler Mandelbort (1963), Fama (1965), Samuelsen (1967), Roll (1970), Embrechts ve diğerleri (1997), Rachev ve Mitnik (2000), McCulloch (1996), iletişim sistemlerinde Stuck ve Kleiner (1974), Zolotarev (1986) ve Nikias ve Shao (1995) tarafından verilmiştir. Normal dağılan bir modelle tanımlanması başarısız olan veri setlerinde alfa kararlı dağılımları tanımlamak iyi olabilir. Kararlı modellere odaklanan birkaç monografi vardır: Zolotarev (1986), Uchaikin ve Zolotarev (1999), Christoph ve Wolf (1992), Samorodnitsky ve Taqqu (1994), Janicki ve Weron (1994), Nikias ve Shao (1995). Aykırı değerlerle ve kalın kuyruklu dağılımlarla veri modelleme Embrechts ve diğerleri (1997), Adler ve diğerleri (1998) ve Reiss ve Thomas (2001) tarafından tartışılmıştır (Nolan, 2009).

Birçok finansal değişkenin normal dağılımdan daha kalın kuyruklu dağılımlarla çok daha iyi modellenebileceği deneysel bir gerçektir. Mikosch ve Vries (2013) hata terimlerinin Pareto dağıldığı bir doğrusal regresyon modeli için elde ettikleri sonuçların ekonomik verilere de uygulanabileceğini göstermişlerdir (Mikosch ve Vries, 2013).

Kararlı dağılımların en mükemmel özelliği normal dağılımla karşılaştırıldığında daha kalın kuyruklu ve tepeli olmasıdır. Fama ve Blume 'nin çalışmalarından sonra kararlı dağılımların özellikleri ve parametrelerinin tahmini Fama ve Roll tarafından ayrıntılı olarak çalışılmıştır (Officer, 1972).

Mandelbort, pamuk, buğday, demiryolu gelirleri, hisse ve döviz oranlarının getirilerinin değişimine çalışmış ve normal dağılım varsayımının savunabileceğinden çok daha fazla aykırı değer bulmuştur. Mandelbort fiyat farklılıkları için kararlı dağılımın uygun bir model olacağını öne sürmüştür. Fama 1957'nin sonundan 26 Aralık 1962 arasındaki periyotta Dow Jones Endüstriyel ortalama indeksinde 30 hissenin günlük getirilerini incelemiştir (M.A.B Ghahfarokhi ve P.B. Ghahfarokhi, 2009).

Frain (2006), günlerin hisse senetlerinin günlük getirilerine etkisini araştırdığı çalışmada, alfa kararlı dağılan artıklar için daha iyi bir uyuma sahip olduğunu göstermiştir.

Bu tezin amacı; hata terimlerinin alfa kararlı dağılıma sahip olduğu basit regresyon modeline bazı sağlam tahmin yöntemleri uygulamak ve performanslarını karşılaştırmaktır. Buradan hareketle alfa kararlı dağılımın farklı yoğunluklarında hangi tahmin yönteminin daha iyi olduğu konusunda araştırmacılara ön bilgi sunmaktır. Çalışmanın 2. bölümünde klasik istatistik yöntemlerinden biri olan regresyon analizi ve bu analizde en yaygın kullanıma sahip En Küçük Kareler (EKK) yönteminden bahsedilmiştir. 3. bölümde regresyon analizinin varsayımlarından sapmalara neden olan aykırı değerler, aykırı değerlerin çeşitleri ve tahmin sonuçlarına etkileri konu edilmiştir. 4. bölümde çalışmamızın asıl amacı olan sağlam yöntemler, bu yöntemlere duyulan ihtiyaç ve en çok tercih edilen bazı sağlam tahmin yöntemleri anlatılmıştır. En Küçük Budanmış Kareler (LTS), M tipi tahmin ediciler, S tahmini, MM Tahmini, En Küçük Medyan Kareler (LMS), En Küçük Mutlak Sapmalar (LAD), Kantil Regresyon ve Bootstrap yöntemi bu başlık altında incelenen metotlardır. 5. bölümde alfa kararlı dağılımlar tanıtılmış, teorik özellikleri ile incelenmiştir. 6. bölümde kararlı regresyon modeli tanıtılmıştır. 7. bölümde hata terimlerinin 12 farklı alfa kararlı dağılıma sahip olduğu basit doğrusal regresyon modeline dördüncü bölümde anlatılan sağlam yöntemler uygulanmış ve performansları karşılaştırılmıştır.

2. REGRESYON ANALİZİ

Regresyon analizi, iki ya da daha fazla değişken arasındaki ilişkiyi açıklamak için bir denklem geliştiren istatistiksel bir tekniktir. Bağımsız değişkendeki değişiklikler bağımlı değişkene aktarılır. Genel olarak bu değişimden bağımlı değişkenin nasıl etkilendiği ile ilgilenilir (Dielman, 1991: 2, 62).

Bu analizin amacı, bağımsız değişkenin bilinen ya da değişmeyen değerlerine dayanarak bağımlı değişkenin ortalama değerini kestirmektir (Gujarati, 1995/1999: 16).

Regresyon öncelikle aşağıdaki sorunlarla ilgilenir:

1. Değişkenler arasındaki ilişkinin şekli ve temsil eden denklemin neye benzediği.
2. İlişkilerin gücü ve yönü. Bunlar eğim katsayılarına bağlı olduğundan bu katsayılar bilinmek zorundadır.
3. Hangi değişkenin önemli olup olmadığı. Bu sorun eğim katsayılarının değerlerini karşılaştırmaya dayanır (Hoffmann, 2010).

$$y = \beta_0 + \beta_1 x + e \quad (2.1)$$

eşitliği bir basit doğrusal regresyon modeli olarak adlandırılır.

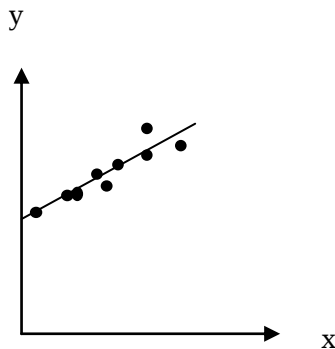
y : Bağımlı değişken

x : Bağımsız değişken

β_0 : Kesişim parametresi

β_1 : Eğim parametresi

e : Rasgele hata terimidir.



Şekil 2.1. Doğrusal regresyon modelinin grafik gösterimi

β_0 ve β_1 'e regresyon katsayıları denir. β_1 eğim parametresi; x ' teki bir birimlik değişimin y 'de meydana getirdiği ortalama değişimdir. $x=0$ olduğunda y bağımlı değişkeninin ortalaması β_0 kesişim parametresini verir. Eş. 2.1 sadece bir bağımsız değişken içerdiğinden “Basit Doğrusal Regresyon Modeli” olarak adlandırılır (Montgomery, Peck ve Wining, 2001: 3).

Parametreler bir modeli karakterize eden bilinmeyen katsayılardır. Parametrelerin tahmini verilerin bir fonksiyonudur ve “istatistik” adını alır.

i.durum için uydurulan değer $\hat{E}(Y/X=x_i)$ tarafından ifade edilir ve $\hat{y}_i$ notasyonu kullanılır.

$$\hat{y}_i = \hat{E}(Y/X=x_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (2.2)$$

i.durum için hata değeri;

$$\hat{e}_i = y_i - \hat{E}(Y/X=x_i) = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i) \quad i=1, \dots, n \quad \text{eşitliği ile verilir.} \quad (2.3)$$

Bu durumda

$$y_i = (\hat{\beta}_0 + \hat{\beta}_1 x_i) + \hat{e}_i \quad (2.4)$$

şeklinde de ifade edilir (Weisberg, 2005: 21).

Regresyon analizinin varsayımları

Regresyon analizi, tahmin denklemini geliştirmek için birtakım varsayımlar gerektirir.

Bunlar genelde regresyon modelindeki hata terimleri e_i 'ler açısından ifade edilmektedir.

Bu varsayımlar;

1. Hata terimlerinin beklenen değeri 0, varyansı σ^2 'dir.

$$E(e_i) = 0$$

$$V(e_i) = E[e_i - E(e_i)]^2 = E[e_i^2] = \sigma^2 \quad (2.5)$$

2. Hata terimleri birbirleri ile ilişkisizdir. Herhangi bir hata çifti (e_i, e_j) birbirine bağlı olmamalıdır. Bunun anlamı e_i hata teriminin değeri biliniyorsa, bunun e_j hata terimi hakkında bize herhangi bir şey söylememesi gerektiğidir. Kısaca

$$\text{cov}(e_i, e_j) = 0 \text{ 'dır. } (i \neq j) \quad (2.6)$$

Bu varsayımın veri toplamada ihlal edilmesi ciddi sorunlara sebebiyet verir. Bu durumda “otokorelasyon” sorunu ile karşı karşıya kalınır. Ardışık bağımlılık olarak da adlandırılan bu sorun iktisadi zaman serilerinin süre durumu ya da ağır hareketlilikleri, önemli bir değişkenin modele dahil edilmemesi, model kurma hatası ve verilerle oynanması durumunda ortaya çıkabilir. Bu sorunun varlığında tahmin ediciler doğrusal sapmasız iken etkin olma özelliklerini yitirirler. Bu sorunun varlığı artık grafikleri ile tespit edilebildiği gibi en yaygın kullanılan yöntem Durbin-Watson testidir.

3. Hata terimleri aynı zamanda X açıklayıcı değişkenleri ile de ilişkisiz olmalıdır. Yani benzer olarak e_i 'nin değeri X 'e bağlı olmamalıdır.

$$\text{cov}(e_i, x_j) = 0 \text{ 'dır.} \quad (2.7)$$

Hipotez testleri ve güven aralıkları için önemli olan bir varsayım hata terimlerinin $e_i \sim N(0, \sigma^2)$ normal dağılımına sahip rasgele değişkenler olduğudur. Çok sayıda bağımsız ve özdeş dağılan rasgele değişken varsa, bu değişkenlerin sayısı sonsuza doğru arttıkça bunların toplam dağılımı birkaç aykırılık dışında normal dağılıma yaklaşır. İşte bu Merkezi Limit Teoreminin hata terimlerinin normal dağıldığı varsayımının gerekçesi olur. Bu varsayımın sağlanamaması en çok aykırı değer ve etkili gözlem varlığında ortaya çıkar (Gujarati, 1995/1999: 59-66; Ryan, 2009: 15,16).

Merkezi Limit Teoremi çok geniş koşullar altında EKK tahminine dayanan sonuçların yaklaşık olarak tüm küçük örneklerde geçerli olduğunu ortaya koyar. EKK tahminleri hata terimleri normal dağıldığı zaman yansız tahminler arasında en yüksek etkinliğe sahip tahmin edicidir. Hata dağılımlarının bazı türleri için özellikle kalın kuyruklu dağılımlarda EKK tahmininin etkinliği önemli derecede azalır. Bu durumda EKK tahmini alternatiflerinden daha düşük bir etkinliğe sahip olacaktır. Kalın kuyruklu dağılımlar aykırı değerlere sebebiyet verdiğinden önemli ölçüde problemdir. EKK'nın tahmininin ispatında

yaygın olarak kullanılan Gauss Markov Teoremi, y_i gözlemlerinin doğrusal fonksiyonlarının yansız tahmin edicilerinde EKK tahmininin en etkin katsayılarla sahip olduğunu ifade eder. Bu sonuç doğrusallık, sabit hata varyansı ve bağımsızlık varsayımlarına bağlıdır fakat normallik gerekli değildir (Fox, 2008: 268).

2.1. En Küçük Kareler Yöntemi

EKK, veri setine ait bir fonksiyon uydurmak, parametrelerin sayısal değerlerini tahmin etmek ve tahminlerin istatistiksel özelliklerini karakterize etmek için yaygın olarak kullanılır. EKK'nın birkaç çeşidi vardır. Bunlardan en basiti sıradan en küçük kareler, daha özelleştirilmiş hali ise her bir gözlemin etkisini ayarlayabildiğinden sıradan EKK 'ya göre daha iyi performans gösteren ağırlıklandırılmış en küçük kareler yöntemidir. En Küçük Kareler metodunun yakın zamanda sunulan çeşitleri ise alternatif en küçük kareler ve kısmi en küçük karelerdir (Abdi, 2010: 705-708).

β_0 ve β_1 'i tahmin ederken sıklıkla En Küçük Kareler Yöntemi kullanılır. Bu yöntemde y_i gözlem değerleri ile regresyon doğrusu arasındaki farkın karelerinin minimum olması istenir. Regresyon analizi için tüm varsayımlar sağlandığında EKK tahmin yöntemi en iyi uyum gösteren minimum varyanslı, yansız doğrusal tahmin edicidir (Sundberg, 2002).

Eş. 2.1 şöyle yazılabilir.

$$y_i = \beta_0 + \beta_1 x_i + e_i \quad (2.8)$$

Eş. 2.1 yığın regresyon modeli olarak, Eş. 2.8 ise n tane (x_i, y_i) gözlem çiftinden oluşan örnek regresyon modeli olarak görülebilir. Bu durumda EKK kriteri;

$$S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.9)$$

Bu eşitliğin sırasıyla β_0 ve β_1 'e göre türevlenirse;

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i \quad (2.10)$$

eşitlikleri elde edilir.

Eş. 2.10 en küçük kareler normal eşitlikleri olarak adlandırılır. Bu normal eşitliklerinin çözümü

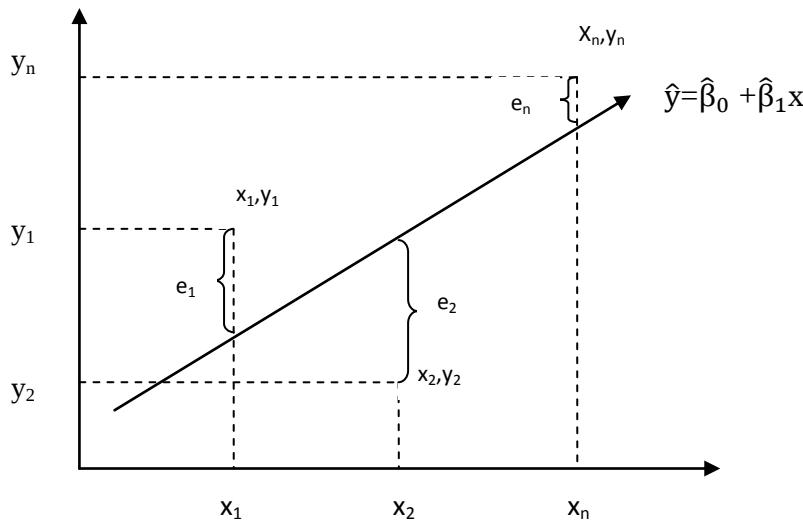
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (2.11)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i)}{n}}{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}} \quad (2.12)$$

Burada;

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{ve} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad y_i \text{ ve } x_i \text{'nin ortalamasıdır.}$$

Eş. 2.11 ve Eş. 2.12 sırasıyla kesişim ve eğimin EKK tahminleridir. (Montgomery ve diğerleri, 2001: 13, 14, 15).



Şekil 2.2. Tahmin edilen regresyon doğrusu

İstatistiksel bakımdan model kurulduktan sonra o modelin geçerliliğinin araştırılması regresyon analizinin önemli bir parçasıdır. Kestirilen modelin gerçek modele ne kadar yaklaştığını belirleyebilmek için, kullanılan EKK yönteminin regresyon varsayımlarını

sağlayıp sağlamadığı kontrol edilmelidir. Kurulan model veriye uygun değilse alınan sonuçlar da yanıltıcı olacaktır (Gürünlü Alma ve Vupa, 2008).

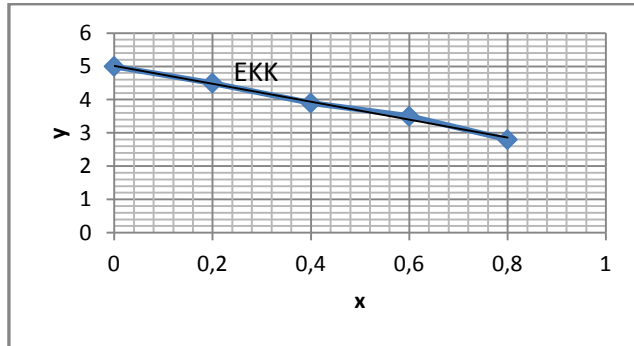
3. AYKIRI DEĞERLER

Aykırı gözlemler verinin çoğunluğundan belirgin bir şekilde sapan değerlerdir. Bu değerler verinin çoğunluğunun uyduğu mekanizmadan farklı bir şekilde oluşmuş kayıt hatası, üretim sırasında bozukluk ya da insandan kaynaklanan sebeplerle meydana gelebilir. Aykırı gözlemler, model kurma hatası, yanlış parametre tahmini veya hatalı analiz sonuçlarına neden olur (Liu, Shah ve Jiang, 2004).

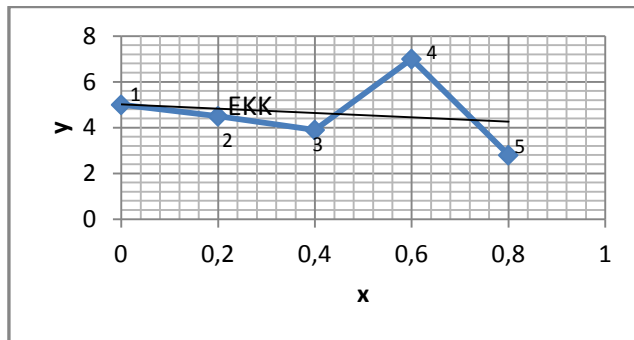
Regresyonda iki tip aykırı değer vardır:

Dikey aykırı değer: Bunlar bağımlı değişkenlerin doğrusal eğiliminden aykırı olan değerlerdir.

Kaldıraç noktaları: Bunlar bağımsız değişken uzayında aykırı olan değerlerdir (Gschwandtner ve Filzmoder, 2012).

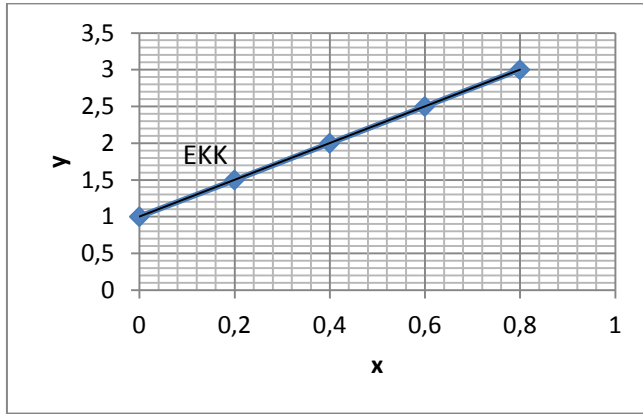


Şekil 3.1. Beş noktalı orijinal veri kümesi ve EKK regresyon doğrusu

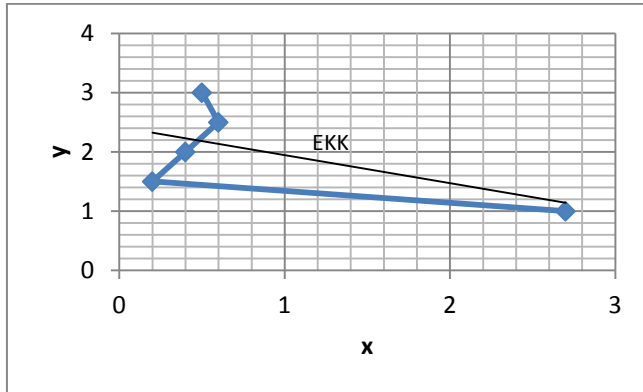


Şekil 3.2. Beş noktalı y-yönünde aykırı gözlem içeren veri kümesi

Şekil 3.1. Bir doğru üzerinde uzanan 5 noktanın dağılım grafiğidir. EKK çözümü veriyi çok iyi uydurur. y_4 noktasının bir değerinin yanlış olduğunu varsayalım. O zaman (x_4, y_4) ideal doğrudan uzak olabilir. Şekil 3.2.'de 4. noktanın doğruyu orijinal durumundan yukarıya taşıdığı bir durum göstermektedir. Bu nokta y- yönünde aykırı değer olarak adlandırılır ve regresyon doğrusuna büyük bir etkisi vardır. Dikey aykırı değerler sıklıkla çok büyük pozitif veya negatif hatalara sahiptirler.

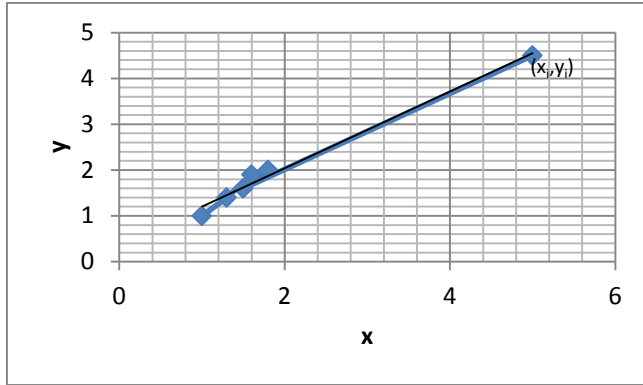


Şekil 3.3. Beş noktalı orijinal veri kümesi ve EKK regresyon doğrusu



Şekil 3.4. Beş noktalı x-yönünde aykırı gözlem içeren veri kümesi

Şekil 3.3. iyi uydurulan EKK doğrusuyla 5 nokta içerir. Birinci noktada bir kayıt hatası yapılırsa Şekil 3.4. elde edilir. Çıkan nokta x-yönünde bir aykırı değer olarak adlandırılır ve EKK doğrusunu yana yatırdığından EKK tahmini üzerinde çok büyük etkisi vardır. Bu nokta kaldıraç noktası olarak adlandırılır. Eğer (x_i, y_i) noktası verinin çoğu tarafından belirlenen regresyon doğrusuna yakın ise “iyi kaldıraç noktası” olarak adlandırılır (Rousseeuw ve Leroy, 1987: 4, 5, 6).



Şekil 3.5. (x_i, y_i) ; iyi kaldıraç noktası.

Aykırı değerlere karşı genel olarak yaklaşım şu şekildedir;

1. Bazı aykırı değerleri çıkarma kurallarıyla veri temizlenir.
2. Kalan gözlemlere klasik tahmin ve test yöntemler uygulanır.

Fakat bu yaklaşım bazı nedenlerden dolayı uygun değildir.

1. Bu iki adımı ayırabilmek nadiren olasıdır. Çok değişkenli regresyon problemlerinde, parametreler için sağlam yöntemler kullanılmadığında aykırı değerleri ayırt etmek zordur.
2. Bazı büyük hataların karıştığı gözlem kümesi normal olmasına rağmen temizlenmiş veri normal dağılmayabilir. Orijinal gözlem kümesi, büyük hata içeren yapı yerine gerçekten normal olmayan bir dağılıma sahipse, durum daha da kötüleşir.
3. En iyi gözlem çıkarma yönteminin performansının, iyi bir sağlam tahmin yönteminin performansına ulaşamadığı deneysel bir gerçektir.
4. Küçük örneklerde çıkarma işlemi sakıncalıdır (Huber, 1981: 4, 5).

Basit regresyon için veri tek değişkenlidir ve dağılım grafiği ile gösterilebilir bu yüzden görsel inceleme ile aykırı değerlere kolayca karar verilebilir. Fakat birkaç açıklayıcı değişkene sahip veri için bu kararı vermek çok zordur. Bu yüzden doğrusal modelden sapan noktaları sağlam regresyona dayanan hatalar belirler. Ayrıca kaldıraç noktalar Rousseeuw ve Zomeren (1990)'da bahsettiği gibi $x_1, x_2, \dots, x_n$ noktalarının uzaklıklarının hesaplanmasıyla belirlenir (Rousseeuw ve Van Driessen, 2006).

Eğer bir gözlem verilerin ortalamasından çok farklı ise ilgili değişken etkili veridir. Bazen bir ya da birkaç veri regresyon modelinin tahmin edilen parametreleri üzerinde etkiye sahip

olabilirler. Diğer bir deyişle etkili veri, kaldırıldığında modelde ciddi deęişmelere neden olan veridir. Her gözlemin çıkarılması regresyon modelinde deęişimlere neden olacaktır ancak etkili gözlemin çıkarılması önemli deęişikliklere sebebiyet verir (Momeni, Nayeri, Ghayoumi ve Ghorbani, 2010).

4. SAĞLAM YÖNTEMLER

İstatistiksel modelin dayandığı varsayımların sağlanmaması durumunda veya veri setinde aykırı değerler bulunması durumunda EKK tahmini tatmin edici sonuçlar vermemektedir. Bu gibi durumlarda, aykırı değerden veya modelden sapmalardan etkilenmeyen sağlam tahminler tercih edilir. Sağlam tahminler verideki aykırı değerlerden ve varsayılan modelden sapmalardan etkilenmeyen tahminlerdir (Altın ve Şenoğlu, 2008).

4.1. Sağlam Yöntemlere Neden İhtiyaç Duyulur?

Sağlam regresyon, hata terimlerinin normal dağılmaması durumunda ve/veya regresyon eşitliğini etkileyen aykırı değerlerin varlığında EKK yöntemine alternatif bir yöntem olarak kullanılabilir. Sağlam regresyon ile ilgili çalışmaların çoğu Andrews, Bickel, Hampel, Huber, Rogers'in 1972'de hata terimleri normal dağılımdan daha kalın kuyruklu dağıldığı zaman EKK tahmininin diğer tahminlere göre iyi olmayan sonuçlar verdiğini gösterdikleri "Princeton Robustness Study" adlı çalışmaları tarafından harekete geçirilmiştir. EKK yaklaşımı aynı zamanda veri seti aykırı değerler içerdiğinde de güvensiz tahminler veren bir yöntemdir. Hemen hemen tüm veri setleri kötü veri içerir ve bu belli uygulama alanlarında süregelen bir problem olabilir. Bu nedenle sağlam tahmin yöntemlerine ihtiyaç duyulur (Ryan, 2009: 421).

Tüm istatistiksel metotlar birtakım varsayımlara dayanır. Bu varsayımlar genellikle istatistiksel modelleme problemi veya veri analizi hakkında istatistikçinin bildiklerini veya tahmin ettiklerini şekillendirmeyi amaçlar. Aynı zamanda farklı bakış açılarıyla (teorik ve uygulamada) çıkan modelleri kullanışlı hale getirmektedir. Model kurmada gözlenen verilerin normal dağılıma sahip olduğu en yaygın kullanılan varsayımdır. Bu varsayım iki yüzyıldır istatistikte mevcuttur ve varyans analizi, çok değişkenli analiz, regresyon gibi tüm klasik metotlarda esastır. Normallik varsayımını kabul ettirmek için Merkezi Limit Teoremi gibi teorik argümanlar ortaya atılmıştır. Normal dağılım için varsayımın ana gerekçesi veri setini yaklaşık olarak temsil etmesidir. Aynı zamanda en çok olabilirlik tahmincileri gibi uygun istatistiksel metotlar için açık formüller sağlamasından dolayı teorik olarak da oldukça uygundur. Tek bir aykırı gözlem bile normallik ve doğrusallık

varsayımı altında en iyi olan klasik metotlar üzerinde büyük bozucu etkilere sahip olabilir (Maronna ve diğerleri, 2006: xv, xvi).

Normallik varsayımı, veri analizlerinde kullanılan istatistiksel yöntemler için temel ilkedir. Fakat çeşitli çalışmalar bu durumun verilerin sadece %10-15 'i için doğru olduğunu göstermektedir. Bu durum hataların normal dağılmaması veya gözlemdeki aykırı değerlerin etkisiyle ilişkilendirilebilir. Regresyon analizinde yaygın olarak EKK yöntemi kullanılır. EKK; aykırı değerlere karşı zayıftır ve bu yöntemde aykırı değerlerden etkilenen sonuç kullanıcıyı yanlış yönlendirecek, yanıltacaktır. Aykırı gözlemlerle ilgili istatistiksel modeller için iki düşünce vardır. Birincisi aykırı değerleri hesaba dâhil eder, ikincisi ise eler. Birçok bilim adamı istatistiksel analizde aykırı değerlerin çıkarılmasına gerek duymayan sağlam istatistiklere güvenir (Momeni ve diğerleri, 2010).

Bir istatistiksel yöntem, istatistiksel modelin varsayımları sağlanmadığında bile iyi performans gösteriyorsa “sağlam” olarak kabul edilir. Verimizin doğrusal regresyon modeline uyduğunu varsayarsak, o zaman EKK tahminleri ve test performansı oldukça iyi olacak, fakat rasgele hataların normalliği varsayımı geçersiz olduğunda ise “sağlam” olmayacaktır (Birkes ve Dodge, 1993: 85).

Sağlam bir tahmin edicinin sağlaması istenen bazı özellikleri vardır,

1. Bu tahmin ediciler, küçük bozulmalara küçük tepkiler vermeli
2. Büyük bozulmaların varlığında bile güvenilir olmalı
3. Veride açıkça görülen aykırı değerleri ayırt edebilmeli
4. Verinin büyük çoğunluğunu en iyi şekilde temsil etmeli
5. Olabildiğince küçük varyansa sahip olmalı
6. Yuvarlama ve gruplamaya karşı düzgün tepkiler vermeli
7. Hesaplaması kolay ve anlaşılır olmalıdır.

(Hampel, 1974; Ryan, 2009: 430).

Sağlam yöntemin amaçları ise şöyle özetlenebilir;

1. Verinin tamamını en iyi tanımlayan yapıyı oluşturmak.
2. Aykırı değerleri tanımlamak veya isteniyorsa altyapıdaki sapmaları iyileştirmek.
3. Yüksek etkili değerleri belirlemek ve bunlar hakkında uyarmak.

4. Ardışık bağımlılık veya daha genel olarak varsayılan korelasyon yapısından sapmalar ile ilgilenmektir (Hampel, Ronchetti, Rousseeuw, Stahel, 1986: 11).

4.2. Sağlam Regresyon

Uygulamada regresyon varsayımlarından sıklıkla sapmalar meydana gelir. Bu sapmalar ciddi ise hataların davranışları incelenmek ve modele veya değişkenlere uygun ayarlamalar yapılmak istenir. Varsayımlardan sapmaların bazı türlerinin uygulamada medyana gelmesi diğerlerine göre daha olasıdır. Örneğin; hataların dağılımı simetrik fakat normal olmayabilir yani normalden daha tepeli ve hafif kuyruklu veya normalden düşük tepeli ve kalın kuyruklu olabilir veya dağılım normal dağılsa bile aykırı değerler meydana gelebilir. Bu ve diğer türlü eksikliklerle mücadele etmek için EKK yönteminin yerine sağlam regresyon yöntemleri kullanılır. Bu metotların varsayımdan sapmalara EKK'dan daha az duyarlı olduğuna dair üstünlüklerinin olduğu iddia edilir (Draper, 1966: 342).

“Sağlam” kelimesinin çağrıştırdığı birçok anlam vardır. Sağlamlık varsayımlardan küçük sapmalara duyarsızlıktır. Bir istatistiksel yöntemin tahmin değeri örnekteki değişimlerden etkilenmiyorsa “sağlamdır” denir (Huber, 1981: 1,7).

Uygulamada sağlam tahminler için aykırı değerlere farklı yaklaşımlar vardır. Kimi araştırmacılar aykırı değerlerin çıkarılmasının mümkün olduğu budama yöntemlerini tercih ederken, kimi araştırmacılar aykırı değerlere ait gözlemlerin hata değerlerine ağırlık vererek sorunu gidermeye çalışan ağırlıklandırma yöntemlerini tercih ederler.

4.3. Sağlam Yöntemin Özellikleri

4.3.1. Kırılma noktası (Breakdown point)

Sonlu bir örnek için kırılma noktası kavramı ilk kez yığın ortalaması tahmininde 1967'de Hodges tarafından sunulmuştur ve tek boyutlu durumlarda sınırlı bir tanım olarak Donoho ve Huber tarafından verilmiştir. Bu tanım sağlam regresyon literatüründe çok kullanılmaktadır. Bir tahminin kırılma noktası tahmini kullanışsız hale getirebilecek aykırı değerlerin en küçük kısmıdır. Bir tahmin için mümkün olan en küçük kırılma noktası $1/n$ 'dir. Ne yazık ki çoğu tahmin edicinin kırılma noktası bu değere sahiptir. Örneğin,

örnek ortalaması bu kırılma noktasına sahiptir. Hampel, Ronchetti, Rousseeuw ve Stahel (1986), verilerin genelde %10'unun büyük hatalar içerdiğini göstermişlerdir, bu yüzden bir tahmin edicinin kırılma noktasının 0,10 'u aşması istenir. Geçmiş yirmi yıl boyunca 0,50 kırılma noktasına sahip sağlam regresyon tahminlerine vurgu yapılmıştır (Ryan, 2009: 427).

En Küçük Medyan Kareler (LMS) ve En Küçük Budanmış Kareler (LTS) 0,50 gibi yüksek kırılma noktası değerine sahip tahmin edicilere örnektir.

Donoho ve Huber tarafından basit sonlu örneklem versiyonu şöyledir;

Z , n tane gözlemden oluşan bir örneklem ve T bir regresyon tahmin edicisi ise,

$$Z = \{(x_{11}, x_{12}, \dots, x_{1p}, y_1), \dots, (x_{n1}, x_{n2}, \dots, x_{np}, y_n)\}$$

Z örneğine T yöntemi uygulanarak $T(Z) = \hat{\theta}$ şeklinde katsayılar vektörü bulunabilir. Burada, orijinal veri değerlerinden herhangi m tanesinin yerine keyfi bir değer yazılarak elde edilen kirlenilmiş bir Z' düşünelim. Böyle bir kirlenme ile neden olunabilen sapma; $\text{sapma}(m; T, Z)$ ile belirtilir.

$$\text{sapma}(m; T, Z) = \sup \|T(Z') - T(Z)\| \quad (4.1)$$

olarak tanımlanır.

Eş. 4.1 'in değeri sonsuz ise m aykırı değer T üzerinde büyük ve tahmin ediciyi bozucu etkisi olduğu söylenebilir ve kırılma noktası olarak açıklanır. Z sonlu örneğinde T tahmin edicisinin kırılma noktası;

$$\epsilon_n^* = (T, Z) = \min \left\{ \frac{m}{n}, \text{sapma}(m; T, Z) \text{ sonsuz} \right\} \quad (4.2)$$

olur.

EKK ve L_1 tahminleri için $\epsilon_n^*(T; Z) = 1/n$ 'dir. $n \rightarrow \infty$ iken bu durum 0'a yaklaşma eğilimi gösterdiğinden EKK %0 kırılma noktasına sahiptir ve aykırı değerlerden oldukça etkilenir (Rousseeuw ve Leroy, 1987: 9,10).

Sonuç olarak kırılma noktası bir tahmin edicinin anlamsız sonuçlar vermeden kullanılabileceği, gerçekte kötü kirlenmenin ne kadar olabileceği hakkında bir fikir verir. Kırılma noktası tahmin edicinin kararlılığının global bir ölçüsüdür. Bundan dolayı, T tahmin edicisi tamamen anlamsız hale gelmeden önce, kirlenmiş modelin varsayılan modelden ne kadar uzaklaşabileceğini gösterir. Bununla birlikte kırılma noktası modele bağlı olmaması anlamında da global bir ölçüdür (Staudte ve Sheather, 1990: 44).

Yüksek kırılma noktalı tahmin ediciler çok sayıda aykırı değer varlığında güvenilir tahminler elde etmek için geliştirilmiştir. %50 kırılma noktası değerine sahiptirler ve dirençli tahmin ediciler olarak bilinirler. Bunlar hem x hem de y yönündeki aykırı değerlerin varlığında güvenilirdirler (Hawkins ve Olive, 2002).

En Küçük Kareler tahmini ve M tahminleri $1/n$ gibi düşük kırılma noktasına sahip olduklarından daha iyi performansa sahip tahminler bulmak için mühim çabalar harcanmıştır (Montgomery ve diğerleri, 2001: 401).

4.3.2. Asimptotik etkinlik

Sağlam regresyon konusunda bir diğer önemli konu etkinliktir. Seçilen bir sağlam yöntemin etkinliği bu teknikle elde edilen hata kareler ortalamasının EKK ile elde edilen hata kareler ortalamasına bölünmesiyle elde edilen oran olarak tanımlanabilir. Açık olarak bu oranın 1'e yakın olması istenmektedir. Literatürde asimptotik etkinliğe vurgu yapılır. Etkinlik rakamları sağlam tahmin edicileri karşılaştırmada faydalı olabilir. Sonlu örnek etkinliklerinde asimptotik değerleri için önemli ölçüde farklılıklar beklenebilir (Ryan, 2009: 427).

Kırılma noktası sağlamlığın bir ölçüsü iken etki fonksiyonu tahmin edici üzerindeki çok küçük bozulmaların etkisini ölçer. Etki fonksiyonu gözlemlerdeki dejenerasyonun neden olduğu asimptotik yanı ölçtüğünden asimptotik değerdeki çok küçük değişimlerin şeklini verir. Büyük örneklemdeki ölçülemeyecek kadar küçük bozulma oranının tahmin edici üzerindeki etkisini gösterir. $(n-1)$ gözlemli $(x_1, \dots, x_{n-1})$ örnekleme ve $T_{n-1}(x_1, \dots, x_{n-1})$ tahmin edicisi olduğunu varsayalım. Örneklemdeki bir tahmin edicinin etki fonksiyonu (EF); $T_n(x_1, \dots, x_{n-1}, x)$ tahmin edicisinin çizimidir.

$$EF(Z; T, F) = \lim_{\varepsilon \rightarrow 0} \frac{T((1 - \varepsilon)F + \varepsilon \Delta_2) - T(F)}{\varepsilon} \quad (4.3)$$

Δ_2 terimi, Z noktasındaki tüm kümeleri içine alan olasılık dağılımıdır. Etki fonksiyonu veride aykırı değerlerin sebep olduğu yanlılığı ifade eder. Dejenere olan dağılımda tahmin edicinin doğrusal bir yaklaşımını verir. Tüm tahmin ediciler bir kırılma noktasına sahiptir fakat etki fonksiyonuna sahip olmak zorunda değildir (Heikkila, 2003).

4.4. L Tipi Tahmin Ediciler

L tipi tahmin ediciler sınıfı, sıra istatistiklerinin doğrusal kombinasyonlarına dayandığından bu isim ile adlandırılırlar.

4.4.1. En küçük mutlak sapmalar yöntemi (Least absolute deviation)

1887’de Edgeworth tarafından çeşitli iyileştirmeler yapılan bu yöntem sadece dikey aykırı değerlere karşı sağlam, kaldıraç değerlere sağlam değildir (Gschwandtner ve Filzmoser, 2012).

En küçük mutlak sapmalar yöntemi, en küçük mutlak hatalar, en küçük mutlak değerler veya L_1 tipi regresyon olarak adlandırılır. En küçük kareler yönteminde, artıkların karelerinin toplamını minimum yapan $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahminleri seçilmek istenirken, en küçük mutlak sapmalar yönteminde artıkların mutlak değerlerinin toplamını minimum yapan $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahminleri seçilir (Radhakrishna ve Toutenberg, 1999: 273).

L_p normu tahminler olarak adlandırılan bu tahminlerde artıkların p . kuvvetlerinin olan $\sum |\hat{e}_i|^p$ ‘nin minimizasyonu istenir. p ’nin 1. kuvveti için amaç fonksiyonu en küçük mutlak sapmaları yani L_1 regresyonu verirken, artıkların 2. kuvveti için amaç fonksiyonu EKK olarak bilinen L_2 regresyonu verir. p ’nin 1 ve 2 arasındaki değerleri için L_p normu tahmini EKK ve LAD arasında bir uzlaşlı olarak sayılabilir.

En Küçük Mutlak Sapmalar yönteminin amaç fonksiyonu;

$$\min \sum_{i=1}^n |\hat{e}_i| = \min \sum_{i=1}^n |y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i| \quad (4.4)$$

şeklinde gösterilebilir.

En Küçük Mutlak Sapmalar Yöntemi'nin ilkeleri aslında En Küçük Kareler Yöntemi' nin ilkelerinden daha basittir. Çünkü $|\hat{e}_i|$ 'nin hesaplanması $\hat{e}_i^2$ 'nin hesaplanmasından daha kolaydır. Bununla beraber En Küçük Kareler Yöntemi'nin tahmin edicilerinin hesaplanması belirli formüllere dayanırken, En Küçük Mutlak Sapmalar Yöntemi'nin tahmin edicilerinin hesaplanması için herhangi bir formül bulunmadığından bu yöntemin tahmin edicilerini hesaplayabilmek için bir algoritma oluşturulması gereklidir. Algoritmanın temel mantığı, verilen bir (x_0, y_0) noktasından geçen tüm doğruların arasından en iyi doğruyu bulmaya dayanır. (x_0, y_0) noktasından geçen en iyi doğru bulunur. Bu yöntem “bir doğru iki noktadan geçer” ilkesine dayanır ve algoritma doğrunun geçtiği diğer bir nokta olan (x_1, y_1) noktasını bulmakla başlar. Daha sonra (x_1, y_1) noktasından geçen doğrular arasından en uygun doğru bulunmakta ve bu doğrunun ikinci noktası olan (x_2, y_2) bulunmaktadır. Algoritma daha iyi doğrular elde edene dek devam eder. Sonuç olarak elde edilen en son doğru bir önceki doğru ile aynı olacaktır. Bulunan doğru, tüm doğrular içinde toplam en küçük mutlak sapmaya sahip olduğundan “En Küçük Mutlak Sapmalar Regresyon Doğrusudur.”

Hesaplama adımları şu şekildedir;

Herhangi bir gözlem seçilir ve bu (x_0, y_0) başlangıç noktası olur. İki noktadan geçen doğrunun eğimi $(y_i - y_0) / (x_i - x_0)$ formülünden yararlanılarak her bir noktanın başlangıç noktası ile ayrı ayrı eğimi hesaplanır. Eğer $x_i = x_0$ ise eğim tanımlanamayacağından $x_i = x_0$ sağlayan noktalar ihmal edilir. Bütün eğimler hesaplandıktan sonra

$$\frac{(y_1 - y_0)}{(x_1 - x_0)} \leq \frac{(y_2 - y_0)}{(x_2 - x_0)} \leq \dots \leq \frac{(y_n - y_0)}{(x_n - x_0)}$$

olacak şekilde artan sırada dizilmektedir.

$T = \sum |x_i - x_0|$ olmak üzere;

$$|x_1 - x_0| + \dots + |x_{k-1} - x_0| < \frac{1}{2} T \quad (4.5)$$

$$|x_1 - x_0| + \dots + |x_{k-1} - x_0| + |x_k - x_0| > \frac{1}{2} T \quad (4.6)$$

Görüldüğü gibi $|x_k - x_0|$ ifadesinin eklenmesiyle $|x_1 - x_0|$ 'ın kümülatif toplamı $T/2$ 'yi geçmektedir. Bu koşulu sağlayan k noktası bulunur. Buna göre (x_0, y_0) noktasından geçen en iyi doğru;

$$\hat{\beta}_1 = \frac{y_k - y_0}{x_k - x_0}$$

$$\hat{\beta}_0 = y_0 - \hat{\beta}_1 x_0 \quad (4.7)$$

olmak üzere $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ olarak belirlenir.

Bir noktadan geçen birden fazla farklı uygun doğrunun olması “tek olmama durumu” (nonuniqueness) olarak tanımlanmaktadır. Bir noktadan geçen en iyi doğrunun aynı zamanda iki veya daha fazla noktadan geçmesi durumu ise “dejenere olma durumu” (degeneracy) olarak tanımlanmaktadır.

Bu tür problemler $\beta^* = \frac{y_k - y_0}{x_k - x_0}$ eğiminin $\frac{y_{k-1} - y_0}{x_{k-1} - x_0}$ eğimine veya $\frac{y_{k+1} - y_0}{x_{k+1} - x_0}$ eğimine eşit olması durumunda ortaya çıkmaktadır. Yukarıda anlatılan algoritma ile birçok veri kümesinin En Küçük Mutlak Sapmalar Regresyon doğrusu bulunabilmektedir fakat “tek olmama” ve “dejenere olma” durumlarında süreç karışmaktadır. Çünkü algoritmanın temel mantığı, verilen bir (x_0, y_0) noktasından geçen tüm doğruların arasından en iyi doğruyu ve bu doğrunun geçtiği diğer noktayı bularak, bulunan noktayı bir sonraki adımda kullanmaktır. Fakat tek olmama durumu mevcut olduğunda, veri noktasından geçen birden çok doğru olduğundan doğrulardan hangisinin seçilip işleme devam edileceği sıkıntı yaratmaktadır.

Dejenere olma durumunda ise, bulunan en iyi doğru birden çok noktadan geçtiği için, bir sonraki adımda kullanılmak üzere birden çok nokta içinden birini seçme sorunu yaşanmaktadır. Doğru seçim yapılmadığında, algoritma döngüye girebilmekte veya gerçekte “en küçük mutlak sapmalar regresyon doğrusu” olmayan bir doğru en küçük mutlak sapmalar regresyon doğrusuymuş gibi kabul edilebilmektedir. “Tek olmama” ve “Dejenere olma” durumlarıyla karşılaşıldığında, yukarıda verilen algoritma kullanılamamaktadır. Veri kümesinde tek olmama durumu mevcutsa, yani bir noktadan geçen birden fazla uygun doğru varsa, bu doğruların her biri için mutlak sapmaların

toplamı hesaplanmakta ve en küçük mutlak sapmaya sahip olan doğru seçilmektedir. Bu algoritmanın uygulanabilirliği, n örnek boyutuna bağlıdır.

Tek olmama durumunda yani birden fazla en küçük mutlak sapmalar regresyon doğrusu bulunması durumunda bunlardan biri keyfi olarak seçilebilmekte veya bunların ortalaması alınabilmektedir. Fakat dejenere olma durumunda hiçbir algoritma kullanılamamaktadır. Regresyon katsayılarının LAD tahmininin en iyi özelliği mutlak hataların toplamını minimum eden tahminler olmasıdır. Aynı zamanda hataların dağılımı Laplace dağılımına sahip olduğunda LAD tahmini β vektörünün en çok olabilirlik tahminidir. Rasgele hataların olasılığının 0'a yakın olduğunu varsayarsak, LAD tahminleri tutarlı ve asimptotik normaldir. LAD tahmininin sağlamlığının bir göstergesi hatalar sonlu varyansa sahip olduğunda tutarlı bir tahmin olmasının gerekli olmadığı gerçeği ile verilir. LAD tahmininin gücü, bağımlı değişkenlere karşı sağlam olmasıdır. Bu nedenle LAD tahminleri, M regresyon ve parametrik olmayan yöntemlerin iteratif algoritmalarında başlangıç tahmini olarak önerilir. LAD metodunu özellikle örnek çapı çok büyük olduğunda veya hataların dağılımı asimetrik veya çok kalın kuyruklu olduğunda kullanmak uygundur (Birkes ve Dodge, 1993: 57-62, 191-192, 205).

Konum tahmininin çok değişkenli olması durumunda L_1 tahmini örnek medyanıdır. Kırılma noktası %50 kadar yüksektir (Heikkilä, 2003).

4.5. M Tipi Tahmin Ediciler

Sağlam tahminlerin temel amacı, veri setinde bulunan yüksek hata değerine sahip aykırı gözlemlerin tahmin üzerindeki etkisini azaltmaktır. Sağlam yöntemler sayesinde aykırı değerler tespit edilebilir ve oluşturacakları etki bastırılabilir. Bu yöntemlerden en yaygın olanı 1964'te Huber tarafından sunulan M tahminidir. M tahmini bir ağırlıklandırma yöntemidir ve büyük artıklara küçük ağırlıklar vererek aykırı değerlerin olumsuz etkisini minimuma indirmeye çalışır.

M regresyondaki "M" maksimum olabilirlik tahmin edicileri ve M tahmini arasındaki ilişki nedeniyle seçilmiştir. M tahmininin temel amacı hata yığınlarının bir düzenlemesi için en iyi performans gösteren tahminleri elde etmektir.

Adını maksimum olabilirlik tahmincilerinden alan bu yöntem, artıkların karelerinin yerine artıkların daha az hızla artan bir fonksiyonunu minimize etmeye çalışır.

$$\text{Min } \sum \rho(y_i - b_0 - b_1 x_i) = \min \sum \rho(\hat{e}_i) \quad (4.8)$$

ρ fonksiyonunun türevi $\psi = \frac{\partial \rho(\hat{e})}{\partial (\hat{e})}$ etki fonksiyonu olarak adlandırılır. Bu durumda çözüm;

$$\sum \psi(y_i - b_0 - b_1 x_i) x_i = 0 \text{ olacaktır.} \quad (4.9)$$

Bu fonksiyon sınırlı ve sürekli olmalıdır (Huber, 1981: 43,54).

Ψ , $\Psi(-\infty) < 0 < \Psi(\infty)$ aralığında monoton artan ise her zaman mümkün bir çözüm vardır. Eğer ψ sürekli ve artansa o zaman çözüm tektir. Aksi takdirde çözüm ya bir nokta ya da bir aralık olacaktır (Maronna ve diğerleri, 2006: 24).

Uygun bir ρ fonksiyonu şu özellikleri taşımalıdır.

- $\rho(\hat{e}) > 0$
- $\rho(0) = 0$
- $\rho(\hat{e}) = \rho(-\hat{e})$
- $\rho(\hat{e}_i) > \rho(\hat{e}_i'), \quad |\hat{e}_i| > |\hat{e}_i'|$

Ağırlık fonksiyonu $w(\hat{e}) = \frac{\psi(\hat{e})}{\hat{e}}$ ve $w_i = w(\hat{e}_i)$ olarak tanımlanır.

Bu durumda tahmin eşitliği şöyle yazılabilir.

$$\sum w_i (y_i - b_0 - b_1 x_i) x_i = 0 \quad (4.10)$$

Çözüm bir ağırlıklandırılmış EKK tahminidir. Ağırlıklar artıklara, artıklar tahmin edilen katsayılar ve tahmin edilen katsayılar ağırlıklara bağlıdır. İteratif çözüm bu nedenle gereklidir (Fox, 2002).

EKK tahmininde $\sum \hat{e}_i^2$ minimum edilirken, en küçük mutlak sapmalar (LAD) tahmininde $\sum |\hat{e}_i|$ minimum edilmeye çalışılır. M tahmininde bu düşünce genelleştirilir ve $\hat{\beta}_0$ ve $\hat{\beta}_1$ 'nin

seçimi $\sum \rho(\hat{\epsilon}_i)$ fonksiyonunun minimumu ile mümkündür. $\rho(\hat{\epsilon})$; $\hat{\epsilon}$ 'nin bir fonksiyonudur. EKK tahmini $\rho(\hat{\epsilon}) = \hat{\epsilon}^2$ ve LAD tahmini $\rho(\hat{\epsilon}) = |\hat{\epsilon}|$ 'nin özel bir hali olarak kabul edilebilir. Huber M tahmini olarak bilinen M tahmini $|\hat{\epsilon}|$ ve $\hat{\epsilon}^2$ arasında uzlaşan bir $\rho(\hat{\epsilon})$ fonksiyonu kullanır. $\hat{\epsilon}$; 0'a yakın ise $\rho(\hat{\epsilon})$; $\hat{\epsilon}^2$ 'ye, 0'dan uzakta ise $|\hat{\epsilon}|$ 'ye eşit olarak her iki yönteminde avantajlarını birleştirmeyi dener.

$$\rho(\hat{\epsilon}) = \begin{cases} \hat{\epsilon}^2, & -k \leq \hat{\epsilon} \leq k \\ 2k|\hat{\epsilon}| - k^2, & \hat{\epsilon} < -k \text{ veya } k < \hat{\epsilon} \end{cases} \quad (4.11)$$

Huber $k=1,5\hat{\sigma}$ kabul eder. Bu durumda tahminin asimptotik etkinliği %96'dır. Burada $\hat{\sigma}$ rasgele hataların standart sapması σ 'nın bir tahminidir. $\rho(\hat{\epsilon})$ fonksiyonunu daha yumuşak bir fonksiyon yapabilmek için $|\hat{\epsilon}|$ yerine $2k|\hat{\epsilon}| - k^2$ kullanılır. σ 'yı tahmin etmek için $\hat{\sigma}=1,485 \cdot \text{MAD}$ kullanılır. Burada MAD; mutlak sapmaların medyanıdır. 1,485 çarpanı hataların dağılımının normal olduğu durumda σ 'nın iyi bir tahmin edici olmasını sağlamak amacıyla seçilmiştir.

Algoritma EKK yöntemi ile β_0 ve β_1 'nin başlangıç tahminleri bulunarak başlar. Bunlar σ ve hataları hesaplamak için kullanılır. σ ve artıklar yardımıyla β_0 ve β_1 'nin iyileştirilmiş tahminleri elde edilir. İyileştirilmiş tahminler yeni artıkları ve iyileştirilmiş $\hat{\sigma}$ 'yı hesaplamak için kullanılır. Bundan sonra yeni artıklar ve yeni $\hat{\sigma}$ ile katsayıların yeni tahminleri hesaplanır. Algoritma, katsayılar aynı ya da oldukça yakın olana kadar devam eder. Daha açık bir şekilde $\hat{\beta}_0$ ve $\hat{\beta}_1$, β_0 ve β_1 mevcut tahminleri olsun. $y_i - (\hat{\beta}_0 - \hat{\beta}_1 x_i)$ ile hatalar hesaplanır ve bunlarla $\hat{\sigma}^0 = 1,485 \text{MAD}$ hesaplanır. y değerlerini büyük sapmalardan temizlemek amacı ile yapılacak düzenleme şu şekildedir. Tahmin edilen regresyon doğrusunda y_i 'nin sapmaları $\hat{\epsilon}_i^0 = y_i - (\hat{\beta}_0^0 + \hat{\beta}_1^0 x_i)$ 'dir. Buradan $y_i = (\hat{\beta}_0^0 + \hat{\beta}_1^0 x_i) + \hat{\epsilon}_i^0$ olur. İyileştirilmiş $y_i^* = (\hat{\beta}_0^0 + \hat{\beta}_1^0 x_i) + \hat{\epsilon}_i^*$ tanımlanır. Burada $\hat{\epsilon}_i^*$, $\hat{\epsilon}_i^0$ 'nin daraltılmasıyla elde edilen düzeltilmiş sapmalardır. Öyle ki sapmalardan hiçbirinin mutlak değeri $1,5\hat{\sigma}^0$ 'den büyük değildir. $\hat{\epsilon}_i^0$; $-1,5\hat{\sigma}^0$ ve $1,5\hat{\sigma}^0$ arasında ise $\hat{\epsilon}_i^* = \hat{\epsilon}_i^0$, $-1,5\hat{\sigma}^0$ 'den küçük ise $\hat{\epsilon}_i^* = -1,5\hat{\sigma}^0$, $1,5\hat{\sigma}^0$ 'den büyük ise $\hat{\epsilon}_i^* = 1,5\hat{\sigma}^0$ olur. Düzeltilmiş $y_1^*, y_2^*, y_3^*, \dots, y_n^*$ verisinden elde edilen EKK tahminleri β_0 ve β_1 'nin iyileştirilmiş tahminleridir. Genelde hesaplamada tahminler arasındaki fark 10^{-4} 'ten küçük olduğunda algoritma durdurulur. Bir M tahmininde ρ fonksiyonu veya türevi $\rho' = \psi$ etki fonksiyonu ile tanımlanır. Tahmine etki

fonksiyonunun şekli karar verdiğinden ψ sıklıkla tercih edilir. (Birkes ve Dodge, 1993: 87-89, 100).

Huber M tahmini için etki fonksiyonu;

$$\psi(\hat{e}) = \begin{cases} \hat{e}, & |\hat{e}| < k \\ k \operatorname{sgn}(\hat{e}), & |\hat{e}| \geq k \end{cases} \quad (4.12)$$

Ağırlık fonksiyonu;

$$w(\hat{e}) = \begin{cases} 1, & |\hat{e}| < k \\ k/|\hat{e}|, & |\hat{e}| \geq k \end{cases} \quad (4.13)$$

M tahminine birçok farklı yaklaşım vardır. Ana düşünce aynı olmakla birlikte farklı kullanılan fonksiyonlardan kaynaklanmaktadır.

Tukey'in Yaklaşımı

ρ fonksiyonu;

$$\rho(\hat{e}) = \begin{cases} 1 - [1 - (\hat{e}/k)^2]^3, & |\hat{e}| < k \\ 1, & |\hat{e}| \geq k \end{cases} \quad (4.14)$$

Etki fonksiyonu;

$$\psi(\hat{e}) = \begin{cases} \hat{e}[1 - (\hat{e}/k)^2]^2, & |\hat{e}| < k \\ 0, & |\hat{e}| \geq k \end{cases} \quad (4.15)$$

Ağırlık Fonksiyonu;

$$w(\hat{e}) = \begin{cases} [1 - (\hat{e}/k)^2]^2, & |\hat{e}| < k \\ 0, & |\hat{e}| \geq k \end{cases} \quad (4.16)$$

(Maronna ve diğerleri, 2006: 29,30).

Tukey 'in fonksiyonu aykırı değerlerin etkisini daha çok bastırabilir. Tukey ayar sabiti olarak $k=4,685\sigma$ değerini kullanır (Arya, 2007).

Hampel'in Yaklaşımı

2-4-8 tahmini olarak da adlandırılan Hampel'in 3 parçalı fonksiyonu şöyle tanımlanır;

ρ fonksiyonu;

$$\rho(\hat{e}) = \begin{cases} \frac{1}{2}\hat{e}^2, & |\hat{e}| < a \\ a|\hat{e}| - \frac{1}{2}a^2, & a < |\hat{e}| \leq b \\ ab - \frac{1}{2}a^2 + (c-b)\frac{a}{2}\left[1 - \frac{c-|\hat{e}|}{c-b}\right], & b < |\hat{e}| \leq c \\ ab - \frac{1}{2}a^2 + (c-b)\frac{a}{2}, & |\hat{e}| > c \end{cases} \quad (4.17)$$

Etki fonksiyonu;

$$\psi(\hat{e}) = \begin{cases} \hat{e} & , \quad |\hat{e}| \leq a \\ a \operatorname{sgn}(\hat{e}) & , \quad a < |\hat{e}| \leq b \\ a \frac{c-|\hat{e}|}{c-b} \operatorname{sgn}(\hat{e}) & , \quad b < |\hat{e}| \leq c \\ 0 & , \quad |\hat{e}| > c \end{cases} \quad (4.18)$$

Ağırlık fonksiyonu;

$$w(\hat{e}) = \begin{cases} 1 & , \quad |\hat{e}| \leq a \\ \frac{a \operatorname{sgn}(\hat{e})}{\hat{e}} & , \quad a < |\hat{e}| \leq b \\ a \frac{c-|\hat{e}|}{c-b} \frac{\operatorname{sgn}(\hat{e})}{\hat{e}} & , \quad b < |\hat{e}| \leq c \\ 0 & , \quad |\hat{e}| > c \end{cases} \quad (4.19)$$

Burada; a, b ve c sabitleri sırasıyla 2, 4 ve 8 olarak belirlenmiştir (Hampel, 1974).

Andrews'in Yaklaşımı

ρ fonksiyonu;

$$\rho(\hat{e}) = \begin{cases} 1 - \cos(\hat{e}), & -\pi \leq \hat{e} \leq \pi \\ 2, & \text{d. d.} \end{cases} \quad (4.20)$$

Etki fonksiyonu;

$$\psi(\hat{e}) = \begin{cases} \sin(\hat{e}), & \pi \leq \hat{e} \leq \pi \\ 0, & \text{d. d.} \end{cases} \quad (4.21)$$

Ağırlık fonksiyonu;

$$w(\hat{e}) = \begin{cases} \sin(\hat{e})/e, & \pi \leq \hat{e} \leq \pi \\ 0, & \text{d. d.} \end{cases} \quad (4.22)$$

(Huber,1981: 54).

Fair'in Yaklaşımı

ρ fonksiyonu;

$$\rho(\hat{e}) = c^2 \left[\frac{|\hat{e}|}{c} - \log \left(1 + \frac{|\hat{e}|}{c} \right) \right] \quad (4.23)$$

Etki fonksiyonu;

$$\psi(\hat{e}) = \frac{\hat{e}}{1 + \frac{|\hat{e}|}{c}} \quad (4.24)$$

Ağırlık fonksiyonu;

$$w(\hat{e}) = \frac{1}{1 + \frac{|\hat{e}|}{c}} \quad (4.25)$$

c sabiti 1,3998 olarak seçilmiştir. Fair fonksiyonu için Roepack paket programı kullanılabilir. İlk üç türevi her yerde sürekli tanımlıdır ve tek bir çözüm verir (Zhang, 1995).

Fair (1974) fonksiyonu, büyük artıklar için En Küçük Kareler ve En Küçük Mutlak Sapmalar yöntemi bir arada kullanılarak oluşturulmuştur. Fair ve Lojistik fonksiyonlarda sonuç konveks bir amaç fonksiyonundadır. Amaç fonksiyonunun konveksliği, sadece doğrusal sınırlamalar tarafından tanımlanabilen bir süreç için doğrusal olmayan verinin uzlaşma probleminde global optimumu garanti eder (Özyurt ve Pike, 2004).

German-McClure Yaklaşımı

ρ fonksiyonu;

$$\rho(\hat{e}) = \frac{\frac{\hat{e}^2}{2}}{1 + \hat{e}^2} \quad (4.26)$$

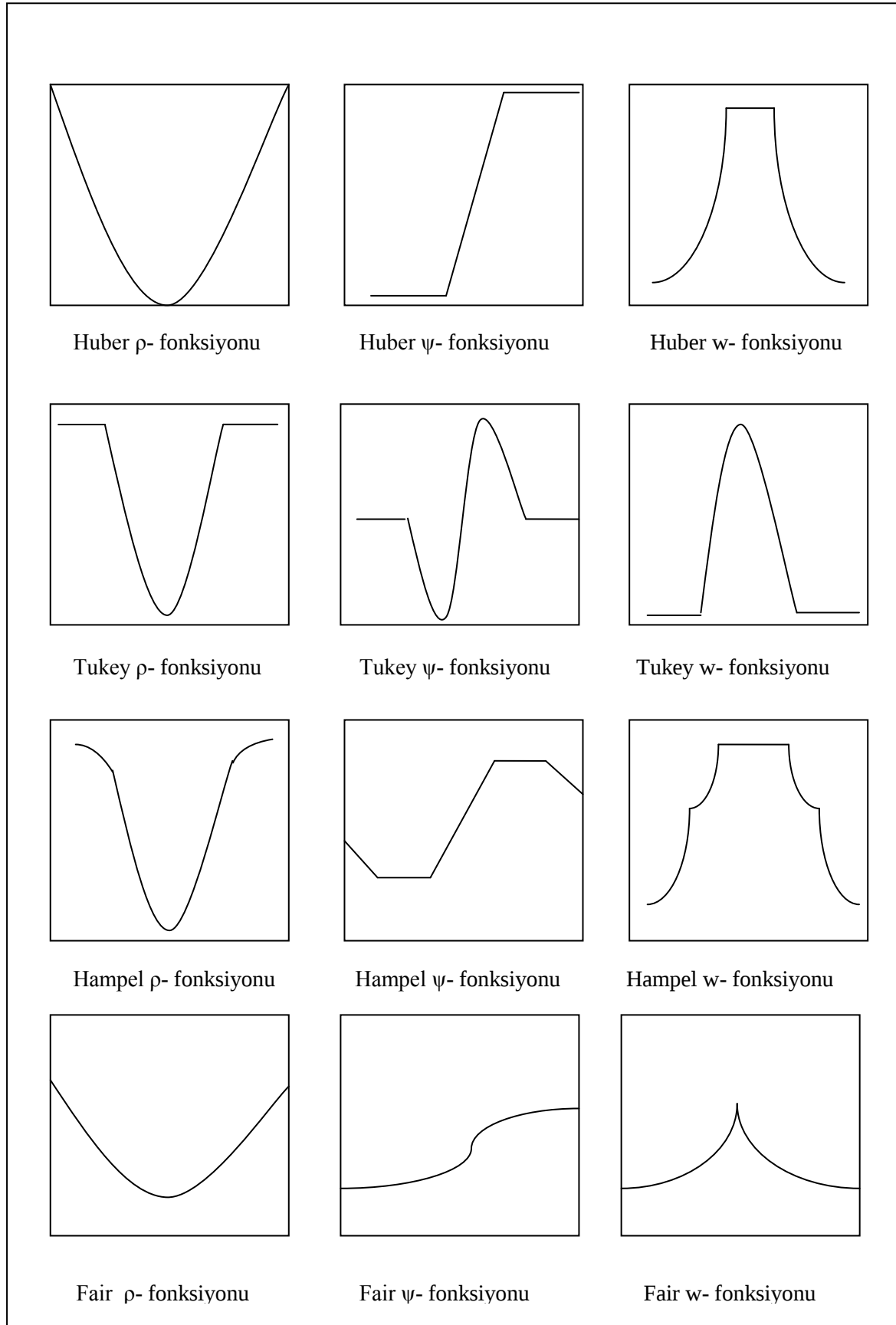
Etki fonksiyonu;

$$\psi(\hat{e}) = \frac{\hat{e}}{(1 + \hat{e}^2)^2} \quad (4.27)$$

Ağırlık fonksiyonu;

$$w(\hat{e}) = \frac{1}{(1 + \hat{e}^2)^2} \quad (4.28)$$

(Zhang, 1995).



Şekil 4.1. Bazı M tahminlerinin ρ , ψ ve w fonksiyonlarının grafik gösterimi (Zhang, 1995)

4.6. S Tipi Tahmin Ediciler

S tahmini terimi ilk kez Rousseeuw ve Yohai (1984) tarafından regresyonda yüksek kırılma noktalı bir tahmin edici sınıfı tanımlamak amacıyla kullanılmıştır. S tahmin sınıfı LMS tahmininin genelleştirilmiş bir hali olarak görülebilir. LMS tahmini açıklayıcı değişken sayısı ne olursa olsun 0,50'ye yakın kırılma noktası değerine sahiptir. LMS tahminine benzer olarak S tahmini de açıklayıcı değişken sayısına bağlı olmaksızın yüksek kırılma noktasına sahiptir. Ayrıca S tahmininin tutarlılığı $n^{1/2}$ iken LMS tahmininin tutarlılığı $n^{1/3}$ 'tür (Tyler, 1999: 659-662).

S tahmini artıkların saçılımının minimizasyonu şeklinde tanımlanır. Amaç fonksiyonu şöyle gösterilebilir;

$$\min s(\hat{e}_1(\theta), \dots, \hat{e}_n(\theta))$$

ölçek tahmini;

$$\hat{\sigma}_{s0} = s(\hat{e}_1(\hat{\theta}), \dots, \hat{e}_n(\hat{\theta})) \quad (4.29)$$

Dağılım $s(\hat{e}_1(\theta), \dots, \hat{e}_n(\theta))$ aşağıdaki eşitliğin çözümü olarak tanımlanır.

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{\hat{e}_i}{s}\right) = K \quad (4.30)$$

ρ fonksiyonu aşağıdaki şartları sağlamalıdır.

1. ρ ; simetrik, sürekli, türevlenebilir ve $\rho(0)=0$ olmalıdır.
2. $c>0$ vardır. Öyle ki; ρ $[0,c]$ aralığında sınırlı artan ve $[c,\infty)$ aralığında sabittir.

ρ fonksiyonu olarak Rousseeuw ve Yohai (1984), aşağıdaki fonksiyonu önermişlerdir.

$$\rho(\hat{e}) = \begin{cases} \frac{\hat{e}^2}{2} - \frac{\hat{e}^4}{2c^2} + \frac{\hat{e}^6}{6c^4}, & |\hat{e}| \leq c \\ \frac{c^2}{6}, & |\hat{e}| > c \end{cases} \quad (4.31)$$

c ayar sabiti 1,547 olarak alınır.

S tahmininin kırılma noktası;

$$\varepsilon^* = \frac{K}{\rho(c)} \quad n \rightarrow \infty \quad (4.32)$$

ile elde edilir (Heikkila, 2003; Rousseeuw ve Leroy, 1987: 136; Ryan, 2009: 435).

Ne yazık ki S tahmini yüksek etkinliğe ve yüksek kırılma noktasına aynı anda sahip olamaz. Hössjer (1992), özellikle 0,5 kırılma noktalı S tahminin hata terimleri normal dağıldığında asimptotik etkinliğinin 0,33'ten büyük olmadığını göstermiştir. Sayısal hesaplamalar biweighted (Tukey) ölçeği 0,29 olduğunda S tahmininin normal dağılımın etkinliğini verdiğini göstermiştir (Maronna ve diğerleri, 2006: 130, 131).

Gürünlü Alma (2011), LTS, MM, S ve M tahminini karşılaştırdığı çalışmasında S tahmininin y-yönündeki aykırı değerlerin yüzdesi düştüğünde diğer sağlam yöntemlere göre daha iyi performans gösterdiğini ancak aykırı değer yüzdesi arttıkça tüm yöntemlerin performansında düşüş olacağını belirtmiştir.

4.7. MM Tahmini

İlk olarak Yohai tarafından 1987'de sunulan MM tahmini giderek daha popüler hale gelmekte ve belki de şimdilerde en sık kullanılan sağlam regresyon tekniklerinden biri olmaktadır. Yüksek kırılma noktası ve yüksek etkinlik özelliklerinin birleştiği bu yöntem hesaplamada birden fazla M tahmini kullandığından MM adını almıştır (Maru ve Modi, 2012).

Bir M tahmini ile aşağıdaki fonksiyonun minimumunu bulmak gerekir.

$$L(\beta_0, \beta_1) = \sum_{i=1}^n \rho\left(\frac{\hat{e}_i}{\hat{\sigma}}\right) \quad (4.33)$$

Yöntemin hesaplama adımları şöyledir;

1. Yüksek kırılma noktalı tutarlı $\widehat{\beta}_0^0$ ve $\widehat{\beta}_1^0$ başlangıç tahminleri elde edilir. Fakat bu tahminler yeterli etkinliğe sahip değildir.
 2. Elde edilen tahminlerin yardımıyla $\widehat{\sigma}_{MM}^0$ sağlam ölçek tahmini hesaplanır.
 3. $\widehat{\beta}_0^0$ ve $\widehat{\beta}_1^0$ başlangıç tahmini ile başlayan iteratif bir yöntem kullanılarak $\widehat{\beta}_0$ ve $\widehat{\beta}_1$ elde edilir.
- $\widehat{\sigma}_{MM}^0$ bir M ölçek tahmini olmalıdır.

$$\frac{1}{n} \sum \rho_0 \left(\frac{\widehat{e}_i}{\widehat{\sigma}_{MM}^0} \right) = 0,5 \quad (4.34)$$

$L(\widehat{\beta}_0, \widehat{\beta}_1) \leq L(\widehat{\beta}_0^0, \widehat{\beta}_1^0)$ olduğunda sonuç elde edilir (Maronna ve diğerleri, 2006: 124, 125).

MM tahminin kırılma noktası her zaman % 50 iken, yanlılığı etkinliği ile artar. Bu nedenle MM tahminin etkinliği ile sağlamlığı arasında bir uzlaşma sağlamak için etkinliğin değeri % 70 olarak varsayılır (Verardi ve Croux, 2009).

4.8. En Küçük Budanmış Kareler Yöntemi (Least Trimmed Squares)

Bu regresyon yöntemi ilk kez 1984'te Rousseeuw tarafından sunulmuştur ve aykırı değerlerin veriden çıkarılmasının mümkün olduğu bir tekniktir. En Küçük Budanmış Kareler yönteminde de katsayılar EKK modeline benzer olarak tahmin edilir. Diğer bir deyişle katsayılar artık karelerin toplamının minimizasyonu şeklinde tahmin edilmek istendiğinde iki yol kullanılabilir (Momeni ve diğerleri, 2010; Rousseeuw 1984).

Fakat LTS'de EKK 'den farklı olarak regresyon modelinde tüm veriler kullanılmaz, artık değeri yüksek veriler çıkarılarak tahmin yapılır (Yaffee, 2002).

LTS tahmini;

$$\min \sum_{i=1}^h \widehat{e}_i^2 = \min \sum_{i=1}^h (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 \quad (4.35)$$

$$h=(k+n+1)/2$$

Burada; n : gözlem sayısı

y_i : i . gözlem değeri

$\hat{y}_i$: i . gözlemin tahmin değeri

h : LTS 'de kullanılan veri sayısı

k : sabit terim dahil olmak üzere parametre sayısıdır.

Burada $\hat{e}_1^2 < \hat{e}_2^2 \dots \dots < \hat{e}_n^2$ sıralı artık karelerdir. Amaç öyle bir tahmin bulmaktır ki sonuçta artıkların kareleri büyükten küçüğe sıralandığında hedef alınan belirli bir sıradaki artıkların karesi başka tahminlerin aynı sıradaki sıralı artık karelerinden daha düşük olsun (Rousseeuw ve Leroy, 1987: 15).

Hesaplama adımları şöyledir;

1. h belirlenir.
2. k gözlemlili tüm mümkün alt kümeler oluşturulur ve regresyon katsayıları hesaplanır.
($\beta_{(1)}, \dots, \beta_{(q)}$)
3. LTS ölçütünden n gözlemin tamamı kullanılarak artıklar hesaplanır.
4. Amaç fonksiyonunu minimum yapan $\hat{\beta}_{(j)}$, LTS tahminini verir (Doornik, 2011).

Alternatif bir hesaplama şekli ise Nevitt ve Tam (1998) tarafından şöyle verilmiştir.

Artıkların kareleri $\hat{e}_i^2$ değerleri hesaplanır ve küçükten büyüğe doğru sıralanır. Sıralamanın ardından $0 < \alpha < 1/2$ olmak koşulu ile artıkların değerlerine ait en büyük α_n değeri silinerek geriye kalan gözlem değerleri için EKK hesaplanması mantığına dayanmaktadır. Uygulamada α değerinin sıklıkla 0,05 veya 0,1 olarak seçildiği görülmektedir. Yani ilk aşamada EKK tekniği kullanılarak hesaplanan büyük değerli artıklara sahip gözlemler veri setinden silinir. Daha sonra kalan verilerle $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin değerleri EKK ile hesaplanır. Genelde h budama oranı α 'ya bağlı olarak hesaplanır. $h = [n(1-\alpha)] + 1$ olduğunda bozulma sınırı $\epsilon^* \cong \alpha$ olacaktır.

Budama oranını belirlemek oldukça önemlidir. Eğer varsayılandan daha fazla veri kirliliği ile karşı karşıya kalınırsa LTS yeterli sağlamlığı gösteremez. Aksine kirlilik varsayılandan daha az ise, parametre tahminleri doğruluk yönünden sorun çıkarır ve tüm iyi noktalar hesaba katılmadığından yeterli etkinliğe sahip olamaz. h değerine göre aykırı değerli veri yapısına bağlı olarak LTS oldukça etkin olabilir. Aykırı değerli noktaların tamamı budanırsa hesaplama açısından EKK ile eşit olur (Schumacker ve diğerleri, 2002).

LTS tahmin edicisinin özellikleri;

1. Her zaman bir çözümü vardır.
2. Eğer parametre sayısı $p > 1$, $h = [n/2] + [(p+1)/2]$ ve gözlemler genel durumda ise yani gözlemlerin p tanesi için β 'nin tek çözüm varsa LTS yönteminin bozulma sınırı $\varepsilon^* = ([n-p]/2 + 1)/n$ şeklinde tanımlanır. Daha büyük h değerleri için LTS' nin bozulma sınırı $\varepsilon^* \approx (n-h)/n$ 'dir. Örneğin, $h \approx 0,75n$ seçildiğinde yöntemin bozulma sınırı $\varepsilon^* \approx 0,25$ olacaktır (Hubert, Rousseeuw, Aelst, 2003: 6). Genel olarak h budama oranı α 'ya bağlı olarak hesaplanır. Örneğin; $h = [n(1-\alpha)] + 1$ olduğunda bozulma sınırı $\varepsilon^* \approx \alpha$ 'dır. Rousseeuw (1987), regresyon modelini bozabilecek aykırı değerlerin en küçük oranını kırılma noktası olarak tanımlamıştır.
3. Eğer $p > 1$ ve $y_i = x_i\beta$ 'yi genel düzendeki gözlemlerden $(n+p+1)/2$ 'den fazlası kesin olarak sağlıyor ise diğer gözlemler ne olursa olsun LTS tahmin edicisinin çözümü β 'ya eşit olur.
4. LTS tahmin edicisi; LMS' nin aksine normal dağılım durumunda da aynı asimptotik etkinliğe sahiptir.
5. LTS yönteminin asimptotik etkinliği eğer $\alpha = 0$ ise sadece %7,2 olarak bulunur. Ancak h değeri n değerine yaklaştıkça LTS' nin etkinliği EKK' nin etkinliğine yaklaşacaktır (Heikkilä, 2003).

Rousseeuw ve Van Driessen çoğu yönden LTS yönteminin LMS yönteminden üstün olduğuna dikkat çekmişlerdir. Örneğin, LTS amaç fonksiyonu LMS'nin amaç fonksiyonundan daha yumuşaktır yani LTS tahmini yerel etkilere karşı daha az hassastır (Hössjer, 1994).

Bu avantajlarına rağmen LTS yönteminin hesaplanması oldukça zordur. LTS amaç fonksiyonunu matematiksel optimizasyonlara elverişli olmaması ve buna bağlı olarak tahminlerin hesaplanmasının zorluğu önemli sorunlar olarak görülmektedir. LTS tahmin edicisinin kesin hesaplamasında h boyutlu alt örneklerin tamamı çekilir. Bu C_h^n kadar alt kümenin elde edilmesi demektir. Her bir alt küme için regresyon doğrusu uydurulur. Buradan, amaç fonksiyonunda belirtildiği gibi her bir örnek için hesaplanan hata karelerin en küçük h tanesinin toplamı bulunur. Bu toplamlar arasından en küçük olanı seçilerek LTS çözümü olarak alınır. Burada karşımıza çıkan en büyük sorun örneklemden çekilecek alt küme sayısının çok büyük boyutlara ulaşabilmesidir. Örnek çapı büyüdükçe işlem

yoğunluğu ve harcanan zaman artacak ve belli bir aşamadan sonra LTS çözümüne ulaşmak imkansız olacaktır (Türkay ve Öztürk, 2005).

4.9. En Küçük Medyan Kareler Yöntemi (Least Median Squares)

Rousseeuw (1984) tarafından geliştirilen En Küçük Medyan Kareler yöntemi için amaç fonksiyonu;

$$\min \text{med } \hat{e}_i^2 \quad (4.36)$$

olarak ifade edilir.

Bu tahminin y yönündeki aykırı değerlere de, x yönündeki aykırı değerler kadar sağlam olduğu anlaşılmıştır. Kırılma noktası %50 gibi yüksek bir olasılık değerine sahiptir fakat asimptotik etkinlik bakımından zayıftır. Bu metodun EKK yönteminden farkı, amaç fonksiyonu $\min \sum [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$ 'de $\sum$ yerine medyan ifadesini kullanmasıdır. Yani metodun amacı artıkların medyanını minimum yapmaktır. LMS Eş. 4.36'da gösterildiği gibi sadece artıkları kullandığından açıklayıcı değişkenler üzerindeki doğrusal dönüşümler bakımından eş değişkenlidir (Rousseeuw ve Leroy, 1987: 14).

Rousseeuw (1984) algoritmasında, verinin tüm mümkün alt kümeleri elde edilir. Her bir alt kümeye EKK uygulanır ve EKK tahminleri üzerinden elde edilen artıkların medyanı bulunur. Bulunan medyan kareler içerisinde en küçük değere sahip olanı en küçük medyan kareler tahminini verir. İşlemler altküme sayısı kadar tekrarlanacağından, büyük veri setleri için hesaplamak olanaksız hale gelecektir. Rousseeuw ve Leroy (1987), belirli kısıtlar altında veriden rasgele çekilen en az bir alt kümenin istenilen sonucu verme olasılığının bire yakın olduğunu ispatlamıştır. Buna göre n birimlik bir veriden p elemanlı m tane altküme seçtiğimizde p tane aykırı olmayan değer içeren en az bir alt kümeye rastlama olasılığı n/p 'nin çok büyük değerleri için aşağıdaki gibi ifade edilir.

$$p = 1 - [1 - (1 - \varepsilon^*)^m]^k \quad (4.37)$$

Burada ε^* verideki kirlenme oranıdır (Sınıksaran ve Aktükün, 2005).

LMS tahmininin özellikleri aşağıdaki gibidir.

1. Her zaman bir çözümü mevcuttur.
2. $p > 1$ ve gözlemler genel durumda ise, o zaman LMS tahmin edicisinin kırılma noktası $([n/2]-p+2)/n$ olur. Kırılma noktası az da olsa n 'e bağlıdır. Tek bir değere sahip olması için genellikle p sabit iken $n \rightarrow \infty$ sınırı dikkate alınır. Bu yüzden EKK tahmininin kırılma noktası %0 iken, LMS tahminininki %50 kadar yüksektir.
3. Eğer $p > 1$ ve gözlemler genel durumda ise $(n-[n/2]+p-1)$ kadarının $y_i = x_i + \beta$ eşitliğini tam olarak sağlayan bir β varsa diğer gözlemler ne olursa olsun LMS tahmin edicisinin çözümü β 'ya eşittir (Rousseeuw, 1984).

4.10. Kantil Regresyon

Kantil Regresyon; Koenker ve Bassett (1978) tarafından hata terimlerinin normal dağıldığı varsayımı konusunda şüpheler olduğunda EKK tahminine alternatif sağlam bir tahmin tekniği olarak sunulmuştur (Levin, 2002).

Kantil Regresyon bağımsız değişkenlerin bir fonksiyonu olarak ifade edilen bağımlı değişkenin koşullu kantil fonksiyon modellerini tahmin etmeye çalışır (Koenker ve Hallock, 2001).

Sabit varyansın geçerli olup olmadığı, kantil regresyon tahminleri ile analiz edilebilir. Standart sabit varyanslı regresyon modelinde regresyon doğrusu y 'nin x 'e göre koşullu beklenen değeridir. Hata terimlerinin dağılımı sabit varyanslı ise kantiller regresyon doğrusuna paraleldir (Saçaklı, 2005).

Kantil regresyon modeli, EKK metodundan daha esnek ve bağımlı değişkenin dağılımına ait kovaryans etkilerini incelemeye imkan sağlar (Çağlayan ve Arikan, 2011).

Kantil regresyon açıklayıcı değişkendeki bir birimlik değişimin bağımlı değişkenin belli bir kantilindeki değişimini tahmin eder.

Klasik En Küçük Kareler tahmininin amacı; $E(Y/X)$ 'in fonksiyonel biçimi ile ilgili bazı varsayımlar altında (örneğin; doğrusallık) " $E(Y/X)$ 'e bağlı bazı X açıklayıcı değişkenlerinde bir Y rasgele değişkeninin koşullu beklentisi nasıldır?" sorusunu

cevaplamaktır. Buna karşın kantil regresyon koşullu dağılımın herhangi bir kantili için böyle bir soruya cevap sağlar (Leping, 2005).

Kantil regresyon, koşullu kantil fonksiyonlar hakkında çıkarım yapmak ve tahmin elde etmek için tasarlanmış bir istatistiksel tekniktir. Klasik doğrusal regresyon metotları koşullu ortalama fonksiyonu için artık karelerin toplamını minimize etmeye dayalı bir modelle tahmin sağlarken, kantil regresyon koşullu medyan fonksiyonu için tüm koşullu kantil fonksiyonları kapsayan bir modelle tahmin etmeye çalışır. Kantil regresyon rasgele değişkenler arasında stokastik ilişkilerin daha kapsamlı bir istatistiksel analizini yapma yeteneğine sahiptir. Kantil regresyon pediatride çocukların büyüme eğrilerinde, ekonomide ise ücret, gelir eşitsizliğindeki eğilim gibi konularda yaygın olarak kullanılır (Koenker, 2002).

Kantil Regresyon modeli artıkların mutlak değerlerinin asimetrik olarak ağırlıklandırılmış toplamını en küçük kılma ilkesi üzerine kuruludur. Kantil regresyon analizini EKK ilkesine tercih etmenin en önemli nedeni, iktisadi değişkenler arasındaki davranışsal ilişkiler analiz edildiğinde, EKK tahmin edicilerinin verideki aşırı değerlere duyarlılığının, normallik varsayımı geçerli olmadığı durumda iyi tahmin ediciler olmasını engellemesidir. Normallik varsayımının yerine getirildiği durumlarda, Kantil Regresyon modeli tahmin edicileri EKK tahmin edicileri ile benzer performans gösterirken, normallik varsayımı ihlal edildiğinde EKK tahmininden daha etkin sonuçlar üretebilmektedir (Çolak ve Öztürkler, 2012).

Herhangi gerçek değerli bir rasgele değişken Y 'nin dağılım fonksiyonu;

$F(y) = P(Y < y)$ ile karakterize edilir.

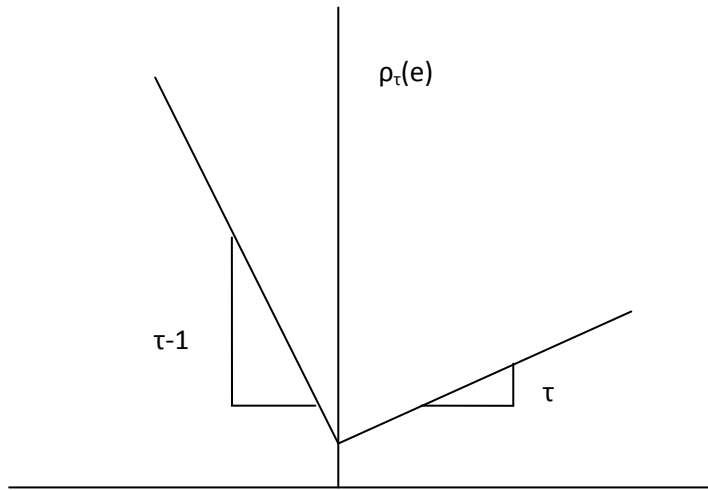
Herhangi $0 < \tau < 1$ için

$$Q(\tau) = \inf \{ y : F(y) > \tau \}$$

Y 'nin τ . kantili olarak adlandırılır. Medyan $Q(1/2)$ merkez rolü oynar. Dağılım fonksiyonu gibi kantil fonksiyonu da Y rasgele değişkeninin tam bir karakterizasyonunu sağlar.

Kantiller basit bir optimizasyon probleminin çözümü olarak formüle edilebilirler. $0 < \tau < 1$ için parçalı doğrusal “check” fonksiyonu ;

$\rho_\tau(e) = e(\tau - 1(e < 0))$ şeklinde tanımlanabilir ve Şekil 4.2'deki gibi gösterilebilir.



Şekil 4.2. Kantil regresyonun ρ fonksiyonu (Koenker, 2005: 5-7)

Doğrusal regresyon modelimiz;

$$y = x_i \beta + e_i \text{ olsun.}$$

Koenker ve Bassett (1978) herhangi β vektörü için kantil regresyon tahminini şöyle tanımlamıştır.

$$\widehat{\beta}_n(\tau) = \min \sum_{i=1}^n \rho_\tau(\widehat{e}_i(\beta)) \quad (4.38)$$

$$\rho_\tau(\widehat{e}(\beta)) = |\widehat{e}(\beta)| [\tau_{\{\widehat{e}(\beta) \geq 0\}} + (1 - \tau)_{\{\widehat{e}(\beta) < 0\}}] = \begin{cases} (1 - \tau)\widehat{e}(\beta), & \widehat{e}(\beta) < 0 \\ \tau\widehat{e}(\beta), & \widehat{e}(\beta) \geq 0 \end{cases} \quad (4.39)$$

$1_{\{A\}}$; A 'nın işaret fonksiyonudur. A, pozitif ise 1 aksi takdirde 0 değerini alır (Neykov, Cizek, Filzmoser ve Neytchev, 2012).

Bu yöntemin sağlamlılığı, hata terimlerinin gerçek büyüklüğüne değil tahminlerin işaretlerine bağlı olmasından kaynaklanmaktadır (Gill, 2002: 249).

Kantil regresyon tahmini çarpıklığa ve normallikten sapmalara karşı sağlamdır. Regresyon katsayılarının hesaplanması ile algoritmalar Koenker (2005), tartıştığı gibi doğrusal

programlama tekniklerine veya Hunter ve Lange (2000) ve Chen (2004) tarafından sunulan maksimizasyon minimizeleme tekniklerine dayanır (Neykov ve diğeri, 2012).

Kantil çizgilerinin aralığı dağılımın sağa veya sola çarpık olup olmadığını gösterir.

Kantil regresyon katsayılarının tahmini için bir optimizeleme problemi oluşturulabilir.

$$u=[y-x\beta] \quad \beta^1=[\beta]_+$$

$$v=[x\beta-y] \quad \beta^2=[-\beta]_+$$

$y=(y_1, y_2, \dots, y_n)'$, $x=(x_1, x_2, \dots, x_n)'$ olmak üzere Eş. 4.39 şu şekilde yeniden formüle

edilebilir.

$$\min_{\beta} \{\tau r'u + (1 - \tau)r'v | y = X\beta + u - v, \{u, v\} \in R_+^n\} = X(\beta^1 - \beta^2) + u - v \quad (4.40)$$

Burada r ; n - boyutlu birim vektördür.

$$A=[X, -X, I, -I]$$

$$z=[\beta^1, \beta^2, u, v]$$

$$c=[0' \ 0' \ \tau r' (1 - \tau)r']' \text{ olmak üzere;}$$

Amaç fonksiyonu:

$$\begin{aligned} \min c'z \quad & Az=y \\ & z \geq 0 \text{ olacaktır} \end{aligned}$$

(Chen ve Wei, 2005; Koenker, 2005: 7-10).

Kantil regresyonunun etki fonksiyonu;

$\tilde{T} = T^{-1}(\tau)$ olmak üzere

$$EF_{\text{kantil}} = \text{sgn}(y - \tilde{T}) / f(T^{-1}(\tau)) \quad (4.41)$$

ile hesaplanır (Koenker, 2005:43).

Kantil regresyon probleminin doğrusal programlama gösterimi hem teorik hem de pratik olarak birkaç önemli çıkarım içerir. Birincisi, sınırlı sayıda simpleks iterasyonla kantil regresyon tahmininin elde edilebileceğinin garanti olduğudur. İkincisi, ortalama tipi regresyon probleminin tersine parametre tahmini sağlamdır. $\hat{\beta}_0$ 'nın çözümü değiştirilmeksizin eğer $y_i - x_i' - \hat{\beta}_0 > 0$ ise y_i, ∞ 'a, $y_i - x_i' - \hat{\beta}_0 < 0$ ise $y_i, -\infty$ 'a gider. Diğer bir deyişle önemli olan bağımlı değişkenin büyüklüğü değil gözlemin tahmin hiperdüzleminin hangi yanında olduğudur. Kantil regresyonun EKK üzerindeki en büyük avantajı heterojen koşullu dağılımlı veri modellemede esnek olmasıdır.

Kantil regresyonun özellikleri;

1. τ 'nın farklı değerlerine bakarak $X=x$ eşitliğini veren Y 'nin tüm şartlı dağılımını karakterize etmek için kantil regresyon kullanılabilir.
2. Kantiller, tekdüze dönüşümlerde eş değişkenlidir. $h(\cdot)$ tekdüze fonksiyon için, $Q_{(h(Y)|X)}(\tau|x) = h_{(Y)|X}(\tau|x)$ 'dir. Örneğin logaritmik gelirin şartlı medyanı, gelirin şartlı meydanının logaritmasıdır.
3. Kantiller, y yönündeki aykırı değerlere karşı sağlamdır.
4. Hata terimleri normal dağılmadığı zaman medyan tahmini, ortalama tahminine göre daha etkin olabilir.
5. Kantil regresyon, değişen varyanslılığın tespitinde kullanılabilir (Leping, 2005).

4.11. Bootstrap (Yeniden Örnekleme)

İstatistik teorisi üç temel soruyu cevaplamak için çalışır.

1. Veri nasıl toplanabilir?
2. Toplanan verileri nasıl analiz etmek ve özetlemek gerekir?
3. Özetlenen veriler ne kadar doğrudur?

Üçüncü soru istatistiksel çıkarsama olarak bilinen sürecin bir parçasını oluşturur. Bootstrap son zamanlarda belirli türde istatistiksel çıkarımlar yapmak için geliştirilmiş bir tekniktir. Klasik istatistik teorisinin karışık hesaplamalarını basitleştirmek için modern bilgisayar gücü gerektirdiğinden son zamanlarda daha geliştirilmiştir (Efron ve Tibsrihani, 1993: 1).

Bradley Efron'un 1979 yılında yayınladığı bootstrap yöntemleri ile ilgili ilk makalesi daha önceki örnekleme yöntemlerinin sentezi ve simülasyona dayalı istatistik analizine yeni bir çalışma alanı oluşturmasıyla istatistik için büyük bir gelişme olmuştur. Bootstrap metotlar; verinin küçük olması, hataların normal dağılmaması gibi varsayımların geçersiz olduğu durumlarda kullanılır (Davison ve Hinkley 1997, ix).

Bootstrap yönteminde temel düşünce eldeki örnekleme, kitle olarak varsayıp buradan belirli sayıda tekrarlı örnekleme yaparak ilgilenilen tahmincinin suni bir örnekleme dağılımını yaratmaktır. Özellikle söz konusu tahmincinin örnekleme dağılımını asimptotik teori ile elde etmek zor ya da olanaksızsa bootstrap yöntemi ya da genel anlamıyla tekrarlı örnekleme yöntemleri güçlü bir potansiyel oluşturmaktadır (Aktükün, 2005).

Bootstrap ve permütasyon istatistikte çok yaygın kullanılır. Yeniden örnekleme günümüzde, güven aralığı, hipotez testleri ve diğer dolaylı problemlerin çözümünde de tercih edilir. Bu yöntem bize kompleks verileri bile analiz etme imkanı verir. Bootstrap sıklıkçı (frequentist) istatistiğe aittir, Bayesian istatistik değildir. Çoğunlukla, bağımsız ve özdeş dağılan veriden yeniden örnekleme yapılarak eşit boyutta belirlenen veri setlerinin oluşturulması şeklinde uygulanır. Bootstrap yöntemin avantajı, kompleks parametre varlığında bile sonucun basit ve anlaşılır olmasıdır. Fakat örneğin bağımsızlığı gibi bazı önemli varsayımlara sahip olması ve garanti olmaması gibi dezavantajları vardır (Sufahani ve Ahmad, 2012).

Bootstrap metodlar, test istatistiklerini karşılaştırmada alternatifler sağlaması ve hesaplama maliyetini düşürmesi sayesinde son yıllarda özellikle ekonometride popüler hale gelmiştir. Bu fikir, çok sayıda simüle edilmiş örnek üretir ve her birinin bootstrap test istatistiklerini hesaplamaya ve daha sonra bootstrap istatistiklerinin ampirik dağılımını belirlemek ve gerçek test istatistiklerini karşılaştırmaya dayanır (Davidson ve MacKinnon, 2006).

Bootstrap örnekler orijinal veriden yerine koyma metodu ile seçilir. $n=4$ birimlik bir veri düşünülürse bu veriden $n^n=4^4=256$ tane bootstrap örnek elde edilir. Fakat bu örnek sayısı bir anda kontrol edilemez hale gelebilir. Örneğin; $n=10$ kadar küçük bir örnek için bootstrap örnek sayısı $10^{10} = 10$ milyara ulaşacaktır ve pratikte bunu birer birer saymak mümkün değildir. Bu nedenle tahminin standart sapması veya ortalamasını hesaplamak için 100 veya 200 bootstrap örnek yeterli olacaktır. Güven aralıklarını hesaplamak içinse 1000 veya 2000 gibi daha fazla bootstrap örneğe ihtiyaç vardır.

Uygulamada bootstrap yöntem aşağıdaki gibi açıklanabilir.

r , bootstrap tekrarlamalarının sayısı olsun. Bu seçilebilecek bootstrap örnek sayısıdır.

Bootstrap sayısı asimptotik olarak $N \rightarrow \infty$ dayanır. Bu nedenle düşük bir r için bile asimptotik olarak geçerli olabilir. Fakat $r \rightarrow \infty$ olduğunda daha kesin sonuç vereceği açıktır (Cameron ve Trivedi, 2005: 361).

n birimlik örnekten rasgele olarak yerine koyma metodu ile yine n gözlemde oluşan bootstrap örnekler $Y_{b1}^*, Y_{b2}^*, \dots, Y_{bn}^*$ ($b=1, \dots, r$) oluşturulur ve bootstrap örnek ortalaması hesaplanır.

$$\bar{Y}_b^* = \frac{\sum_{i=1}^n Y_{bi}^*}{n} \quad (4.42)$$

r bootstrap örneğin bootstrap ortalamasının standart sapmasının tahmini ise

$$SE^*(\bar{Y}^*) = \sqrt{\frac{\sum_{i=1}^n (Y_b^* - \bar{Y}^*)^2}{r-1}} \quad (4.43)$$

$$\bar{\bar{Y}}^* \equiv \frac{\sum_{b=1}^r \bar{Y}_b^*}{r} \quad (4.44)$$

Bir θ parametresinin tahmini $\hat{\theta}$ 'yı tahmin etmek için bootstrap algoritma şu şekildedir.

1. S veri toplama şeklini göstermek üzere popülasyondan örnekler elde edilir.

$S(\text{populasyon}) \Rightarrow \text{örnek}$

$\hat{\theta}$ tahmini gözlenen örneğin bir $s(.)$ fonksiyonudur. Veri toplama yöntemi büyük bir popülasyondan bağımsız rasgele örneklemedir.

2. Populasyonun yerini alan örnekten r-kez tekrarlanarak bootstrap örnekler elde edilir.

$$S(\text{örnek}) \left\{ \begin{array}{l} \Rightarrow \text{bootstrap örnek}_1 \\ \Rightarrow \text{bootstrap örnek}_2 \\ \dots \\ \Rightarrow \text{bootstrap örnek}_r \end{array} \right.$$

3. Her bootstrap örnek için tahmin $\hat{\theta}_b^*$ hesaplanır.

4. $\hat{\theta}$ tahminin örneklem dağılımını tahmin etmek için $\hat{\theta}_b^*$ 'nın dağılımı kullanılır. Örneğin; $\hat{\theta}$ 'nın bootstrap standart hatası;

$$SE^*(\hat{\theta}^*) \equiv \sqrt{\frac{\sum_{b=1}^r (\hat{\theta}_b^* - \bar{\theta}^*)^2}{r-1}}$$

$$\bar{\theta}^* \equiv \frac{\sum_{b=1}^r \hat{\theta}_b^*}{r}$$

Sonuç olarak; Y_{bi}^* bootstrap gözlemler; Y_i orijinal gözlemlere, $\bar{Y}_b^*$ bootstrap ortalama; $\bar{Y}$ orijinal örneğin ortalamasına, orijinal örneğin ortalaması populasyon ortalaması μ 'ye ve bootstrap örnek ortalamasının dağılımı orijinal populasyondan çekilen n birimlik örnek için (bilinmeyen) örnek ortalamasının dağılımına benzer.

Bu yöntem kolaylıkla regresyon modellerine genişletilebilir.

1. Orijinal örnekten regresyon katsayıları tahmin edilir ve her bir gözlem için uydurulan değerler ve artıklar hesaplanır.

$$\hat{Y}_i = A + B_{1 \times i1} + \dots + B_{1 \times in}$$

$$\hat{E}_i = Y_i - \hat{Y}_i$$

2. Artıkların bootstrap örnekleri seçilir. $\hat{e}_b^* = [\hat{E}_{b1}^*, \hat{E}_{b2}^*, \dots, \hat{E}_{bn}^*]'$, ve buradan bootstrap Y değerleri $y_b^* = [Y_{b1}^*, Y_{b2}^*, \dots, Y_{bn}^*]'$ hesaplanır, burada $Y_{bi}^* = \hat{Y}_i + \hat{E}_{bi}^*$ dir.

3. Bootstrap Y deęerleri ve sabit X deęerlerinin regresyonu ile bootstrap regresyon katsayıları edilir.
4. Regresyon katsayıları için gven aralıęı ve bootstrap standart hatalar hesaplanır (Fox, 2008: 597-598).

Mallinckrodt ve dięerleri (2006), bilgisayar teknolojilerinin geliřmesi sonucunda analiz srecinin daha etkin ve hızlı yapılabilirlięinden yola ıkarak bootstrap rnekleminin 1000'den az olmaması kořulunun daha rasyonel bir fikir olacaęı grřndedirler (Burmaoęlu, Polat ve Meydan, 2013).

5. KARARLI DAĞILIMLAR

Kararlı dağılımlar; birçok cazip matematiksel özelliklere sahip kalın kuyruklara ve çarpıklığa izin veren, olasılık dağılımlarının zengin bir sınıfıdır. Bu sınıf 1920'lerde Levy tarafından bağımsız ve özdeş dağılan terimlerin toplamı çalışmasında karakterize edilmiştir. Birkaç dağılım dışında yoğunluk ve dağılım fonksiyonlarının kapalı formlarının olmaması, geçmişte uygulayıcılar açısından büyük engel oluşturmıştır. Ancak günümüzde kararlı yoğunlukları, dağılım fonksiyonları ve kantilleri hesaplamak için güvenilir bilgisayar programları vardır.

Kararlı dağılımlar birçok fiziksel ve ekonomik sistem için bir model olarak önerilmiştir. Bir sistemin tanımlanmasında kararlı dağılımların kullanılmasının birçok nedeni vardır. İlki; normal olmayan kararlı bir model beklemek için sağlam teorik nedenlerin olmasıdır. Örneğin; dönen aynanın yansıması Cauchy dağılımı, Brown Hareketi için çarpma sayısı Levy dağılımına, yıldızların yerçekimi alanı Holtsmark dağılımına uyar. İkincisi; aynı dağılıma sahip bağımsız terimlerin normalleştirilmiş toplamalarının tek mümkün limiti olduğunu ifade eden Genelleştirilmiş Merkezi Limit Teoremi, gözlenen niceliklerin birçok küçük terimin toplamı olduğunu öne sürer. Örneğin; bir menkul kıymetin günlük, haftalık ya da aylık değişimleri işlem den işleme değişimlerin toplamıdır ve böyle sistemleri tanımlamak için kararlı dağılımlar kullanılmalıdır. Üçüncüsü ise; deneysel birçok veri setinin kalın kuyruklu ve çarpık yapı göstermesidir. Bu özellikler ve Merkezi Limit Teoremi ile birlikte güçlü deneysel kanıtlar, kararlı dağılımların kullanımını savunmak için kullanılır (Nolan, 2005).

Genellikle uygulamalarda Gauss dağılımının kullanımını savunmak için kullanılan Merkezi Limit argümanı aynı zamanda Gauss olmayan kararlı modelin seçimini desteklemek içinde kullanılabilir. Gauss ve Gauss olmayan kararlı dağılımlar arasındaki en önemli fark kararlı dağılımların kalın kuyruklu, her zaman sonsuz varyanslı ve bazı durumlarda birinci momentlerinin sonsuz olmasıdır. Bir diğer fark Gauss dağılım için ortalamanın simetrik olması zorunluyken, Gauss olmayan dağılımların asimetriye veya çarpıklığa olanak tanımasıdır. O halde kalın kuyruklu modeller veya asimetrisinin olduğu bazı uygulamalarda kararlı model uygun bir aday olabilir (Janicki ve Weron, 1994).

Eğer bağımsız, aynı dağılıma sahip rasgele değişkenlerin toplamı bir limit dağılımına sahipse, bu limit dağılımı kararlı sınıfın bir üyesidir. Kararlı dağılımların en önemli özelliği toplama altında değişmez olmalarıdır. Yani aynı alfa karakteristik üssüne sahip bağımsız kararlı değişkenlerin toplamı, aynı karakteristik üsse sahip kararlı dağılımı verir (Fama ve Roll, 1968).

5.1. Tanım: Normal rasgele değişkenlerin önemli bir özelliği, normal iki rasgele değişkenin toplamının yine normal rasgele değişken olmasıdır. Bunun sonucu olarak X normal dağılıyorsa X_1 ve X_2 ve herhangi pozitif a ve b sayıları için

$$aX_1 + bX_2 = cX + d \quad d \in \mathbb{R} \text{ ve } c \text{ bir pozitifdir.} \quad (5.1)$$

(d sembolü eşitliğin aynı dağılıma sahip olduğunu ifade eder.)

Bir rasgele X değişkeni için Eş. 5.1 sağlanıyorsa X - kararlıdır denir. Eğer a ve b 'nin tüm seçimleri için bu eşitlik $d=0$ içinde sağlanıyorsa X tesadüfi değişkenine “kesin kararlıdır” denir. Eğer X değişkeni kararlı ve 0 etrafında simetrik dağılıyorsa yani $X \stackrel{d}{=} -X$ oluyorsa “simetrik kararlıdır” denir (Nolan, 2009).

5.2. Tanım: Aynı dağılıma sahip $X_1, X_2, \dots, X_n$ rasgele değişkenlerinin ancak ve ancak $n \geq 2$ için

$$X_1 + X_2 + \dots + X_n = c_n X + d_n \quad (5.2)$$

Eş. 5.2'yi sağlayan bir pozitif c_n ve $d_n \in \mathbb{R}$ varsa, kararlı dağılım ailesine aittir denir (Zolotarev, 1986: 12-28).

5.3. Tanım: Bir X değişkeninin çekim alanı varsa, bağımsız ve özdeş dağılan $Y_1, Y_2, \dots, Y_n$ rasgele değişkenleri ve pozitif tanımlı $\{d_n\}$ dizisi ile gerçek sayılarda tanımlı $\{a_n\}$ dizisi varsa X rasgele değişkenin kararlı dağılıma sahip olduğunu söylenir. Şöyle ki;

$$\frac{Y_1 + Y_2 + \dots + Y_n}{d_n} + a_n \stackrel{d}{\Rightarrow} X \quad (5.3)$$

$\Rightarrow$ notasyonu dağılımın yakınsadığını belirtir. X , Gauss dağılımına sahip ve bağımsız ve özdeş dağılan Y_i 'ler sınırlı varyansa sahip olduğunda Eş. 5.3 sıradan merkezi limit teoremini ifade eder.

5.4. Tanım: Bir X rasgele değişkeninin parametreleri $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$ ve aşağıdaki gibi bir karakteristik fonksiyona sahipse;

$$E(e^{iuX}) = \begin{cases} \exp\left(-\sigma^\alpha |\theta|^\alpha \left[1 - i\beta(\text{sign}\theta) \left(\tan \frac{\pi\alpha}{2}\right)\right] + i\mu\theta\right) & , \alpha \neq 1 \\ \exp\left(-\sigma |\theta| \left[1 + i\beta \frac{2}{\pi} (\text{sign}\theta) \ln(|\theta|)\right] + i\mu\theta\right) & , \alpha = 1 \end{cases} \quad (5.4)$$

X değişkeni kararlı dağılıma sahiptir denir. Kararlı dağılımların yüksek değişkenliği modelleme için önemli bir rol oynamalarının sebeplerinden biridir. Herhangi bir kararlı rasgele değişken varış zamanlarını kapsayan Poisson sürecinin sonsuz toplamı olarak ifade edilebilir. Bu şekilde bir gösterimde α kararlılık indeksi rol oynar. Benzer gösterimler kararlı rasgele vektörler ve kararlı stokastik süreçler içinde kurulabilir (Samorodnitsky ve Taqqu, 1994: 2,5).

Kararlı dağılımlar sınıfının tamamı α , β , γ ve δ olmak üzere 4 parametre tarafından dizayn edilirler.

α -parametresi; karakteristik üs olarak da adlandırılan bu parametre kuyruğun kalınlığına karar verir. $\alpha \in (0,2]$ aralığında tanımlıdır. Parametrenin değeri 0'a yaklaştıkça dağılımın kuyruğu kalınlaşır. $\alpha = 2$ ise dağılım normal dağılıma eşittir. $0 < \alpha \leq 2$ olduğunda dağılımın ikinci momenti ya da ortalaması mevcuttur. Fakat $\alpha < 1$ olduğunda ortalama sonsuzdur.

β -parametresi; dağılımın asimetrisinin bir ölçüsüdür. $\beta \in [-1,1]$ aralığında tanımlıdır. $\beta = 0$ ise dağılım simetriktir. $\beta < 0$ ise dağılım sola, $\beta > 0$ ise sağa çarpıktır. $|\beta|$ 'nın değeri 1'e yakınsarken çarpıklığın derecesi artmaktadır.

γ ölçek parametresi; dağılımın yayılımının bir ölçüsüdür. Ölçek parametresi tüm kararlı dağılımlar için sonlu değerler aldığı halde $\alpha < 2$ 'nin tüm değerleri için sonsuz olur.

δ konum parametresi; $\delta \in (-\infty, \infty)$ aralığında tanımlanır. $\alpha > 1$ olduğunda dağılımın ortalamasına, $\alpha \leq 1$ olduğunda ise medyana eşittir.

Ölçek parametresi γ , $\gamma > 0$ için δ etrafında dağılımı genişletir veya daraltırken konum parametresi δ dağılımı sağa veya sola kaydırır.

Bir kararlı rasgele değişkenin kümülatif dağılım fonksiyonu X şöyle yazılabilir.

$$P(X < x) = S(x; \alpha, \beta, \gamma, \delta) = S((x - \delta)/c; \alpha, \beta, 1, 0) \quad (5.5)$$

Bu eşitlik şöyle basitleştirilebilir;

$$s(x; \alpha, \beta, \gamma, \delta) = (1/c) s_{\alpha, \beta}((x - \delta)/c) \quad (5.6)$$

Eğer bir rasgele değişken $S(x; \alpha, \beta, \gamma, \delta)$ dağılımına sahipse $X \sim S(\alpha, \beta, \gamma, \delta)$ olur (McCulloch, 1998).

Yoğunluklarına ait kapalı formlarının yazılabileceği 3 kararlı dağılım; Normal, Cauchy ve Levy dağılımlarıdır.

X tesadüfi değişkenine ait yoğunluk fonksiyonu;

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (-\infty < x < \infty) \quad (5.7)$$

ise o zaman $x \sim N(\mu, \sigma^2)$ dir.

Normal dağılım $\alpha=2$, $\beta=0$ parametreleri ile kararlıdır.

X tesadüfi değişkenine ait yoğunluk fonksiyonu;

$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2} \quad (-\infty < x < \infty) \quad (5.8)$$

ise o zaman $X \sim \text{Cauchy}(\gamma, \delta)$

Fizikte Lorenz dağılımı olarak da bilinen Cauchy dağılımı $\alpha=1$, $\beta=0$ parametreleri ile kararlı dağılır.

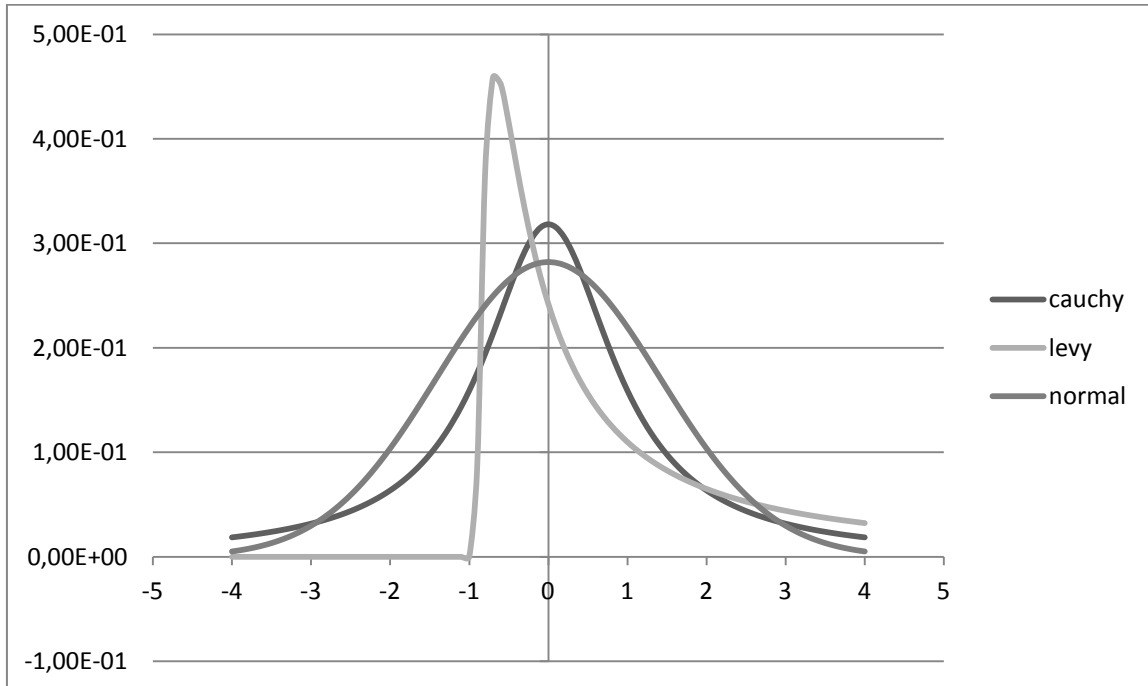
X tesadüfi değişkenine ait yoğunluk fonksiyonu;

$$f(x) = \sqrt{\frac{\delta}{2\pi}} \frac{1}{(x - \delta)^{3/2}} \exp\left(-\frac{\gamma}{2(x - \delta)}\right) \quad (-\delta < x < \infty) \quad (5.9)$$

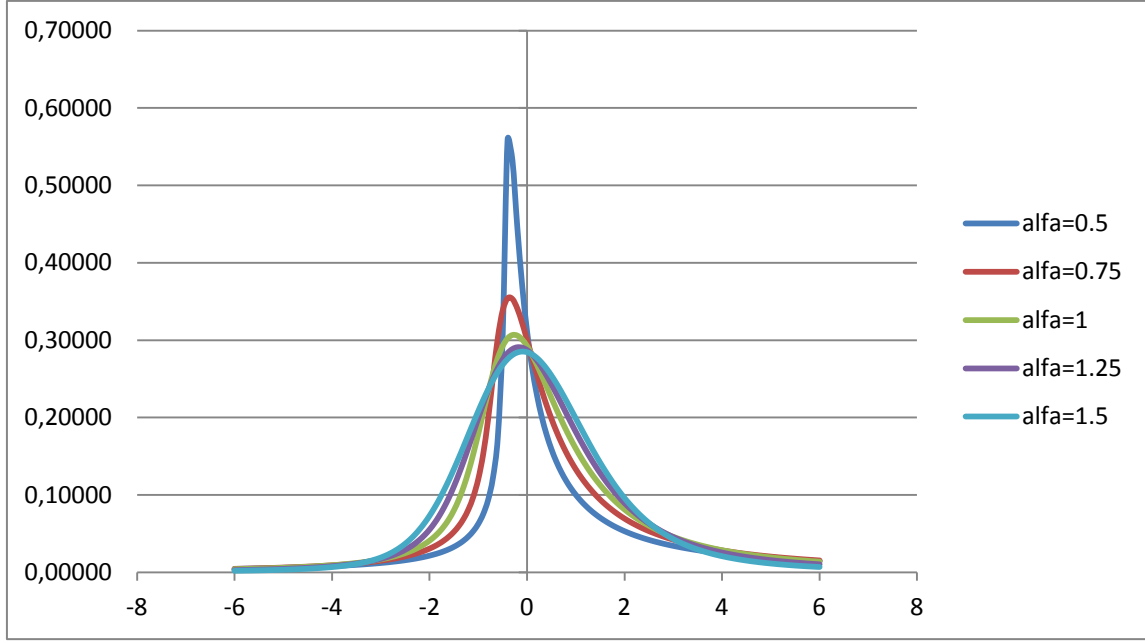
ise $X \sim \text{Levy}(\gamma, \delta)$ olmaktadır.

$\alpha=1/2$, $\beta=1$ parametreleri ile kararlı dağılmaktadır.

Hem normal dağılım hem de Cauchy dağılımı simetrik ve çan şeklindeki eğrilerdir. Bu iki dağılım arasındaki temel fark, Cauchy dağılımının daha kalın kuyruklara sahip olmasıdır. Özellikle standart normal dağılım 3'ün üzerinde değer alma olasılığı çok az olduğu halde Cauchy dağılımının 3'ün üzerinde değere sahip olma olasılığı daha yüksektir. Bu iki dağılımdan gelen örneklemelerde 3'ün üzerindeki değerler Cauchy dağılımında standart normal dağılıma göre ortalama 100 kat daha fazla olacaktır. Bu sebeple kararlı dağılımlar kalın kuyruklu olarak adlandırılırlar. Normal ve Cauchy dağılımının aksine Levy dağılımı oldukça çarpıktır ve dağılımın sahip olduğu olasılığın tamamı $x>0$ üzerine yoğunlaşmaktadır. Levy dağılımı da Cauchy dağılımına göre daha kalın kuyrukludur (Nolan, 2009).



Şekil 5.1. Standart Normal $N(0,1)$, Cauchy $(1,0)$, Levy $(1,0)$ yoğunluklarının grafikleri



Şekil 5.2. $S_\alpha(0,5;1;0)$ parametrelemesinde, $\alpha=0,5; 0,75; 1; 1,25$ ve $1,50$ için kararlı yoğunluklar

5.1. Kararlı Dağılımların Parametrelendirilmesi

Kararlı dağılımlar için birçok parametrelendirme söz konusudur ve bu durum bazı karışıklıklara sebebiyet vermektedir. Parametrelenedeki çeşitlilik, tarihi gelişimden ve birçok problemin kararlı dağılımların özelleştirilmiş biçimlerinin kullanılarak analiz edilmesinden kaynaklanmaktadır. Farklı durumlarda farklı parametrelmelerin kullanılmasının geçerli sebepleri vardır. Sayısal çalışma veya veri modellemek için ayrı, dağılımın basit cebirsel özellikleri gerektiğinde ayrı ve kararlı kuramın analitik özellikleri çalışılmak istendiğinde ayrı bir parametrelendirme tercih edilebilir.

Kararlı dağılımlar için son zamanlarda en çok $S_\alpha(\sigma, \beta, \mu)$ notasyonu kullanılır. Bu notasyonu $S(\alpha, \beta, \gamma, \delta; k)$ formunda kullanmak daha çok tercih edilir. Bunun üç sebebi vardır. İlki; genel gösterimde α farklı ve sabit olarak belirlenir. Uygulamada $\alpha, \beta, \gamma, \delta$ bilinmemektedir ve bunların tahminine ihtiyaç vardır. Yeni notasyon bunu vurgular. İkincisi; ölçek parametresi, standart sapma değildir ve konum parametresi de genelde ortalama değildir. Bu yüzden ölçek için σ yerine γ ve konum için μ yerine δ parametresi kullanılır. Üçüncüsü farklı parametreler arasındaki fark net olmalıdır, k bunun içindir. Kararlı dağılım kullanıcılarının hangi parametrelendirmeyi kullandıklarını açıkça belirtmeleri gerekir (Nolan, 2009).

Kararlı dağılımların en az on farklı parametrelendirmesi vardır. $S(\alpha, \beta, \gamma, \delta_0; 0)$ ve $S(\alpha, \beta, \gamma, \delta_1; 1)$ tarafından gösterilen parametrelendirmeleri tanımlarsak; ilki daha iyi sayısal davranış ve sezgisel anlamlara sahip olduğundan tüm uygulamalarda kullanılır, ikincisi ise daha çok literatürde kullanıldığından bu parametremeyi anlamak çok önemlidir. İki parametrelende α , β ve γ aynı anlamda iken fark sadece konum parametresindedir. İkisini ayırt etmek için hangi parametremenin olduğunu belirten bir alt simge kullanılır. $S(\alpha, \beta, \gamma, \delta_0; 0)$ için δ_0 , $S(\alpha, \beta, \gamma, \delta_1; 1)$ için δ_1 kullanılır.

$X \sim S(\alpha, \beta, \gamma, \delta_0; 0)$ bir rasgele değişken ise sahip olduğu karakteristik fonksiyonu;

$$E(e^{iuX}) = \begin{cases} \exp\left(-\gamma^\alpha |u|^\alpha \left[1 + i\beta \left(\tan \frac{\pi\alpha}{2}\right) (\text{signu})(|\gamma u|^{1-\alpha} - 1)\right] + i\delta u\right), & \alpha \neq 1 \\ \exp\left(-\gamma |u| \left[1 + i\beta \frac{2}{\pi} (\text{signu}) \ln(\gamma |u|)\right] + i\delta u\right) & , \alpha = 1 \end{cases} \quad (5.10)$$

Eğer; $X \sim S(\alpha, \beta, \gamma, \delta_1; 1)$ bir rasgele değişken ise sahip olduğu karakteristik fonksiyonu;

$$E(e^{iuX}) = \begin{cases} \exp\left(-\gamma^\alpha |u|^\alpha \left[1 - i\beta \left(\tan \frac{\pi\alpha}{2}\right) (\text{signu})\right] + i\delta u\right) & , \alpha \neq 1 \\ \exp\left(-\gamma |u| \left[1 + i\beta \frac{2}{\pi} (\text{signu}) \ln(|u|)\right] + i\delta u\right) & , \alpha = 1 \end{cases} \quad (5.11)$$

Konum parametreleri şöyle ilişkilendirilir;

$$\delta_0 = \begin{cases} \delta_1 + \beta \gamma \tan \frac{\pi\alpha}{2}, & \alpha \neq 1 \\ \delta_1 + \beta \frac{2}{\pi} \gamma \ln \gamma, & \alpha = 1 \end{cases} \quad \delta_1 = \begin{cases} \delta_0 + \beta \gamma \tan \frac{\pi\alpha}{2}, & \alpha \neq 1 \\ \delta_0 + \beta \frac{2}{\pi} \gamma \ln \gamma, & \alpha = 1 \end{cases} \quad (5.12)$$

Dikkat edilirse; $\beta=0$ olduğunda parametreler çakışır. $\alpha \neq 1$ ve $\beta \neq 0$ olduğunda parametreler $\alpha \rightarrow 1$ giderken sonsuza giden $\beta \gamma \tan \frac{\pi\alpha}{2}$ değişimiyle farklılaşır. Özellikle $S(\alpha, \beta, \gamma, \delta_1; 1)$ için yoğunluk $\alpha \rightarrow 1$ giderken ∞ veya $-\infty$ gitme eğilimindedir. α ; 1'e yakın olduğunda bu aralıkta kararlı yoğunlukları ve kümülatifleri hesaplamak zordur ve parametre tahminleri güvenilirmezdir. $\alpha=1$ olduğunda sıfır parametremesi basit bir ölçek verirken, 1 parametremesinde bu mevcut değildir. Uygulamalardan hareketle tüm dört parametrenin ortak olduğu $S(\alpha, \beta, \gamma, \delta_0; 0)$ kullanmak tercih edilir. $S(\alpha, \beta, \gamma, \delta_1; 1)$ parametremesi kullanıldığında argümanlar cebirsel olarak basittir. Her iki parametremenin kullanılması kaçınılmaz görüldüğünden kullanıcılar hangisini kullandıklarını açıkça belirtmelidirler (Nolan, 2003).

$\alpha=2$ ve $\beta=0$ ile kararlı normal dağılımı $N(\mu, \sigma^2)=S(2,0,\sigma/\sqrt{2},0;0)=S(2,0,\sigma/\sqrt{2},0;1)$

$\alpha=1$ ve $\beta=0$ ile kararlı Cauchy dağılımı $\text{Cauchy}(\gamma, \delta)=S(1,0,\gamma, \delta;0)=S(1,0,\gamma, \delta;1)$

$\alpha=1/2$ ve $\beta=1$ ile kararlı Levy dağılımı için $\text{Levy}(\gamma, \delta)=S(1/2,1,\gamma, \gamma+\delta;0)=S(1/2,1,\gamma, \delta;1)$ olur

(Nolan, 2003).

$S(\alpha, \beta, \gamma, \delta; k)$ için yoğunluk fonksiyonu

$$f(x|\alpha, \beta, \gamma, \delta; k) = \begin{cases} [\delta - \gamma \tan \frac{\pi \alpha}{2}, \infty) & \alpha < 1, \beta = 1, k = 0 \\ (-\infty, \delta + \gamma \tan \frac{\pi \alpha}{2}] & \alpha < 1, \beta = -1, k = 0 \\ [\delta, \infty) & \alpha < 1, \beta = 1, k = 1 \\ (-\infty, \delta] & \alpha < 1, \beta = -1, k = 1 \\ (-\infty, \infty) & \text{aksi durumlarda} \end{cases} \quad (5.13)$$

Eş. 5.13'teki gibi verilir (Nolan, 2005).

5.2. Kararlı Dağılımların Parametre Tahmini

Son zamanlara kadar hesaplama zorluklarından dolayı pratikte kararlı dağılımları kullanmak zordu. Günümüzde bu zorlukların çoğu kararlı yoğunlukları, kümülatif dağılım fonksiyonları ve nicelikleri hesaplayabilen STABLE programı ile ortadan kaldırılmıştır. Programda kullanılan temel metot 1997'de Nolan tarafından tanımlanmıştır. Daha sonra program Chambers, Mallows ve Stuck (1976), kararlı rasgele değişkenlerle simülasyon metodunu kapsayan kararlı parametre tahminleri geliştirilmiştir.

Kararlı dağılımlar ile ilgili bir diğer problem sadece $p < \alpha$ olduğunda momente sahip olan $\alpha \in (0,2)$ karakteristik üssü ile ilgilidir. Özellikle, normal dağılıma durumu dışında bu rasgele değişkenlerin varyansı sonsuzdur. Bu durum iyi asimptotik özelliklerden yoksun olmakla beraber, iyi güven aralıkları ve doğru çıkarımlar yapmak oldukça zordur. Uygulamalarda en çok ilgilenilen parametre kuyruğun kalınlığına karar veren alfa karakteristik üssüdür.

Önerilen α tahmin yöntemleri genel olarak üç kategoride toplanır. Bunlar; en çok olabilirlik, kantil metotları ve karakteristik fonksiyonuna dayanan metotlardır.

“Hill tahmini” adı verilen metoda dayanan birçok tahmin vardır. Pictet, Dacorogna ve Muller (1998), Pickands tahmini (Pickands, 1975, Dekkers ve De Haan, 1989), Hill tahmini, De Haan ve Resnick tahmini (1980) ve modifiye Hill tahmini (Dekkers ve diğerleri, 1990) gibi dört tahminin performanslarını bir simülasyon çalışması ile detaylandırmışlardır. Hill tipi tahmin ediciler örnekteki k . en büyük gözleme dayanır. Bu tahmin edici iyi asimptotik özelliklere sahiptir fakat pratikte yeterli değildir. Bu tahmin ediciler k 'nın seçimi ile ilgili önemli farklılıklara sahiptir. DuMouchel (1973), en çok olabilirlik tahmini tipi bir algoritma sunmuştur. Algoritma teoride üstün olmasına rağmen hesaplamada zahmetlidir.

Kantil metotları, sıra istatistiklerine dayanan ve hesaplaması kolay metotlardır. İlk olarak Fama ve Roll (1978) tarafından sunulan bu metot, daha sonra McCulloch (1986) tarafından geliştirilmiştir. McCulloch tahminleri $\alpha \in (0,6;0,2)$ aralığında çalışmıştır. Koutrouvelis (1981), karakteristik fonksiyonu kullanmanın en iyi performans gösteren teknik olduğunu göstermiştir. Bu metodun geliştirilmiş versiyonu Kogon ve Williams (1998) tarafından sunulmuştur (Jammalamadaka ve Iyer, 2003).

McCulloch α ve β 'nın tahmininde yüzdelikleri kullanır. $S(x;\alpha,\beta,\gamma,\delta)$ kararlı dağılımından çekilmiş n birimlik x_i örneği olsun. x_p p . yığın yüzdeliği ise $\hat{x}_p$ örnek yüzdeliğidir ve x_p 'nin tutarlı bir tahmin edicisidir.

$$v_\alpha = \frac{x_{.95} - x_{.05}}{x_{.75} - x_{.25}}$$

$$v_\beta = \frac{x_{.95} - x_{.05} - 2x_{.5}}{x_{.95} - x_{.05}} \quad (5.14)$$

Örnek değerleri ise;

$$\hat{v}_\alpha = \frac{x_{.95} - x_{.05}}{x_{.75} - x_{.25}}$$

$$\hat{v}_\beta = \frac{x_{.95} - x_{.05} - 2x_{.5}}{x_{.95} - x_{.05}} \quad (5.15)$$

α ve β 'nin bir fonksiyonu olan v_α ve v_β , γ ve δ 'dan bağımsızdır.

$$v_\alpha = \Phi_1(\alpha, \beta)$$

$$v_\beta = \Phi_2(\alpha, \beta) \quad (5.16)$$

Bu eşitlikler ters çevrilerek v_α ve v_β cinsinden yazılır ise

$$\alpha = \varphi_1(v_\alpha, v_\beta)$$

$$\beta = \varphi_2(v_\alpha, v_\beta) \quad (5.17)$$

O halde α ve β 'nın tutarlı tahmin edicileri;

$$\hat{\alpha} = \varphi_1(\hat{v}_\alpha, \hat{v}_\beta)$$

$$\hat{\beta} = \varphi_2(\hat{v}_\alpha, \hat{v}_\beta) \quad (5.18)$$

Ölçek parametresinin tahmini ise;

$$v_\gamma = \frac{x_{.75} - x_{.25}}{\gamma}$$

$$v_\gamma = \Phi_3(\alpha, \beta) \quad (5.19)$$

eşitliklerinin yardımı ile şöyle hesaplanır;

$$\hat{\gamma} = \frac{\hat{y}_{.75} - \hat{y}_{.25}}{\Phi_3(\hat{\alpha}, \hat{\beta})} \quad (5.20)$$

Konum parametresi için;

$$v_{\delta} = \frac{\delta - x_{.5}}{\gamma}$$

$$v_{\delta} = \Phi_4(\alpha, \beta) \quad (5.21)$$

eşitlikleri ile tahmin;

$$\hat{\delta} = \hat{\gamma} \cdot \Phi_4(\hat{\alpha}, \hat{\beta}) \cdot \hat{x}_{.5} \quad (5.22)$$

ile bulunur.

$\hat{\delta}$, $\hat{\gamma}$ tahmin değerleri, v_{δ} , v_{γ} , α ve β değerlerine göre McCulloch (1986), çalışmasında mevcut tablolar yardımıyla bulunur (McCulloch, 1986).

6. KARARLI REGRESYON

Hata terimleri normal dağıldığında en çok olabilirlik çözümü ve EKK çözümü aynıdır, fakat normal dağılmadığında iki metot arasında fark vardır. Sağlam tahmin metotları aykırı değerleri ortadan kaldıran ya da önemini azaltan bazı teknikler kullanır. Bir sağlam tahmin metoduna karşın, sağlam modelleme hem regresyon katsayıları ile hem de hata dağılımının uyumu ile ilgilenir. Ölçülemeyen terimler eğer sonlu varyansa sahipse o zaman Merkezi Limit Teoremi hata terimlerinin yaklaşık olarak normal olacağını söyler. Fakat ölçülemeyen terimler kalın kuyruklu ise Genelleştirilmiş Merkezi Limit Teoremine göre hata terimleri tahminen kararlı olacaktır (Nolan ve Ojeda Revah, 2013).

Kararlı dağılımla en çok karşılaşılan ekonometride sıklıkla basit doğrusal regresyon modeli kullanılır.

$$y_j = BX_j + e_j \quad j=1,2,\dots,T \quad (6.1)$$

y_j ; j . bağımlı değişken, X stokastik olmayan bağımsız değişken, B bir parametre ve e_j , 0 ortalamalı bağımsız ve özdeş dağılan rasgele değişkenlerdir. Ekonomistler B parametresini EKK metodunu kullanarak tahmin ederler. Hesaplamadaki kolaylığının yanı sıra bu metot, Merkezi Limit Teoremi ve Gauss-Markov Teoreminin varlığında popülerlik kazanmıştır. e_j 'ler Eş. 6.1'i sağlıyor ve sınırlı varyanslı bir dağılıma sahipse Gauss-Markov Teoremini ifade eder ve o zaman EKK tahmini B 'nin tüm doğrusal yansız tahminleri, minimum varyansa sahiptir. Merkezi Limit Teoremi hipotez testi için bir temel oluşturduğundan önemlidir. Bu durumda sınırlı varyanslı her bir dağılımın bağımsız olarak dağılmış çok sayıda toplamı normal dağılma eğilimi gösterecektir. Göz ardı edilen bağımsız değişkenlerin çok sayıda toplamının y_j üzerindeki etkisini göstermek için Eş. 6.1'deki e_j 'lerin söz konusu olduğu hatırlandığında bu teoremin uygunluğu görülür. Göz ardı edilen bu değişkenler sonlu varyansla dağılıyorsa, eklenmeleri durumunda e_j 'ler normal dağılacaktır. Bu her zamanki F ve t testlerini yapmamıza izin verir. Bu aynı zamanda, En Küçük Kareler yönteminin en çok olabilirlik tahmini olduğu anlamına gelir.

e_j veya e_j 'lerin temelini oluşturan değişkenlerin sınırlı varyansa sahip olduğu varsayımı EKK kullanımında önemli bir rol oynar. Sonsuz varyanslı e_j 'ler tarafından etkileri özetlenen, göz ardı edilen değişkenler varsa, MLT nin varsayımlarını tutturmak başarısız

olacak ve e_j 'ler normal dağılmayacaktır. Dahası e_j 'lerin sonsuz bir varyansla karakterize edilmesi muhtemeldir. Bu nedenle Gauss-Markov teoreminin minimum varyansı açıkladığı kısmı veya en küçük karelerin etkinlik özelliği artık uygulanamaz. Bu nedenle sonsuz varyanslar, hem EKK 'nın etkinliği hem de normal dağılımın uygulanabilirliği hipotez testi için kaybolur (Blattberg ve Sargent, 1971).

Önceleri, kuyruklara uzak birtakım gözlemlerin aykırı değer olarak çıkarılması nedeniyle birçok sürecin yaklaşık normal olduğu düşünülürdü. Fakat bütün gözlemler ele alındığında birçok süreçte kuyruklarda normal dağılımdan daha fazla olasılık kümelenir. Bu tür kalın kuyruk davranışı, kararlı olmayan normal dağılımın ve diğer bazı dağılımların tipik özelliğidir. Sürecin davranışının kararlı dağılım ile açıklanmasının avantajı, değişkenlerin doğrusal bileşenlerinin kararlı dağılım varsayımı altında analiz edilmesinin daha kolay olmasıdır (bu her zaman geçerli değildir). Birkaç tane bağımsız değişken söz konusu olduğunda, çok değişkenli kararlı dağılımlar gerekmektedir (Press, 2005: 150,151).

e 'nin elemanları bağımsız ve özdeş normal olmayan α kararlı dağılımla dağılıyorsa o zaman EKK tahmini B, α kararlı dağılıma sahiptir. Böylece standart EKK'nın çıkarımları geçerli değildir. Logan ve ark. (1978), t-istatistiklerinin asimptotik özelliklerini aşağıdaki gibi ortaya koymuşlardır.

1. Dağılım fonksiyonun kuyruğu $\pm\infty$ 'da normaldir.
2. $0<\alpha<1$ için ± 1 'de $|1 \mp x|^{-\alpha}$ ve $\beta \neq \pm 1$ için sonsuz tekilliklere sahiptir ve $1<\alpha<2$ olduğunda dağılım ± 1 'de tepeye ulaşır.
3. $\alpha \rightarrow 2$ iken dağılım normale eğilim gösterir ve tepe kaybolur.

DuMouchel (1971, 1973, 1975), belli ön koşullara bağlı olarak α -kararlı dağılım parametrelerinin en çok olabilirlik tahmininin, en çok olabilirlik tahmincisinin genel asimptotik özelliklerine sahip olduğunu göstermiştir.

$e_i = y_i - \sum_{j=1}^k x_{ij}\beta_j$ varsayarsak hata terimleri $(\alpha, \beta, \gamma, \delta)$ parametreleriyle α -kararlı dağılıma sahiptir. α -kararlı dağılımların yoğunluk fonksiyonu $s(\alpha, \beta, \gamma, \delta)$ ile gösterilir ise o zaman e_i 'lerin yoğunluk fonksiyonu şöyle gösteririz:

$$s(e_i, \alpha, \beta, \gamma, \delta) = \frac{1}{\gamma} s\left(\frac{y_i - \sum_{j=1}^k x_{ij}\beta_j}{\gamma}, \beta, 1, 0\right) \quad (6.2)$$

Beklendiği gibi EKK tüm gözlemlere eşit ağırlıklar verir. α -kararlı süreç için tahmin edici dağılım merkezine yüksek değerlere küçük ağırlıklar verir, α azalırken bu etki artar. Bu sonuç, α -kararlı dağılımların konumunu ölçerken budanmış ortalamaları bir Monte Carlo çalışmasıyla tamamlayan Fama ve Roll (1968) tarafından elde edilmiştir. Bu çalışma ile şu sonuçları bulmuşlardır: $\alpha=1,1$ ve $0,25$ budamayla tüm n 'ler için ortalamalar baskın durumdadır. $\alpha=1,3$ ve $1,5$ ise $0,50$ budamayla en iyidir ve $\alpha=1,9$ olduğunda $0,75$ budamayla tahminler diğer tahminlere göre daha az düzgün dağılır. Son olarak Normal dağıldığında ($\alpha=2$ olduğunda) ortalama en iyi tahmin edicidir. Tabi aynı zamanda minimum varyanslı ve yansızdır (Frain, 2006).

Eş. 6.1'de ki hata terimleri e_i 'ler daha az ya da çok sayıda bağımsız gözlenmeyen değişkenin toplamı olarak düşünülebilir. Merkezi Limit Teoremine göre bağımsız ve özdeş dağılan rasgele değişkenlerin toplamı sınırlı bir dağılıma sahipse bu sınırlı dağılım kararlı sınıfın bir üyesi olmalıdır (Zolaterov, 1986: 12-28).

Bağımsız ve özdeş dağılan hata terimlerinin 0 medyan ve γ ölçeği ile simetrik kararlı dağıldığını varsayarsak $e_i \sim s(\alpha, 0, \gamma, 0)$ demektir. DuMouchel (1973,1975) bu durumda en çok olabilirlik tahmininin etkin ve asimptotik normal tahminler verdiğini göstermiştir. Asimptotik varyanslar Fisher'ın bilgi matrisi tarafından kontrol edilir veya log olabilirlik fonksiyonun Hessian matrisi tarafından eş zamanlı hesaplanır. Konum parametresinin en çok olabilirlik tahmini $n^{-1/2}$ oranında doğru parametreye yakınsar. $\alpha < 2$ olduğunda katsayılar vektörünün tahmininde EKK yerine en çok olabilirlik tahmini kullanılır. $\alpha > 1$ olmak koşuluyla EKK örnek büyüdükçe doğru parametreye $n^{1/\alpha-1}$ oranında yakınsar. En çok olabilirlik ile $n^{-1/2}$ oranına ulaşılabilir. EKK tamamen kullanışsız değildir fakat en çok olabilirlik tahminine göre 0 asimptotik etkinliğe sahiptir. Bozulmalar kalın kuyruklu olduğunda EKK etkin değildir. $f(e)$; Eş. 6.1'deki hata terimlerinin olasılık yoğunluk fonksiyonudur. Simetrik kararlı durumlar için $f(e) = s_{\alpha,0}(e/\gamma)/\gamma$ olur. Gauss veya normal dağılım en bilinen ve kolay anlaşılır kararlı dağılımlardır, bu nedenle geçmişte doğrusal regresyon modelinde hataların kontrolünde her zaman kabul edilmişlerdir. Fakat Student'in gösterdiği gibi analizlerin birçoğu aykırı değerler yani, büyük hataları gösteren gözlemler içerir. Normal olmayan kararlı dağılımların hepsi normalden daha kalın kuyruklara sahiptir ve bu yüzden aykırı değerlerin varlığı bulunduğu ve gözlenmeyen değerlerin toplamı için normal dağılımın uygun genişletmeleri vardır (McCulloch, 1998).

7. UYGULAMA

Bu bölümde amaç, hata terimlerinin alfa kararlı dağılıma sahip olduğu basit doğrusal regresyon modeline sağlam tahmin yöntemlerini uygulamak ve bu yöntemlerin performanslarını değerlendirmektir. Aynı zamanda kararlı dağılımın farklı parametre değerleri için sonuçların nasıl bir değişim gösterdiğini izlemektir.

$n=15, 20, 30$ ve 40 olmak üzere dört örnek çapı düzeyi oluşturulmuştur. Büyük örnek çapı için LTS 'in hesaplanmasındaki güçlük örnek çapının artırılmasına imkan vermemiştir. Her örnek için hata terimlerinin on iki farklı olasılık dağılımı kullanılmıştır. Bu bağlamda dağılımlar; $\alpha=0,7, \beta=0$ simetrik kararlı dağılım, $\alpha=1, \beta=0$ simetrik kararlı dağılım, $\alpha=1,3, \beta=0$ simetrik kararlı dağılım, $\alpha=1,7, \beta=0$ simetrik kararlı dağılım, $\alpha=0,7, \beta=1$ simetrik olmayan kararlı dağılım, $\alpha=0,7, \beta=-1$ simetrik olmayan kararlı dağılım, $\alpha=1, \beta=1$ simetrik olmayan kararlı dağılım, $\alpha=1, \beta=-1$ simetrik olmayan kararlı dağılım, $\alpha=1,3, \beta=1$ simetrik olmayan kararlı dağılım, $\alpha=1,3, \beta=-1$ simetrik olmayan kararlı dağılım, $\alpha=1,7, \beta=1$ simetrik olmayan kararlı dağılım, $\alpha=1,7, \beta=-1$ simetrik olmayan kararlı dağılımdır.

(x,y) ikilisinden oluşan veri setinde baz alınacak doğrusal model $y_i = 2 + x_i + e_i$ olarak seçilmiştir. $x_i=1,2,3,\dots,n$ olmak üzere ardışık seriden üretilmiş bir vektördür. Bu durumda kesişim parametresi $\beta_0=2$ ve eğim parametresi $\beta_1=1$ olmaktadır. e_i hata terimleri ise yukarıda verilen on iki olasılık dağılımdan alınmıştır.

Senaryo 1: $y_i = 2 + x_i + e_i, e_i \sim S_{0,7}(0,1,0), x_i=1,2,\dots,n$

Senaryo 2: $y_i = 2 + x_i + e_i, e_i \sim S_1(0,1,0), x_i=1,2,\dots,n$

Senaryo 3: $y_i = 2 + x_i + e_i, e_i \sim S_{1,3}(0,1,0), x_i=1,2,\dots,n$

Senaryo 4: $y_i = 2 + x_i + e_i, e_i \sim S_{1,7}(0,1,0), x_i=1,2,\dots,n$

Senaryo 5: $y_i = 2 + x_i + e_i, e_i \sim S_{0,7}(1,1,0), x_i=1,2,\dots,n$

Senaryo 6: $y_i = 2 + x_i + e_i, e_i \sim S_{0,7}(-1,1,0), x_i=1,2,\dots,n$

Senaryo 7: $y_i = 2 + x_i + e_i, e_i \sim S_1(1,1,0), x_i=1,2,\dots,n$

Senaryo 8: $y_i = 2 + x_i + e_i, e_i \sim S_1(-1,1,0), x_i=1,2,\dots,n$

Senaryo 9: $y_i = 2 + x_i + e_i, e_i \sim S_{1,3}(1,1,0), x_i=1,2,\dots,n$

Senaryo 10: $y_i = 2 + x_i + e_i, e_i \sim S_{1,3}(-1,1,0), x_i=1,2,\dots,n$

Senaryo 11: $y_i = 2 + x_i + e_i, e_i \sim S_{1,7}(1,1,0), x_i=1,2,\dots,n$

Senaryo 12: $y_i = 2 + x_i + e_i, e_i \sim S_{1,7}(-1,1,0), x_i=1,2,\dots,n$

Kararlı dağılımları türetirken Nolan (2009)'ın kararlı rasgele değişken simülasyonu için önerdiği eşitlik referans alınmıştır.

Buna göre,

Simetrik kararlı değişkenler

$$Z = \begin{cases} \frac{\sin \alpha \Theta}{(\cos \Theta)^{1/\alpha}} \left[\frac{\cos ((\alpha-1)\Theta)}{W} \right]^{(1-\alpha)/\alpha}, & \alpha \neq 1 \\ \tan \Theta, & \alpha \geq 1 \end{cases} \quad (7.1)$$

$S(\alpha,0;0)=S(\alpha,0;1)$ dağılımına sahiptir.

Simetrik olmayan kararlı değişkenler

$$Z = \begin{cases} \frac{\sin \alpha (\theta_0 + \Theta)}{(\cos \alpha \theta_0 \cos \Theta)^{1/\alpha}} \left[\frac{\cos (\alpha \theta_0 + (\alpha-1)\Theta)}{W} \right]^{(1-\alpha)/\alpha}, & \alpha \neq 1 \\ \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta \Theta \right) \tan \Theta - \beta \log \left(\frac{\frac{\pi}{2} W \cos \Theta}{\frac{\pi}{2} + \beta \Theta} \right) \right], & \alpha \geq 1 \end{cases} \quad (7.2)$$

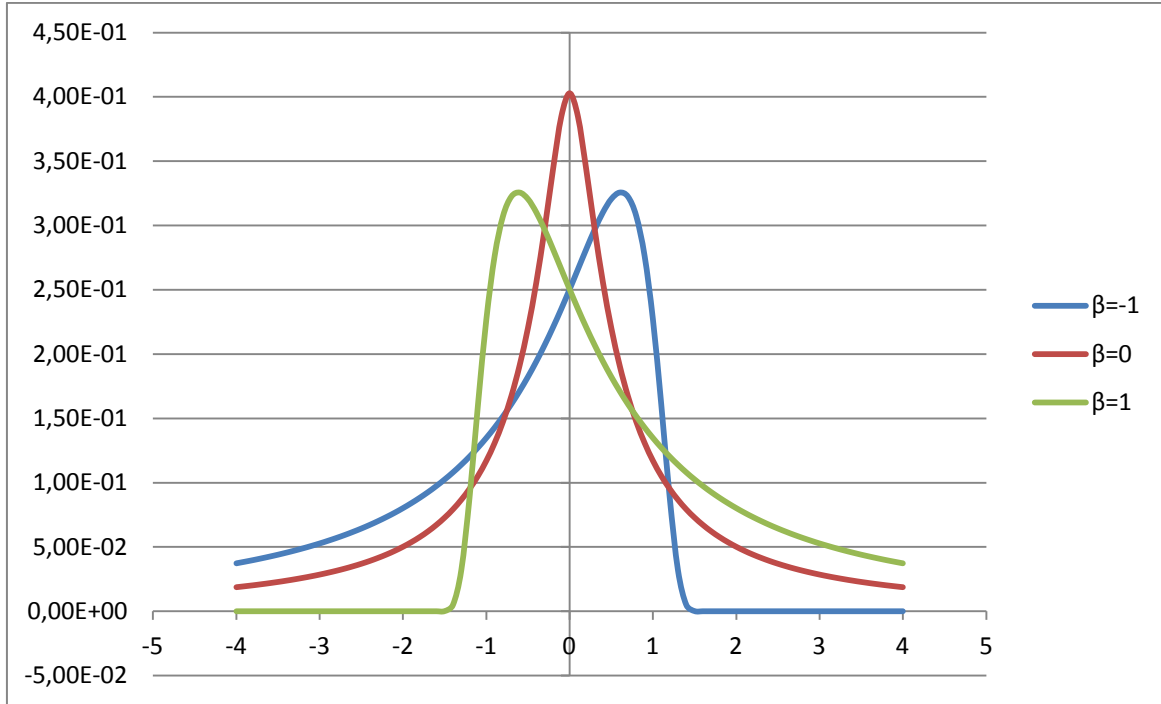
$S(\alpha,\beta;1)$ dağılımına sahiptir.

U_1 ve U_2 Uniform(1,0) dağılan rasgele değişkenler olmak üzere $\theta = \pi \left(U_1 - \frac{1}{2} \right)$, $W = -\log U_2$ 'dir.

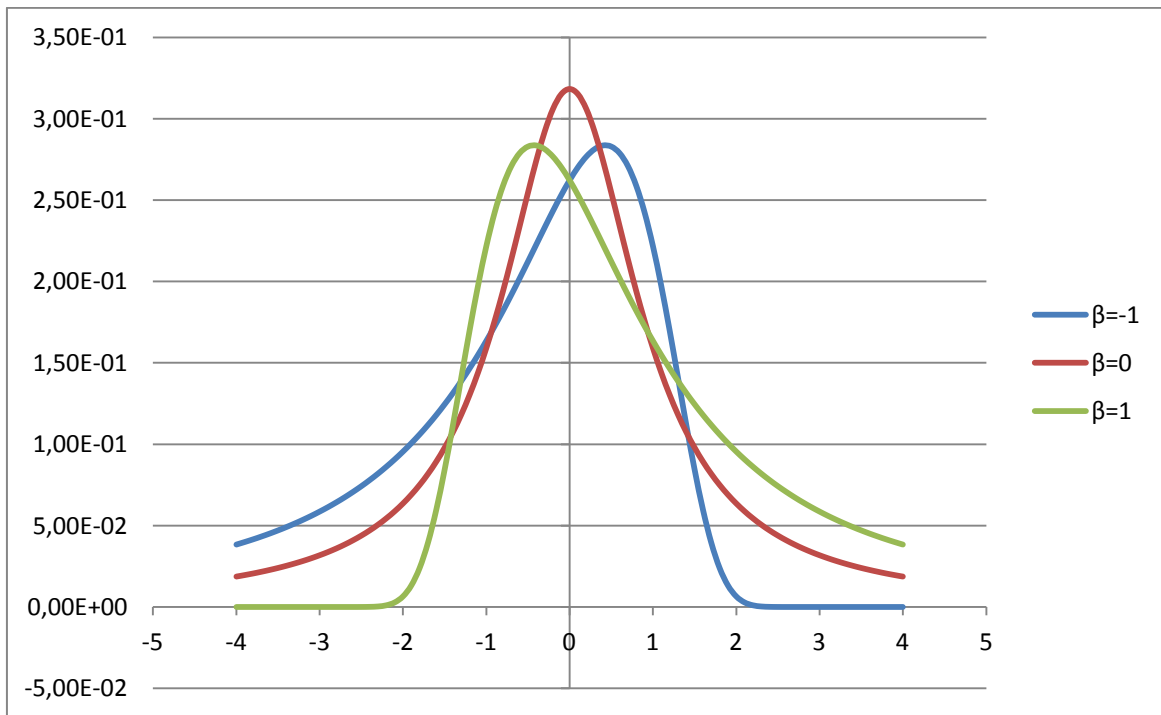
Veri üretme aşamasının ardından her bir veri setine, EKK, M (Andrews, Fair, Hampel, Huber, Tukey), MM, S, LTS, LMS, LAD, Bootstrap ve Kantil tahmin yöntemleri uygulanarak parametre tahminleri elde edilmiştir. Bu işlemler 100 kez tekrarlanmıştır.

Karşılaştırma kriteri olarak tahmin edicilerin ortalaması, varyansı ve hata kareler ortalaması alınmıştır.

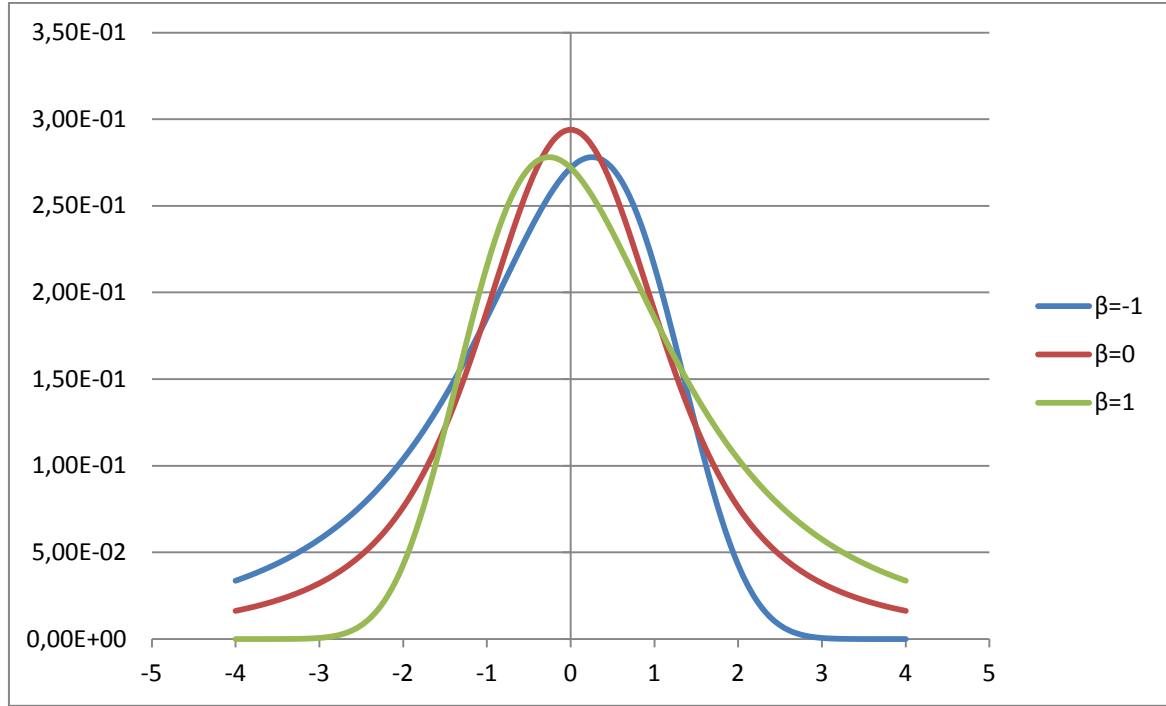
Çalışma içerisinde hata terimlerinin sahip olduğu farklı alfa kararlı dağılımlar için yoğunluklar aşağıdaki grafikler ile gösterilmiştir.



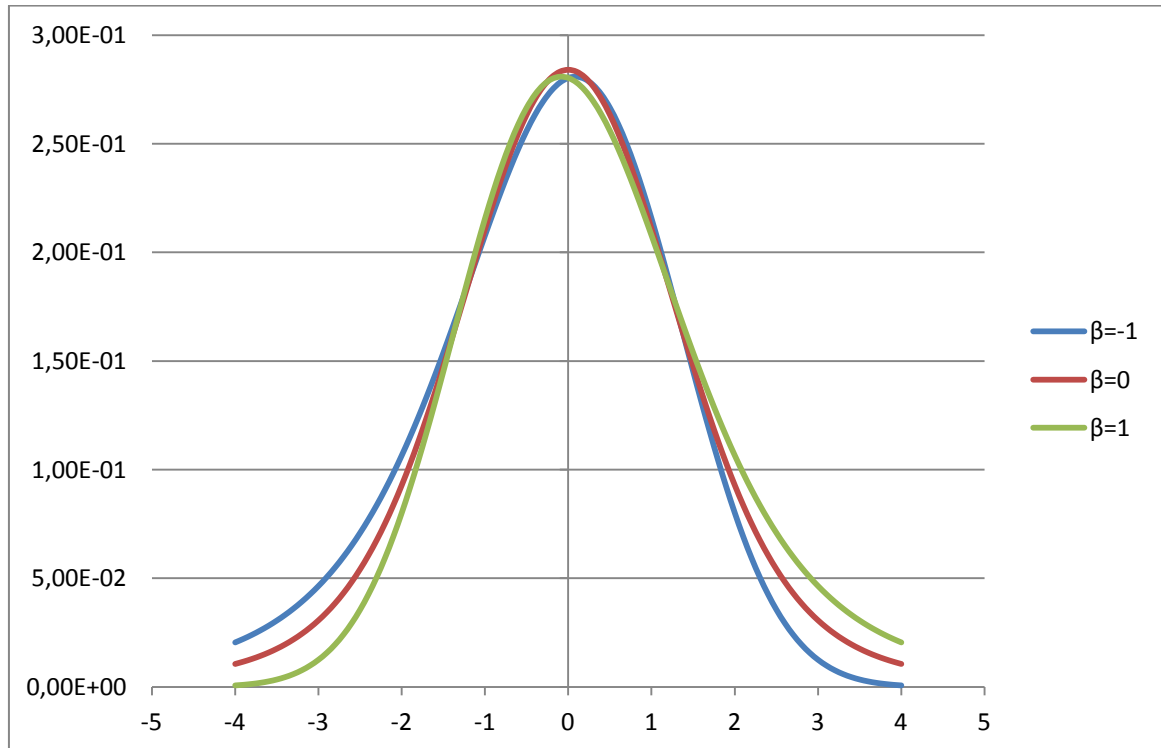
Şekil 7.1. $S_{0,7}(-1;1;0)$, $S_{0,7}(0;1;0)$, $S_{0,7}(1;1;0)$ için kararlı yoğunluklar



Şekil 7.2. $S_{1,7}(-1;1;0)$, $S_{1,7}(0;1;0)$, $S_{1,7}(1;1;0)$ için kararlı yoğunluklar



Şekil 7.3. $S_{1,3}(-1;1;0)$, $S_{1,3}(0;1;0)$, $S_{1,3}(1;1;0)$ için kararlı yoğunluklar



Şekil 7.4. $S_{1,7}(-1;1;0)$, $S_{1,7}(0;1;0)$, $S_{1,7}(1;1;0)$ kararlı yoğunluklar

Simülasyon çalışmasından elde edilen sonuçlar tablolar halinde sunulmuştur.

Çizelge 7.1. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,750722	1,017629	2,533221	0,028573	2,595361	0,028884
Tukey	1,749566	1,017973	2,537617	0,028584	2,600334	0,028907
Fair	1,421432	1,053586	7,843073	0,086748	8,177814	0,089619
Huber	1,529965	1,040963	4,523430	0,047938	4,744363	0,049615
Hampel	1,389337	1,078515	7,239369	0,113009	7,612279	0,119174
LMS	1,978641	1,004025	2,065175	0,029924	2,065632	0,029941
MM	1,865322	1,010522	1,581108	0,020115	1,599247	0,020226
S	1,978298	0,998900	1,338918	0,016446	1,339389	0,016447
LTS	0,487424	1,028327	154,2713	4,256239	156,5592	4,257041
LAD	1,712159	1,023301	2,339917	0,022763	2,422769	0,023306
Kantil	23,11643	17,20541	52518,01	10742,58	52963,91	11005,19
Bootstrap	-31,09843	5,141367	50884,16	729,1513	51979,67	746,3022

Çizelge 7.2. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,773355	1,015031	1,415447	0,012121	1,466815	0,012347
Tukey	1,772563	1,015270	1,409359	0,012046	1,461087	0,012280
Fair	1,407013	1,046278	10,77198	0,068233	11,12362	0,070375
Huber	1,402692	1,045174	5,051185	0,029750	5,407961	0,031791
Hampel	2,261468	0,981233	68,19414	0,430435	68,26250	0,430787
LMS	1,994186	1,002149	1,043040	0,009418	1,043074	0,009422
MM	1,845855	1,007831	1,050299	0,007595	1,074059	0,007657
S	1,946378	1,002822	0,666432	0,005394	0,669307	0,005402
LTS	1,172726	1,296669	677,5809	14,06962	678,2653	14,15764
LAD	1,842088	1,015165	0,777101	0,009841	0,802037	0,010071
Kantil	34,82525	0,409469	25428,66	77,44894	26506,16	77,79767
Bootstrap	1279,889	-267,921	157947689,6	7143003,6	159580689,2	7215322,25

Çizelge 7.3. Senaryo 1 ($S_{0,7}(0,1,0)$) $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,070967	0,995398	0,838987	0,002736	0,844023	0,002757
Tukey	2,079076	0,994907	0,848304	0,002770	0,854557	0,002795
Fair	2,179102	0,991015	2,266012	0,006962	2,298090	0,007042
Huber	2,147568	0,99159	1,496981	0,004949	1,518757	0,005019
Hampel	2,035761	0,995704	1,860716	0,006641	1,861995	0,006659
LMS	2,022507	1,001215	0,805376	0,002527	0,805883	0,002528
MM	2,132606	0,992208	0,743095	0,002443	0,760680	0,002503
S	2,047461	0,999228	0,499397	0,001599	0,501650	0,001600
LTS	0,922786	0,981653	217,9911	0,106568	219,1515	0,106904
LAD	2,061862	0,997928	0,736503	0,002247	0,740330	0,002250
Kantil	14,90313	0,962621	1635,589	3,182050	1802,080	3,183447
Bootstrap	4424,195	-509,816	1,86E+09	25623486	1,88E+09	2,59E+07

Çizelge 7.4. Senaryo 1 ($S_{0.7}(0,1,0)$), $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,982689	1,001631	0,340960	0,000654	0,341260	0,028884
Tukey	1,980467	1,001693	0,338656	0,000648	0,339038	0,028907
Fair	1,899023	1,002858	0,895979	0,001673	0,906175	0,089619
Huber	1,911823	1,002932	0,621100	0,001222	0,628875	0,049615
Hampel	1,888466	1,004923	1,055803	0,002201	1,068243	0,119174
LMS	1,989026	0,999898	0,391121	0,000727	0,391241	0,029941
MM	1,974483	1,00141	0,343730	0,000613	0,344381	0,020226
S	2,037753	0,999313	0,165521	0,000302	0,166946	0,016447
LTS	1,783645	1,002028	10,27282	0,022056	10,31963	4,257041
LAD	2,003216	0,999211	0,232806	0,000482	0,232816	0,023306
Kantil	12,3637	0,961784	316,3271	0,498669	423,7334	11005,19
Bootstrap	35,39597	-0,04836	124616,3	152,7454	125731,5	746,3022

Çizelge 7.5. Senaryo 2 ($S_1(0,1,0)$), $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,993884	1,001175	1,270894	0,018117	1,270932	0,018118
Tukey	2,000216	1,000599	1,255006	0,017975	1,255006	0,017976
Fair	2,002978	0,993007	2,147380	0,027351	2,147389	0,027401
Huber	2,025293	0,993975	1,514183	0,020507	1,514823	0,020544
Hampel	1,939576	1,001816	2,931430	0,037053	2,935081	0,037057
LMS	1,984748	0,995586	1,936967	0,029343	1,937200	0,029363
MM	1,998332	1,000293	1,242913	0,015547	1,242916	0,015548
S	1,951627	1,001139	1,924416	0,025233	1,926756	0,025235
LTS	1,855873	0,992478	8,834806	0,056353	8,855578	0,056409
LAD	1,997565	0,994211	1,187945	0,015654	1,187951	0,015688
Kantil	5,820378	10,10735	12794,95	9443,238	12809,54	9526,182
Bootstrap	-2,58519	2,437291	2826,585	240,3454	2847,609	242,4112

Çizelge 7.6. Senaryo 2 ($S_1(0,1,0)$), $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,980046	1,006519	1,063635	0,006735	1,064033	0,006777
Tukey	1,977241	1,006691	1,065524	0,006769	1,066042	0,006813
Fair	1,848216	1,019464	2,062281	0,012917	2,085319	0,013296
Huber	1,902223	1,014861	1,620467	0,009915	1,630028	0,010136
Hampel	1,896557	1,018652	2,211772	0,013921	2,222472	0,014269
LMS	2,081649	0,991586	1,368752	0,009501	1,375419	0,009571
MM	1,978985	1,002419	0,724818	0,004534	0,725259	0,004539
S	2,043639	0,993842	0,960014	0,005715	0,961919	0,005753
LTS	1,977051	1,016129	7,332983	0,045321	7,333509	0,045581
LAD	1,944603	1,005045	1,190114	0,007333	1,193183	0,007358
Kantil	7,763595	0,981163	298,9787	1,009609	332,1977	1,009964
Bootstrap	0,862254	1,014731	211,0953	2,554966	212,3898	2,555183

Çizelge 7.7. Senaryo 2 ($S_1(0,1,0)$), $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,088496	0,995219	0,447940	0,001928	0,455771	0,001950
Tukey	2,087125	0,995271	0,450017	0,001935	0,457608	0,001957
Fair	2,007829	1,001209	0,679405	0,002782	0,679466	0,002783
Huber	2,007892	1,000828	0,522379	0,002234	0,522442	0,002234
Hampel	2,016708	0,999509	0,742758	0,003141	0,743037	0,003141
LMS	2,192462	0,987983	0,907659	0,003272	0,944701	0,003416
MM	2,095258	0,994953	0,390574	0,001648	0,399648	0,001673
S	2,022852	0,998259	0,719617	0,002542	0,720139	0,002545
LTS	2,159735	0,991672	1,239904	0,005616	1,265420	0,005685
LAD	2,014176	0,999831	0,436362	0,001866	0,436562	0,001865
Kantil	7,524184	0,981135	101,6814	0,246528	132,1980	0,246883
Bootstrap	-2,52625	1,290434	1142,577	3,051578	1163,064	3,135929

Çizelge 7.8. Senaryo 2 ($S_1(0,1,0)$), $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,919872	1,002375	0,349625	0,000683	0,356046	0,000689
Tukey	1,919663	1,002378	0,348427	0,000679	0,354881	0,000684
Fair	1,868901	1,002703	0,454959	0,000757	0,472146	0,000764
Huber	1,864150	1,003350	0,408003	0,000678	0,426459	0,000689
Hampel	1,823875	1,004872	0,506521	0,000815	0,537541	0,000839
LMS	2,111463	0,994299	0,440736	0,000875	0,453159	0,000907
MM	1,926858	1,002237	0,324923	0,000625	0,330273	0,000629
S	2,028189	0,997459	0,302783	0,000624	0,303578	0,000630
LTS	1,718121	1,008745	1,113885	0,001598	1,193341	0,001674
LAD	1,935680	1,000974	0,318812	0,000608	0,322949	0,000609
Kantil	5,087743	1,012295	14,69123	0,020555	24,22539	0,020707
Bootstrap	1,692360	0,979403	81,54288	0,069604	81,63752	0,070028

Çizelge 7.9. Senaryo 3 ($S_{1,3}(0,1,0)$), $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,001579	0,995528	0,996687	0,012919	0,996689	0,012939
Tukey	2,002029	0,995621	0,997518	0,012898	0,997522	0,012917
Fair	1,974551	0,996731	1,033602	0,013256	1,034250	0,013267
Huber	1,979687	0,997508	0,976263	0,012897	0,976676	0,012904
Hampel	1,951164	1,000487	1,181029	0,015855	1,183414	0,015855
LMS	2,278433	0,965298	2,740761	0,031584	2,818286	0,032788
MM	2,019799	0,993438	0,881222	0,011101	0,881614	0,011144
S	1,942709	0,999323	1,986826	0,018186	1,990108	0,018186
LTS	1,833399	1,006093	1,401604	0,017211	1,429360	0,017248
LAD	1,935914	0,998816	1,164691	0,013607	1,168798	0,013608
Kantil	6,146730	0,918947	106,9836	0,651144	124,1789	0,657714
Bootstrap	4,446218	0,829517	421,0177	1,103323	427,0017	1,132387

Çizelge 7.10. Senaryo 3 ($S_{1,3}(0,1,0)$), $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,908191	1,007201	0,833802	0,005786	0,842231	0,005838
Tukey	1,906003	1,007417	0,822414	0,005731	0,831249	0,005786
Fair	1,879672	1,010505	0,999490	0,007136	1,013969	0,007247
Huber	1,881368	1,010717	0,913539	0,006605	0,927613	0,006719
Hampel	1,854072	1,011929	0,989527	0,007354	1,010822	0,007496
LMS	1,867594	1,014189	1,569989	0,010445	1,587521	0,010647
MM	1,921583	1,007467	0,816233	0,005402	0,822382	0,005458
S	1,906186	1,010579	1,307343	0,009565	1,316144	0,009677
LTS	1,885456	1,009101	1,573879	0,009821	1,586999	0,009904
LAD	1,904506	1,009838	0,998310	0,007062	1,007429	0,007159
Kantil	4,797289	1,051713	21,06853	0,179479	28,89336	0,182153
Bootstrap	1,911540	1,034516	9,383101	0,111988	9,390926	0,113179

Çizelge 7.11. Senaryo 3 ($S_{1,3}(0,1,0)$), $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,024186	0,997575	0,462215	0,001632	0,462800	0,001638
Tukey	2,024844	0,997534	0,461424	0,001629	0,462041	0,001636
Fair	2,054668	0,994451	0,632094	0,002209	0,635082	0,002239
Huber	2,047030	0,995259	0,545462	0,001934	0,547674	0,001956
Hampel	2,044161	0,995029	0,670247	0,002305	0,672197	0,002329
LMS	1,948531	1,005970	0,911101	0,002876	0,913750	0,002912
MM	2,016393	0,997925	0,427966	0,001509	0,428234	0,001514
S	1,940022	1,004057	0,821908	0,002665	0,825506	0,002681
LTS	1,974381	0,999372	0,678363	0,002484	0,679019	0,002485
LAD	2,059062	0,995092	0,504841	0,001856	0,508329	0,001879
Kantil	4,937727	0,984879	8,485776	0,022264	17,11602	0,022493
Bootstrap	2,275084	0,976234	8,028992	0,018656	8,104663	0,019220

Çizelge 7.12. Senaryo 3 ($S_{1,3}(0,1,0)$), $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,971177	1,001596	0,335899	0,000629	0,336729	0,000632
Tukey	1,971279	1,001562	0,336415	0,000629	0,337240	0,000632
Fair	1,985563	1,000732	0,364646	0,000674	0,364855	0,000675
Huber	1,983250	1,000929	0,355742	0,000657	0,356022	0,000657
Hampel	1,977215	1,001154	0,400865	0,000783	0,401385	0,000784
LMS	2,008069	0,999154	0,648252	0,001182	0,648318	0,001183
MM	1,977767	1,001445	0,318414	0,000603	0,318908	0,000605
S	2,068314	0,997535	0,567252	0,000911	0,571918	0,000917
LTS	1,978161	1,000612	0,417806	0,000917	0,418283	0,000918
LAD	1,998982	1,000749	0,310712	0,000575	0,310713	0,000576
Kantil	4,538165	1,005292	3,424831	0,007691	9,867114	0,007719
Bootstrap	2,328117	0,969125	9,728351	0,058504	9,836012	0,059457

Çizelge 7.13. Senaryo 4 ($S_{1,7}(0,1,0)$), $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,055449	1,001854	0,585314	0,007837	0,588388	0,007841
Tukey	2,055755	1,001841	0,584772	0,007829	0,587880	0,007832
Fair	2,071013	0,999259	0,598327	0,008671	0,603369	0,008672
Huber	2,064989	0,999825	0,591909	0,008173	0,596132	0,008173
Hampel	2,053706	1,000627	0,619803	0,008944	0,622687	0,008945
LMS	2,072157	1,005476	1,951719	0,024313	1,956925	0,024343
MM	2,025154	1,005389	0,578655	0,008060	0,579287	0,008089
S	2,066382	1,007369	1,339411	0,017046	1,343818	0,017100
LTS	2,068481	1,000485	0,642141	0,008896	0,646831	0,008896
LAD	2,020216	1,007502	0,754335	0,009668	0,754743	0,009724
Kantil	4,446297	0,975959	11,74801	0,085563	17,73238	0,086141
Bootstrap	2,202974	0,977855	1,519803	0,035641	1,561001	0,036132

Çizelge 7.14. Senaryo 4 ($S_{1,7}(0,1,0)$), $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,007778	1,003133	0,360664	0,002585	0,360725	0,002595
Tukey	2,006003	1,003276	0,360629	0,002586	0,360665	0,002597
Fair	1,965814	1,009295	0,379623	0,002918	0,380791	0,003004
Huber	1,974883	1,007686	0,354469	0,002711	0,355100	0,002769
Hampel	1,971178	1,009194	0,409364	0,003081	0,410194	0,003166
LMS	1,925198	1,010486	1,546377	0,010979	1,551972	0,011089
MM	1,985344	1,004854	0,356882	0,002610	0,357097	0,002633
S	2,036656	1,004580	1,144236	0,008012	1,145579	0,008033
LTS	1,973493	1,009700	0,499511	0,004306	0,500214	0,004400
LAD	1,997943	1,003520	0,472303	0,003209	0,472307	0,003221
Kantil	3,881021	1,033450	2,683990	0,027682	6,222231	0,028801
Bootstrap	1,916306	1,018897	0,612159	0,005894	0,619164	0,006251

Çizelge 7.15. Senaryo 4 ($S_{1,7}(0,1,0)$), $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,944568	1,003547	0,304446	0,001039	0,307518	0,001051
Tukey	1,943820	1,003563	0,301052	0,001033	0,304208	0,001045
Fair	1,954676	1,003072	0,305249	0,001060	0,307303	0,001069
Huber	1,947188	1,003512	0,295943	0,001056	0,298732	0,001068
Hampel	1,936076	1,004206	0,308262	0,001081	0,312348	0,001098
LMS	2,087205	0,996834	1,183488	0,003317	1,191093	0,003327
MM	1,947759	1,003318	0,303759	0,001056	0,306488	0,001067
S	2,038297	0,997202	0,867146	0,002651	0,868613	0,002658
LTS	1,945690	1,003088	0,356413	0,001208	0,359362	0,001218
LAD	1,940844	1,003541	0,460796	0,001417	0,464295	0,001429
Kantil	3,937401	1,003704	1,457799	0,004804	5,211322	0,004818
Bootstrap	1,949023	1,003019	0,705770	0,002005	0,708369	0,002014

Çizelge 7.16. Senaryo 4 ($S_{1,7}(0,1,0)$), $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,017925	0,998870	0,239399	0,000407	0,239721	0,000408
Tukey	2,018589	0,998854	0,238691	0,000405	0,239037	0,000406
Fair	2,033749	0,998383	0,240959	0,000421	0,242098	0,000424
Huber	2,022988	0,998580	0,238987	0,000404	0,239516	0,000406
Hampel	2,030609	0,998266	0,242954	0,000422	0,243889	0,000425
LMS	2,096653	0,997239	0,963248	0,001754	0,972589	0,001762
MM	2,023454	0,998854	0,236409	0,000406	0,236959	0,000408
S	2,048691	0,999228	0,657612	0,001129	0,659983	0,001129
LTS	2,042016	0,998171	0,285119	0,000477	0,286884	0,000480
LAD	2,075729	0,997695	0,323297	0,000566	0,329032	0,000572
Kantil	3,982132	1,000486	1,226794	0,003419	5,155642	0,003419
Bootstrap	2,030092	0,995966	0,586496	0,002108	0,587402	0,002124

Çizelge 7.17. Senaryo 5 $S_{0,7}(1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	5,233137	1,017379	4,420617	0,043544	14,873793	0,043846
Tukey	5,236325	1,017425	4,420479	0,043512	14,894280	0,043815
Fair	6,615055	1,132281	23,22526	0,548863	44,523994	0,566361
Huber	6,233723	1,064409	11,77434	0,167601	29,698747	0,171750
Hampel	6,192424	1,151448	14,48957	0,297033	32,065985	0,319969
LMS	4,280169	1,024192	3,067213	0,049221	8,2663834	0,049807
MM	4,687314	1,024333	1,788323	0,027701	9,0099816	0,028294
S	4,238813	1,007186	1,895878	0,025626	6,9081595	0,025678
LTS	10,12906	0,983301	415,9169	4,948637	481,99852	4,948916
LAD	5,158500	1,033382	5,139489	0,068045	15,115613	0,069159
Kantil	476,9106	-20,3605	10963850	51371,95	11189389,91	51828,22
Bootstrap	92,60099	-3,78362	240546,3	1553,108	248754,8472	1575,991

Çizelge 7.18. Senaryo 5 $S_{0,7}(1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	5,255341	0,989611	3,544583	0,026482	14,141831	0,026590
Tukey	5,268125	0,989104	3,596173	0,026464	14,276815	0,026582
Fair	6,879172	1,026085	14,66170	0,160887	38,468015	0,161567
Huber	6,569032	1,009166	8,474026	0,073298	29,350083	0,073382
Hampel	7,208805	1,015241	21,89695	0,150674	49,028609	0,150906
LMS	4,361062	0,984307	1,980217	0,012481	7,5548324	0,012727
MM	4,932386	0,988925	2,250714	0,011949	10,849604	0,012072
S	4,318710	0,989545	1,586737	0,008937	6,9631544	0,009045
LTS	8,274806	1,117596	58,84506	0,603841	98,218254	0,617670
LAD	5,235504	1,005589	3,913552	0,043786	14,382038	0,043816
Kantil	40,40695	2,939189	16118,69	155,5719	17593,785	159,3323
Bootstrap	25,63178	5,048096	142705,04	1215,725	143263,50	1232,1123

Çizelge 7.19. Senaryo 5 $S_{0,7}(1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	5,194739	0,994524	1,679668	0,003329	11,886025	0,003358
Tukey	5,202683	0,99436	1,671799	0,003302	11,928979	0,003334
Fair	6,674947	1,015876	7,148683	0,031188	29,003809	0,031440
Huber	6,340117	1,007924	3,817644	0,011818	22,654256	0,011880
Hampel	7,084349	1,013799	6,638124	0,014791	32,488732	0,014982
LMS	4,124115	1,002034	0,972925	0,002594	5,4847889	0,002598
MM	4,922397	0,996163	1,365814	0,002499	9,9062162	0,002514
S	4,158667	1,000231	0,768576	0,001613	5,4284202	0,001613
LTS	7,378816	1,025934	26,14328	0,096148	55,074943	0,096821
LAD	5,047505	1,009077	2,496097	0,008306	11,783381	0,008388
Kantil	39,20853	1,212209	6361,987	15,61621	7746,4621	15,661245
Bootstrap	38,43474	2,139859	39194,09	155,4740	40521,583	156,77328

Çizelge 7.20. Senaryo 5 $S_{0,7}(1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	5,106051	0,995093	1,57659	0,001382	11,224148	0,001406
Tukey	5,115101	0,995021	1,596581	0,001381	11,300436	0,001406
Fair	6,674767	0,997136	4,952606	0,007307	26,806054	0,007316
Huber	6,316753	0,996859	3,376369	0,003641	22,010732	0,003651
Hampel	7,111221	0,993599	5,545737	0,005951	31,670322	0,005992
LMS	4,108737	0,995264	0,620309	0,000799	5,067084	0,000822
MM	4,756132	0,999598	0,605229	0,000898	8,201494	0,000898
S	4,141122	0,996755	0,386273	0,000506	4,97068	0,000516
LTS	8,563995	1,009229	21,939964	0,055431	65,026	0,055516
LAD	5,153961	0,993079	1,624356	0,002669	11,571832	0,002717
Kantil	35,84887	1,194493	1970,6973	12,013368	3116,4435	12,051195
Bootstrap	-88148,5	8007,317	7,81E+11	6,37E+09	7,89E+11	6,43E+09

Çizelge 7.21. Senaryo 6 $S_{0,7}(-1,1,0)$ $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	-1,075565	0,973494	5,377788	0,055485	14,836888	0,056188
Tukey	-1,080120	0,973702	5,383619	0,055556	14,870756	0,056248
Fair	-2,898943	0,968703	28,40644	0,149296	52,406084	0,150276
Huber	-2,270282	0,967538	9,564753	0,082623	27,800061	0,083677
Hampel	-2,814425	0,951242	14,20544	0,141172	37,384126	0,143549
LMS	0,132655	0,964777	2,236679	0,051683	5,7236581	0,052923
MM	-0,612945	0,973002	2,211967	0,033778	9,0394481	0,034507
S	0,066071	0,967834	1,371749	0,025384	5,1118295	0,026418
LTS	-57,30248	6,023606	285961,1	2495,723	289477,90	2520,960
LAD	-1,025784	0,953565	7,828740	0,086836	16,984106	0,088993
Kantil	0,804576	0,990654	0,434712	0,006092	1,8637513	0,006179
Bootstrap	303,52066	-70,9538	8487434,1	411887,4	8578348,8	417064,7

Çizelge 7.22. Senaryo 6 $S_{0,7}(-1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Beklenen Değer		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	-1,252230	0,995129	4,177111	0,021714	14,754112	0,021737
Tukey	-1,255501	0,995292	4,167678	0,021656	14,765971	0,021678
Fair	-2,513742	0,935997	12,20300	0,134499	32,576869	0,138595
Huber	-2,386120	0,977813	6,350972	0,045862	25,589023	0,046354
Hampel	-2,693213	0,975376	7,098564	0,060945	29,124814	0,061551
LMS	-0,179331	1,004588	1,447747	0,007978	6,1972319	0,007999
MM	-0,845594	0,998313	2,123043	0,010481	10,220451	0,010484
S	-0,184820	1,000548	1,136331	0,006625	5,9097724	0,006625
LTS	-5,446750	0,968517	122,7580	0,958896	178,21217	0,959887
LAD	-1,253172	0,987783	5,175106	0,035966	15,758237	0,036115
Kantil	0,7642879	0,998077	0,372455	0,002589	1,8994395	0,002592
Bootstrap	-330,7105	18,49278	8756293,07	35503,81	8866989,3	35809,80

Çizelge 7.23. Senaryo 6 $S_{0,7}(-1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	-0,893685	0,990146	1,464967	0,003669	9,838383	0,003767
Tukey	-0,904020	0,990348	1,473777	0,003662	9,907109	0,003756
Fair	-3,494609	1,028680	35,01241	0,053989	65,20315	0,054812
Huber	-2,630841	1,009598	8,984452	0,011634	30,42914	0,011727
Hampel	-2,788786	0,985197	7,884032	0,015238	30,81651	0,015458
LMS	-0,099369	1,001038	0,565459	0,001909	4,972809	0,001911
MM	-0,590859	0,990576	0,783723	0,002503	7,496274	0,002592
S	-0,078512	0,999709	0,460332	0,001557	4,780543	0,001558
LTS	-5,163412	1,049725	99,18890	0,192335	150,5034	0,194808
LAD	-0,937260	0,993704	1,919193	0,006629	10,54669	0,006669
Kantil	0,838372	0,996418	0,132291	0,000441	1,481672	0,000454
Bootstrap	-12,267294	-1,347617	36118,35	484,8261	36321,91	490,3375

Çizelge 7.24. Senaryo 6 $S_{0,7}(-1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	-0,890164	0,999575	0,758302	0,000997	9,111351	0,000997
Tukey	-0,893820	0,999495	0,762807	0,001002	9,137000	0,001003
Fair	-2,381352	0,997047	2,745607	0,004766	21,94185	0,004775
Huber	-2,041550	0,997405	1,543220	0,002227	17,87734	0,002234
Hampel	-2,771155	0,998054	3,136742	0,004587	25,90066	0,004591
LMS	0,023015	0,998324	0,429988	0,000847	4,338460	0,00085
MM	-0,692923	1,000629	0,621047	0,000773	7,872878	0,000774
S	-0,061884	1,001546	0,439533	0,000678	4,690896	0,000681
LTS	-4,235850	1,014731	21,31636	0,026227	60,20219	0,026444
LAD	-0,870138	0,997128	1,045809	0,002004	9,283502	0,002013
Kantil	0,786185	1,000046	0,106627	0,000230	1,579973	0,00023
Bootstrap	-13,824863	0,03671	10975,37	36,51023	11225,80	37,43817

Çizelge 7.25. Senaryo 7 $S_1(1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,854734	0,996314	1,302538	0,015650	2,033108	0,015664
Tukey	2,856327	0,996635	1,301847	0,015621	2,035143	0,015632
Fair	3,545611	0,991440	3,888779	0,039749	6,277692	0,039823
Huber	3,410995	0,989702	2,493006	0,023465	4,483912	0,023571
Hampel	3,643902	0,989464	3,905451	0,033071	6,607864	0,033182
LMS	2,519939	0,967467	2,219032	0,029122	2,489369	0,030181
MM	2,651840	0,995543	1,215144	0,016076	1,640039	0,016096
S	2,298247	0,989161	1,654168	0,024433	1,743119	0,024551
LTS	3,537344	1,000332	5,486104	0,050464	7,849530	0,050464
LAD	2,767397	1,008972	1,581912	0,025402	2,170810	0,025483
Kantil	22,38791	1,181663	4654,762	50,03143	5070,429	50,06443
Bootstrap	6,116058	1,053848	266,1512	4,389240	283,0931	4,392139

Çizelge 7.26. Senaryo 7 $S_1(1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,906212	0,999208	1,263402	0,008256	2,084623	0,008257
Tukey	2,904782	1,000590	1,282831	0,008777	2,101462	0,008777
Fair	3,544260	0,992027	3,121240	0,021784	5,505980	0,021848
Huber	3,345894	0,997201	2,038062	0,013846	3,849493	0,013853
Hampel	3,583776	0,995361	2,348962	0,017408	4,857309	0,017429
LMS	2,169948	1,003609	2,236909	0,013353	2,265791	0,013366
MM	2,620674	1,003986	1,095718	0,006935	1,480955	0,006950
S	2,172807	1,004456	1,842484	0,010816	1,872347	0,010835
LTS	4,044668	0,971999	10,06674	0,058318	14,24741	0,059102
LAD	2,885587	1,001140	2,075757	0,017453	2,860022	0,017454
Kantil	14,11632	0,952399	364,0315	2,209519	510,8368	2,211785
Bootstrap	6,751954	0,913431	142,8467	0,810221	165,4277	0,817715

Çizelge 7.27. Senaryo 7 $S_1(1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,740874	0,996679	0,662812	0,001614	1,211706	0,001625
Tukey	2,746092	0,996739	0,664744	0,001618	1,221397	0,001629
Fair	3,333102	0,996558	1,378378	0,003593	3,155539	0,003605
Huber	3,202082	0,996439	1,044513	0,002455	2,489513	0,002467
Hampel	3,488697	0,997030	1,559537	0,003674	3,775756	0,003682
LMS	2,184176	0,993416	1,032233	0,003002	1,066154	0,003046
MM	2,634769	0,996571	0,620143	0,001469	1,023075	0,001480
S	2,176193	0,993427	0,673652	0,001885	0,704696	0,001928
LTS	3,422723	0,991329	1,973533	0,005166	3,997672	0,005242
LAD	2,737994	0,994322	1,018962	0,002697	1,563596	0,002729
Kantil	12,31496	1,010849	203,6093	0,518058	310,0077	0,518175
Bootstrap	9,673986	0,878592	753,0373	0,520193	811,9274	0,534933

Çizelge 7.28. Senaryo 7 $S_1(1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,761402	0,997705	0,483259	0,000838	1,062991	0,000843
Tukey	2,764739	0,997721	0,483035	0,000836	1,067862	0,000841
Fair	3,307637	1,000641	1,118857	0,002196	2,828772	0,002196
Huber	3,206980	0,999782	0,751103	0,001243	2,207905	0,001243
Hampel	3,518770	0,999519	1,266201	0,002057	3,572864	0,002057
LMS	1,948140	1,001479	0,528856	0,001062	0,531545	0,001064
MM	2,652136	0,997130	0,421005	0,000778	0,846286	0,000787
S	2,063268	0,999310	0,399593	0,000871	0,403596	0,000871
LTS	3,521544	0,999333	1,582605	0,003235	3,897701	0,003236
LAD	2,719440	0,997189	0,610389	0,001516	1,127983	0,001524
Kantil	10,03137	1,054530	63,6337	0,141138	128,1367	0,144111
Bootstrap	12,13964	0,799153	3008,766	3,555106	3111,578	3,595445

Çizelge 7.29. Senaryo 8 $S_1(-1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	0,810572	1,019564	2,423996	0,027274	3,838735	0,027656
Tukey	0,802186	1,020104	2,431995	0,027279	3,866754	0,027684
Fair	0,223028	1,029250	4,002280	0,047393	7,159911	0,048248
Huber	0,344196	1,028059	3,254514	0,035316	5,996202	0,036103
Hampel	0,130146	1,023850	4,124086	0,049495	7,620441	0,050064
LMS	1,481210	1,034765	3,428516	0,031164	3,697659	0,032373
MM	1,117192	1,023658	1,775089	0,015725	2,554438	0,016285
S	1,687823	1,014541	1,652787	0,016893	1,750242	0,017105
LTS	0,187988	1,032963	5,685087	0,050189	8,968474	0,051275
LAD	0,902498	1,030884	2,882820	0,031071	4,087331	0,032025
Kantil	2,797167	1,008766	0,649701	0,007509	1,285176	0,007587
Bootstrap	-0,480212	0,717955	295,0813	9,509965	301,2327	9,589515

Çizelge 7.30. Senaryo 8 $S_1(-1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Beklenen Değer		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,270584	0,993488	1,284378	0,007409	1,816426	0,007451
Tukey	1,269687	0,993296	1,277049	0,007379	1,810408	0,007424
Fair	0,620749	1,004835	2,212462	0,014979	4,114792	0,015003
Huber	0,770374	1,001109	1,738655	0,010481	3,250634	0,010482
Hampel	0,477672	1,006054	2,310280	0,013633	4,627763	0,013670
LMS	1,912275	0,993834	1,645428	0,010992	1,653124	0,011029
MM	1,557712	0,987690	0,811487	0,007037	1,007105	0,007188
S	1,903037	0,992360	1,218766	0,008338	1,228168	0,008397
LTS	0,507546	0,995634	4,082241	0,029067	6,309659	0,029086
LAD	1,343167	0,992118	1,413098	0,010791	1,844528	0,010853
Kantil	2,977558	0,990215	0,502216	0,003779	1,457836	0,003876
Bootstrap	-1,46969	0,942692	46,28067	0,721410	58,31942	0,724694

Çizelge 7.31. Senaryo 8 $S_1(-1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,377165	0,990885	0,651887	0,001907	1,039809	0,001989
Tukey	1,372044	0,990973	0,647893	0,001897	1,042222	0,001978
Fair	0,735157	0,994393	1,258059	0,003856	2,857888	0,003888
Huber	0,869005	0,993501	0,818132	0,002515	2,097283	0,002558
Hampel	0,579871	0,996373	1,191347	0,003640	3,208113	0,003653
LMS	1,935857	1,000154	0,995136	0,003415	0,999249	0,003415
MM	1,495835	0,991813	0,563980	0,001629	0,818163	0,001696
S	1,932344	0,997470	0,784545	0,002576	0,789123	0,002582
LTS	0,570987	1,003898	2,610913	0,005369	4,652993	0,005384
LAD	1,413977	0,993088	0,867044	0,003275	1,210467	0,003323
Kantil	2,989626	0,994127	0,236613	0,000866	1,215973	0,000901
Bootstrap	-2,85029	1,043192	165,5149	0,418307	189,0402	0,420173

Çizelge 7.32. Senaryo 8 $S_1(-1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,328345	1,001531	0,403033	0,000572	0,854153	0,000574
Tukey	1,324956	1,001524	0,404549	0,000574	0,860233	0,000576
Fair	0,758181	1,003121	0,916359	0,001439	2,458473	0,001449
Huber	0,890647	1,002361	0,660490	0,000932	1,891155	0,000937
Hampel	0,650319	1,001924	1,054423	0,001238	2,876063	0,001242
LMS	2,081487	0,996629	0,438898	0,000841	0,445538	0,000853
MM	1,421991	1,001459	0,350379	0,000539	0,684475	0,000541
S	1,988509	1,000369	0,275156	0,000519	0,275288	0,000519
LTS	0,589816	1,000297	1,568128	0,002505	3,556748	0,002505
LAD	1,347876	1,001774	0,600698	0,001113	1,025964	0,001116
Kantil	2,975365	0,998146	0,182702	0,000327	1,134039	0,000331
Bootstrap	-4,71458	1,086177	162,9120	0,204547	207,9977	0,211973

Çizelge 7.33. Senaryo 9 $S_{1,3}(1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	0,577335	1,007553	1,128874	0,011615	3,152851	0,011672
Tukey	0,577971	1,007692	1,127628	0,011576	3,149795	0,011635
Fair	0,681019	1,023253	1,562049	0,016958	3,301762	0,017499
Huber	0,691091	1,016594	1,338434	0,013113	3,051678	0,013388
Hampel	0,786503	1,022038	1,651333	0,017181	3,123907	0,017667
LMS	0,187076	0,996291	2,116572	0,029734	5,403266	0,029748
MM	0,363931	1,012526	0,802794	0,009864	3,479517	0,010021
S	0,141240	1,000855	1,509359	0,020294	4,964348	0,020294
LTS	0,668678	1,005753	1,199614	0,012989	2,972033	0,013023
LAD	0,442777	1,010888	1,318267	0,015213	3,743212	0,015332
Kantil	5,265980	1,317734	111,2486	2,724517	121,9152	2,825471
Bootstrap	0,809924	1,104681	12,77502	0,263247	14,19131	0,274205

Çizelge 7.34. Senaryo 9 $S_{1,3}(1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	0,550791	1,007216	0,831748	0,005535	2,931954	0,005587
Tukey	0,552554	1,007192	0,832684	0,005543	2,927784	0,005595
Fair	1,006843	0,992738	1,947585	0,011784	2,933946	0,011837
Huber	0,871334	0,998363	1,238834	0,008139	2,512722	0,008142
Hampel	0,980826	1,000284	1,181168	0,008073	2,219883	0,008073
LMS	0,241883	1,002979	1,940385	0,013976	5,031359	0,013985
MM	0,410644	1,008368	0,621193	0,004500	3,147245	0,004570
S	0,240169	0,997824	1,386043	0,010423	4,483049	0,010428
LTS	0,976674	0,996479	2,432622	0,013612	3,479818	0,013624
LAD	0,643067	0,995097	1,018801	0,007749	2,860068	0,007774
Kantil	7,757614	0,864750	86,57522	0,410225	119,7253	0,428518
Bootstrap	2,386585	0,959022	26,30460	0,132943	26,45405	0,134622

Çizelge 7.35. Senaryo 9 $S_{1,3}(1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	0,488445	1,000693	0,440156	0,001198	2,724955	0,001199
Tukey	0,491127	1,000728	0,439479	0,001197	2,716178	0,001198
Fair	0,764924	1,001479	0,730018	0,002213	2,255429	0,002215
Huber	0,684478	1,002069	0,570814	0,001629	2,301414	0,001633
Hampel	0,821433	1,002831	0,740544	0,002070	2,129564	0,002078
LMS	0,121846	0,997975	1,074781	0,003714	4,602242	0,003718
MM	0,419003	1,001071	0,392135	0,001184	2,891687	0,001186
S	0,103952	0,998942	0,680099	0,002380	4,275097	0,002381
LTS	0,590072	1,005609	0,797149	0,002158	2,785045	0,002189
LAD	0,475791	1,000639	0,520429	0,001894	2,843642	0,001895
Kantil	5,978664	0,974715	53,36799	0,123730	69,19776	0,124369
Bootstrap	1,673559	1,008164	9,933328	0,049761	10,03989	0,049827

Çizelge 7.36. Senaryo 9 $S_{1,3}(1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	0,455130	1,002153	0,261062	0,000439	2,647684	0,000444
Tukey	0,457497	1,002145	0,260286	0,000438	2,639601	0,000443
Fair	0,720156	1,001654	0,345062	0,000635	1,983062	0,000637
Huber	0,650363	1,001879	0,271817	0,000492	2,093337	0,000496
Hampel	0,791861	1,001415	0,378879	0,000623	1,838479	0,000625
LMS	-0,08108	1,005426	0,644570	0,001375	4,975460	0,001405
MM	0,402766	1,002635	0,246378	0,000421	2,797534	0,000428
S	0,017987	1,003141	0,451942	0,000879	4,380316	0,000889
LTS	0,712143	1,002087	0,340745	0,000747	1,999322	0,000751
LAD	0,392309	1,003506	0,309102	0,000654	2,893773	0,000666
Kantil	4,806546	1,010273	17,10475	0,038614	24,98145	0,038719
Bootstrap	2,174033	0,982528	14,16272	0,009647	14,19301	0,009953

Çizelge 7.37. Senaryo 10 $S_{1,3}(-1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	3,419848	0,993656	1,392122	0,014669	3,408091	0,014709
Tukey	3,414215	0,994023	1,392429	0,014661	3,392434	0,014697
Fair	3,255289	0,989694	1,897965	0,021016	3,473716	0,021122
Huber	3,260142	0,992696	1,782942	0,018643	3,370890	0,018696
Hampel	3,152386	0,988742	2,313873	0,024284	3,641867	0,024411
LMS	3,565499	1,019205	3,765862	0,035297	6,216649	0,035666
MM	3,541173	0,997468	1,144022	0,012075	3,519237	0,012081
S	3,624754	1,020904	2,613242	0,024159	5,253069	0,024596
LTS	3,430618	0,982333	1,424423	0,018016	3,471089	0,018328
LAD	3,477075	0,997488	1,724339	0,017430	3,906091	0,017436
Kantil	5,217554	0,983880	1,112955	0,014527	11,46561	0,014787
Bootstrap	2,808391	0,967539	8,100092	0,111437	8,753587	0,112491

Çizelge 7.38. Senaryo 10 $S_{1,3}(-1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	3,326379	1,010408	0,827374	0,004355	2,586657	0,004464
Tukey	3,321942	1,010605	0,831446	0,004371	2,578977	0,004484
Fair	3,040123	1,013543	1,230777	0,007632	2,312632	0,007815
Huber	3,103639	1,012627	1,015768	0,005589	2,233787	0,005748
Hampel	2,981570	1,011645	1,274130	0,007996	2,237611	0,008132
LMS	3,808562	1,006825	1,992102	0,014154	5,262998	0,014200
MM	3,426613	1,011029	0,761429	0,004240	2,796656	0,004362
S	3,753727	1,007778	1,691485	0,011245	4,767042	0,011306
LTS	3,061093	1,010127	1,761747	0,015426	2,887665	0,015529
LAD	3,258796	1,021219	1,119209	0,007077	2,703776	0,007527
Kantil	5,000735	1,008580	0,600022	0,003929	9,604436	0,004004
Bootstrap	2,284643	1,009383	14,11055	0,084448	14,19157	0,084536

Çizelge 7.39. Senaryo 10 $S_{1,3}(-1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	3,597289	0,994905	0,419052	0,001103	2,970384	0,001129
Tukey	3,596047	0,994915	0,419017	0,001105	2,966384	0,001131
Fair	3,450975	0,988629	0,486209	0,001787	2,591538	0,001916
Huber	3,480051	0,990722	0,427282	0,001375	2,617833	0,001461
Hampel	3,374377	0,989694	0,494085	0,001733	2,382996	0,001839
LMS	3,920426	0,995975	1,043802	0,003557	4,731837	0,003573
MM	3,654385	0,994980	0,401018	0,001097	3,138008	0,001123
S	3,975134	0,995075	0,705237	0,002528	4,606393	0,002553
LTS	3,545525	0,990739	0,524092	0,001704	2,912739	0,001789
LAD	3,695602	0,991306	0,526717	0,001615	3,401782	0,001690
Kantil	5,103461	0,999987	0,361602	0,001049	9,993073	0,001049
Bootstrap	2,730925	0,977072	4,517207	0,019064	5,051458	0,019590

Çizelge 7.40. Senaryo 10 $S_{1,3}(-1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	3,370294	1,004753	0,375307	0,000494	2,253012	0,000517
Tukey	3,368508	1,004747	0,375172	0,000495	2,247987	0,000517
Fair	3,166814	1,003833	0,426452	0,000712	1,787907	0,000727
Huber	3,219836	1,003995	0,389922	0,000589	1,877922	0,000605
Hampel	3,106907	1,003759	0,443488	0,000701	1,668731	0,000715
LMS	3,854024	1,002038	0,842492	0,001489	4,279898	0,001494
MM	3,425140	1,004425	0,355338	0,000489	2,386363	0,000508
S	3,817156	1,002282	0,695816	0,001181	3,997871	0,001186
LTS	3,206387	1,003888	0,453450	0,000703	1,908819	0,000719
LAD	3,408019	1,005476	0,475715	0,000715	2,458235	0,000745
Kantil	5,168578	0,999030	0,410614	0,000707	10,45049	0,000708
Bootstrap	2,534105	0,996047	2,656239	0,007886	2,941507	0,007902

Çizelge 7.41. Senaryo 11 $S_{1,7}(1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,614762	1,018854	0,669259	0,008909	0,817668	0,009265
Tukey	1,615265	1,018846	0,667191	0,008893	0,815212	0,009249
Fair	1,661291	1,021678	0,817199	0,011479	0,931922	0,011949
Huber	1,653225	1,021277	0,739181	0,010482	0,859434	0,010935
Hampel	1,710609	1,018215	0,825726	0,010524	0,909474	0,010855
LMS	1,313597	1,025407	2,078825	0,021958	2,549974	0,022604
MM	1,563701	1,018369	0,728968	0,008759	0,919324	0,009097
S	1,437634	1,013704	1,599221	0,016745	1,915476	0,016933
LTS	1,544224	1,022874	0,697379	0,009395	0,905111	0,009918
LAD	1,558933	1,019833	0,982652	0,012397	1,177192	0,012791
Kantil	4,586210	1,019549	51,09310	0,384612	57,78164	0,384995
Bootstrap	1,827690	1,025166	2,177679	0,028076	2,207370	0,028709

Çizelge 7.42. Senaryo 11 $S_{1,7}(1,1,0)$, $n=20$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,679989	0,999368	0,466156	0,002834	0,568563	0,002834
Tukey	1,680151	0,999379	0,466165	0,002834	0,568468	0,002834
Fair	1,734549	1,000252	0,540688	0,003200	0,611152	0,003200
Huber	1,725200	0,999885	0,499478	0,002922	0,574993	0,002922
Hampel	1,766512	1,000575	0,550171	0,003066	0,604687	0,003066
LMS	1,470627	0,996489	1,775998	0,010422	2,056234	0,010435
MM	1,630599	0,998715	0,446400	0,002710	0,582857	0,002712
S	1,525979	0,993641	1,312134	0,008686	1,536830	0,008726
LTS	1,675200	0,998891	0,575763	0,003176	0,681258	0,003177
LAD	1,613545	1,000967	0,659753	0,004313	0,809101	0,004314
Kantil	4,546805	1,007134	10,79072	0,104328	17,27694	0,104379
Bootstrap	1,867629	1,004974	1,116392	0,008994	1,133915	0,009019

Çizelge 7.43. Senaryo 11 $S_{1,7}(1,1,0)$, $n=30$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,705502	0,999348	0,363599	0,001119	0,450328	0,001120
Tukey	1,707934	0,999305	0,361652	0,001111	0,446955	0,001111
Fair	1,790289	0,999352	0,428350	0,001274	0,472329	0,001275
Huber	1,760165	0,999535	0,395683	0,001176	0,453203	0,001176
Hampel	1,805908	0,999739	0,426393	0,001269	0,464064	0,001269
LMS	1,447978	1,002987	1,127829	0,003674	1,432558	0,003682
MM	1,685302	0,999335	0,363598	0,001132	0,462632	0,001133
S	1,607589	0,995799	0,813992	0,002948	0,967978	0,002965
LTS	1,665042	0,999728	0,461702	0,001445	0,573899	0,001445
LAD	1,739607	0,997814	0,500958	0,001619	0,568762	0,001625
Kantil	4,070167	1,006585	3,601134	0,012445	7,886728	0,012488
Bootstrap	1,992729	1,000322	0,948590	0,002561	0,948643	0,002561

Çizelge 7.44. Senaryo 11 $S_{1,7}(1,1,0)$, $n=40$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	1,636544	1,003115	0,243351	0,000415	0,375452	0,000424
Tukey	1,637334	1,003108	0,243343	0,000415	0,374869	0,000425
Fair	1,708064	1,002961	0,277009	0,000461	0,362237	0,000469
Huber	1,700474	1,002482	0,259147	0,000433	0,348863	0,000439
Hampel	1,762821	1,001963	0,268185	0,000432	0,324438	0,000436
LMS	1,392278	1,002700	0,949185	0,001724	1,318511	0,001731
MM	1,611277	1,003169	0,251070	0,000432	0,402176	0,000441
S	1,372870	1,005512	0,605768	0,001139	0,999060	0,001169
LTS	1,632083	1,002685	0,277069	0,000501	0,412432	0,000507
LAD	1,562509	1,004189	0,360666	0,000592	0,552064	0,000609
Kantil	4,217246	0,997579	3,520846	0,005242	8,437026	0,005247
Bootstrap	1,939560	1,001491	0,586377	0,000954	0,590029	0,000955

Çizelge 7.45. Senaryo 12 $S_{1,7}(-1,1,0)$, $n=15$ için simülasyon sonuçları

Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,253175	1,002765	0,709816	0,008572	0,773913	0,008579
Tukey	2,251974	1,002699	0,707453	0,008529	0,770944	0,008536
Fair	2,192772	1,002356	0,782854	0,009422	0,820015	0,009427
Huber	2,192324	1,003379	0,729855	0,008741	0,766844	0,008753
Hampel	2,155789	1,002459	0,776334	0,009047	0,800605	0,009053
LMS	2,538031	0,983978	2,747607	0,03196	3,037084	0,032217
MM	2,338467	1,001199	0,771783	0,009291	0,886342	0,009293
S	2,428816	0,999474	1,847078	0,021083	2,030961	0,021084
LTS	2,201355	1,011635	0,935932	0,011078	0,976476	0,011213
LAD	2,314234	0,999061	0,925505	0,011167	1,024248	0,011168
Kantil	4,060139	0,991296	1,177721	0,014510	5,421897	0,014586
Bootstrap	2,072395	0,975502	6,741485	0,167886	6,746726	0,168486

Çizelge 7.46. Senaryo 12 $S_{1,7}(-1,1,0)$, $n=20$ için simülasyon sonuçları

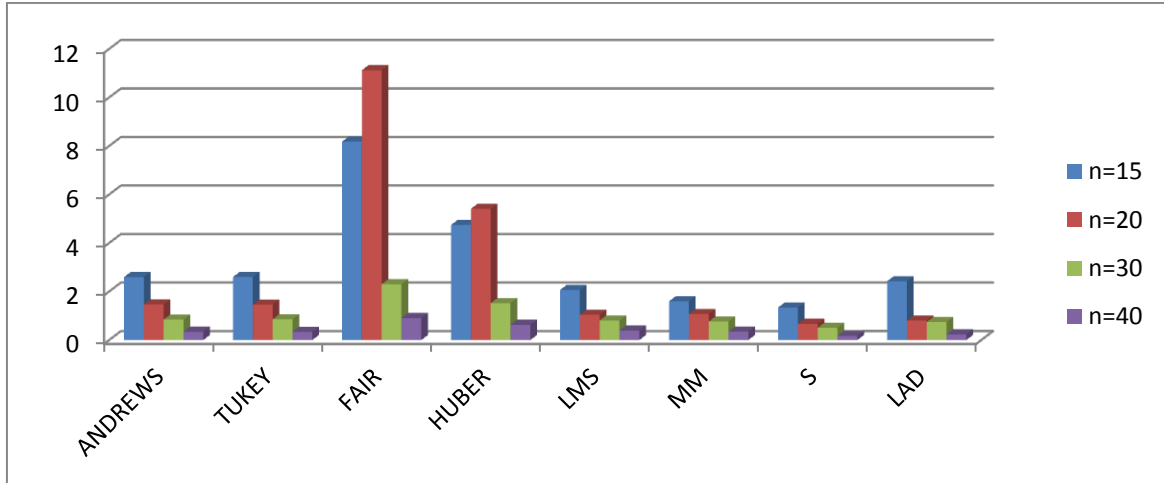
Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,248409	1,005244	0,545068	0,003695	0,606775	0,003723
Tukey	2,247156	1,005291	0,545268	0,003699	0,606354	0,003727
Fair	2,179213	1,005312	0,627252	0,004102	0,659369	0,00413
Huber	2,189961	1,005681	0,576157	0,003862	0,612242	0,003894
Hampel	2,146589	1,005316	0,598604	0,003985	0,620093	0,004013
LMS	2,378570	1,006418	1,601397	0,012487	1,744712	0,012528
MM	2,291188	1,005149	0,561748	0,003918	0,646538	0,003945
S	2,273946	1,019476	1,461627	0,009081	1,536673	0,009461
LTS	2,235828	1,006841	0,685318	0,004494	0,740933	0,00454
LAD	2,284793	1,006912	0,809567	0,004916	0,890674	0,004964
Kantil	4,025428	1,004098	1,370329	0,009074	5,472688	0,00909
Bootstrap	2,014554	1,003955	1,383417	0,008226	1,383628	0,008242

Çizelge 7.47. Senaryo 12 $S_{1,7}(-1,1,0)$, $n=30$ için simülasyon sonuçları

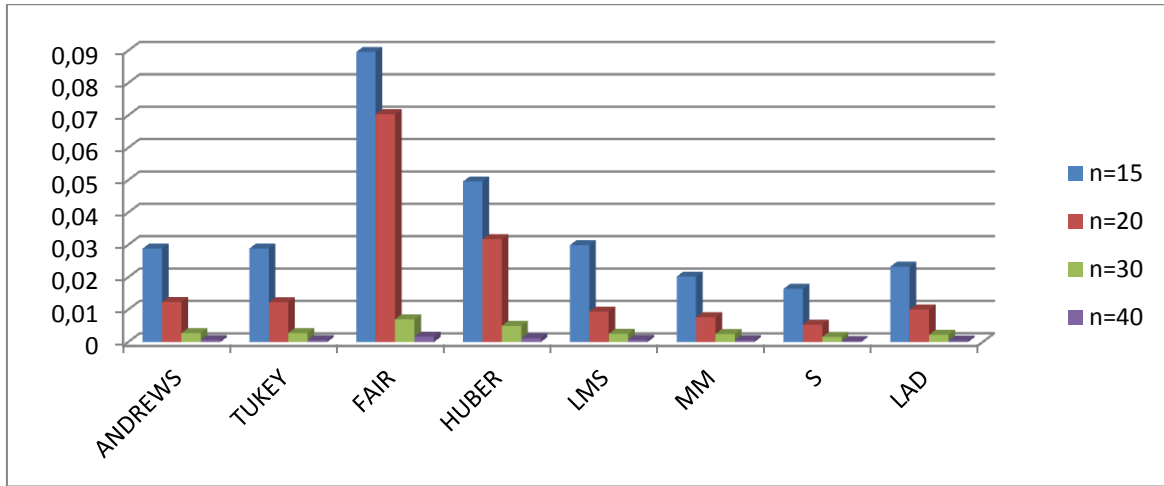
Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,309752	0,997029	0,340882	0,001067	0,436829	0,001075
Tukey	2,308462	0,997042	0,339957	0,001065	0,435106	0,001074
Fair	2,206413	0,998507	0,336019	0,001096	0,378625	0,001098
Huber	2,232501	0,99812	0,325744	0,001059	0,379801	0,001062
Hampel	2,181371	0,998609	0,350341	0,001111	0,383236	0,001113
LMS	2,528051	0,997866	1,148457	0,003401	1,427294	0,003405
MM	2,333118	0,996827	0,334308	0,001041	0,445275	0,001051
S	2,50046	0,998004	0,804706	0,002561	1,055165	0,002565
LTS	2,313456	0,998947	0,412372	0,001386	0,510626	0,001387
LAD	2,288539	0,999370	0,390509	0,001301	0,473764	0,001301
Kantil	3,850309	1,004539	0,583408	0,002149	4,007053	0,002169
Bootstrap	1,983403	1,000959	0,706798	0,002009	0,707073	0,002011

Çizelge 7.48. Senaryo 12 $S_{1,7}(-1,1,0)$, $n=40$ için simülasyon sonuçları

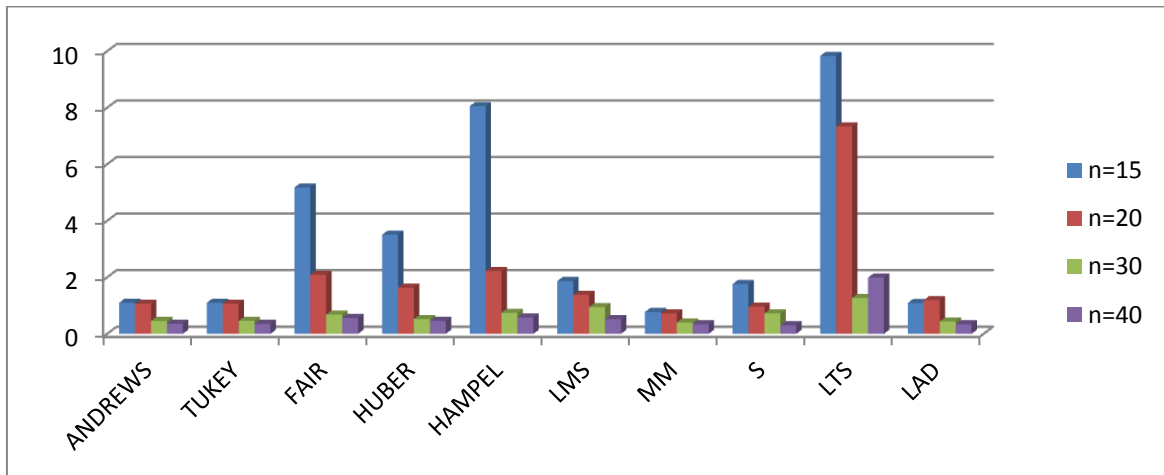
Yöntemler	Ortalama		Varyans		HKO	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
Andrews	2,292764	1,001571	0,208799	0,000338	0,294511	0,000341
Tukey	2,291382	1,001577	0,209262	0,000338	0,294165	0,000341
Fair	2,204887	1,002052	0,227638	0,000374	0,269617	0,000378
Huber	2,234069	1,001555	0,216766	0,000354	0,271554	0,000356
Hampel	2,187579	1,001785	0,226622	0,000369	0,261807	0,000372
LMS	2,552524	0,999807	0,946597	0,001578	1,251879	0,001578
MM	2,304039	1,001830	0,215601	0,000349	0,308041	0,000352
S	2,472835	1,002543	0,639225	0,001105	0,862798	0,001112
LTS	2,323337	1,000969	0,215427	0,000369	0,319973	0,000371
LAD	2,266743	1,004114	0,285399	0,000469	0,356551	0,000487
Kantil	4,048769	1,000132	0,303375	0,000584	4,500832	0,000584
Bootstrap	2,031056	1,000626	0,582736	0,001215	0,583701	0,001216



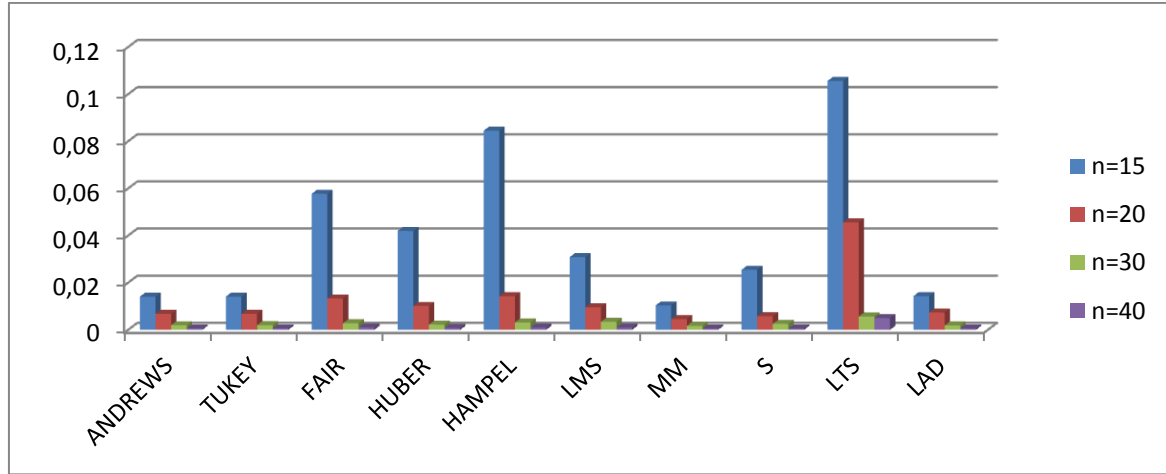
Şekil 7.5. Senaryo1 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



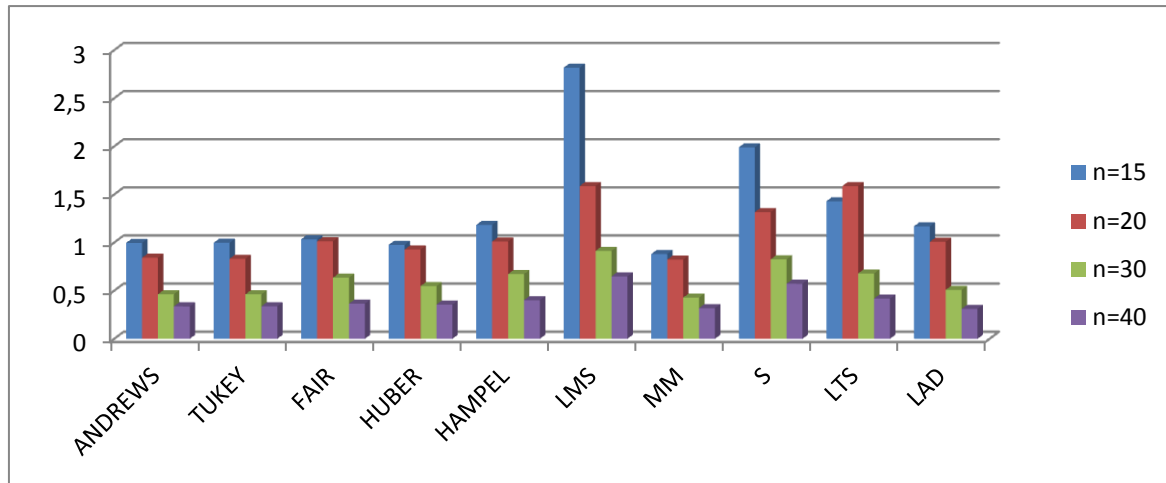
Şekil 7.6. Senaryo1 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



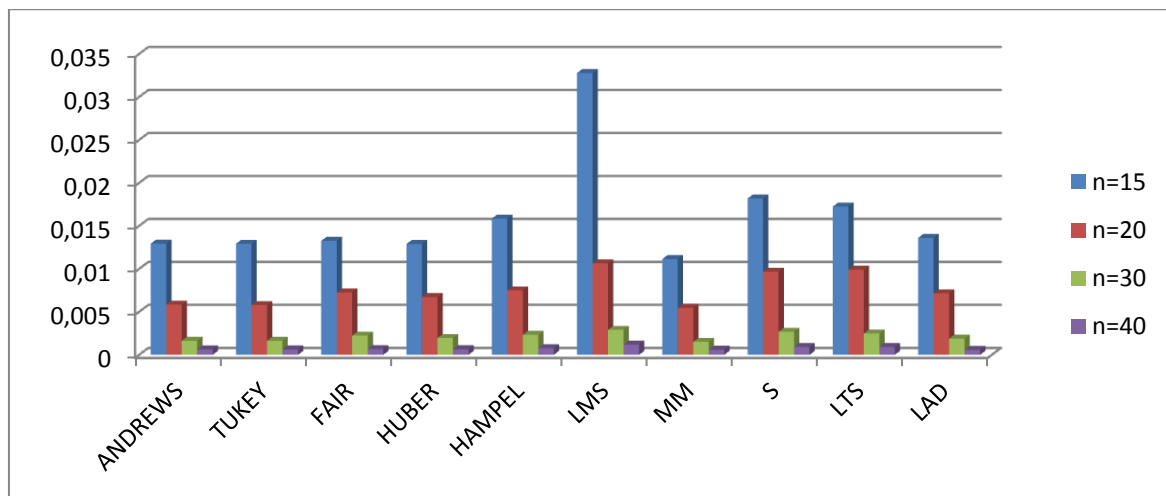
Şekil 7.7. Senaryo2 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



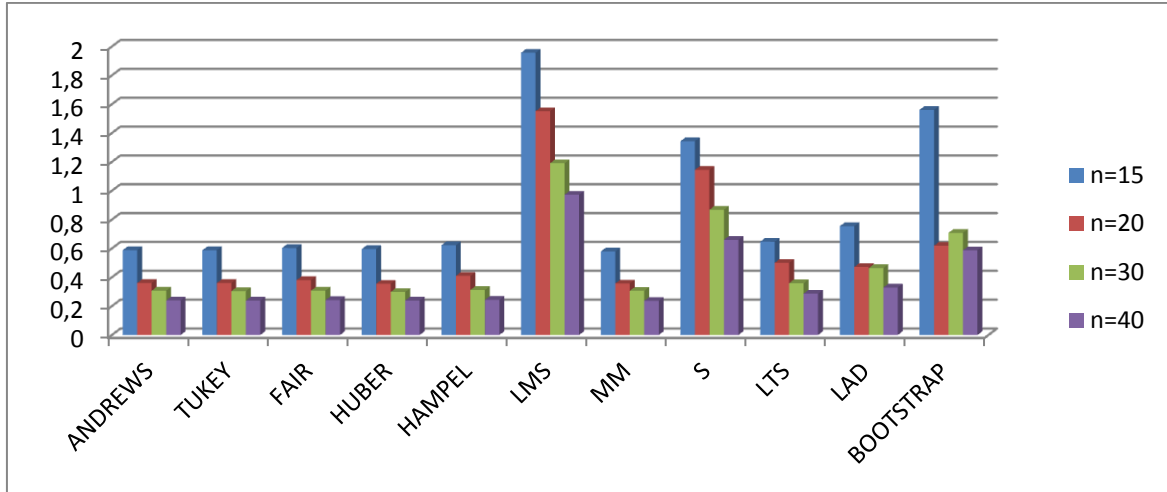
Şekil 7.8. Senaryo2 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



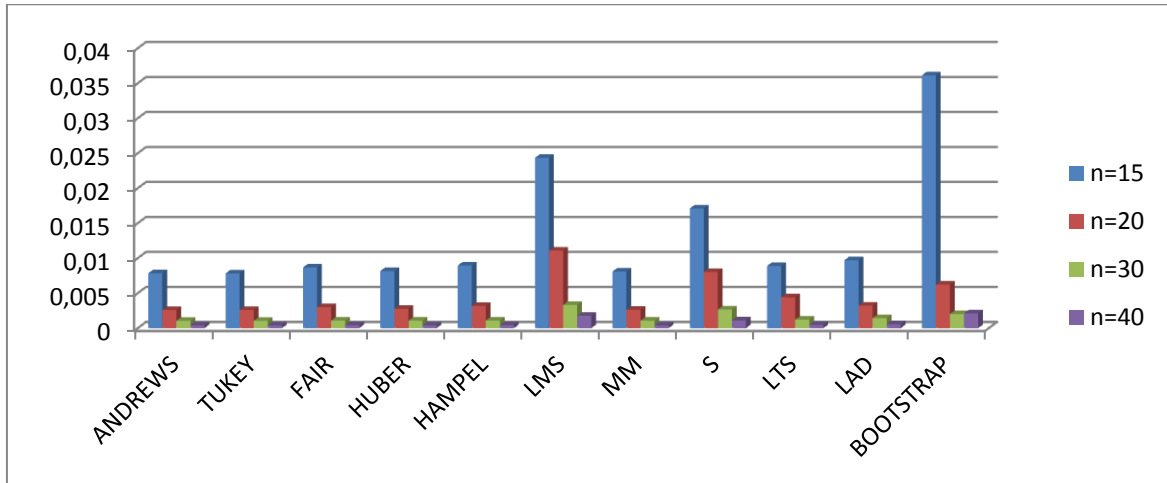
Şekil 7.9. Senaryo 3 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



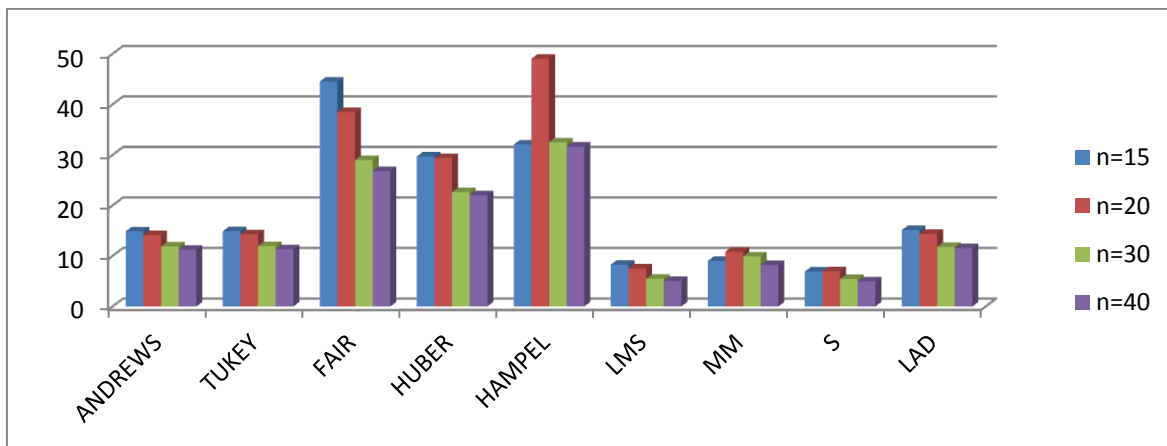
Şekil 7.10. Senaryo 3 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



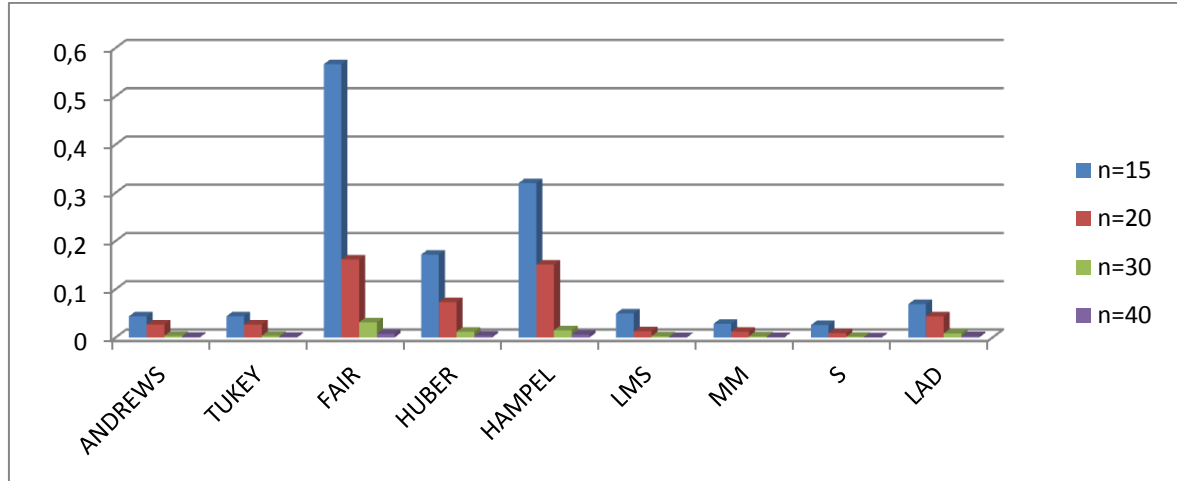
Şekil 7.11. Senaryo 4 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



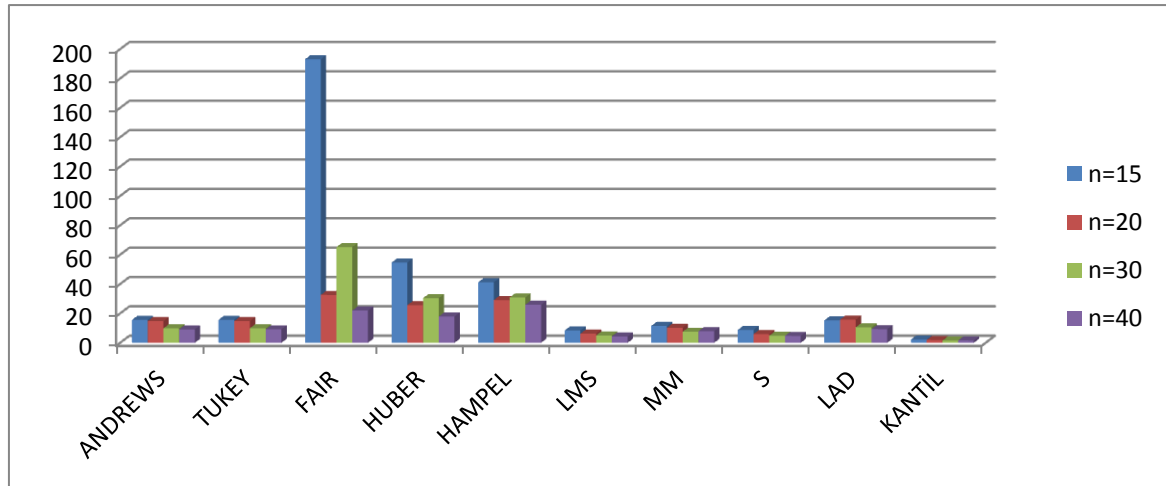
Şekil 7.12. Senaryo 4 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



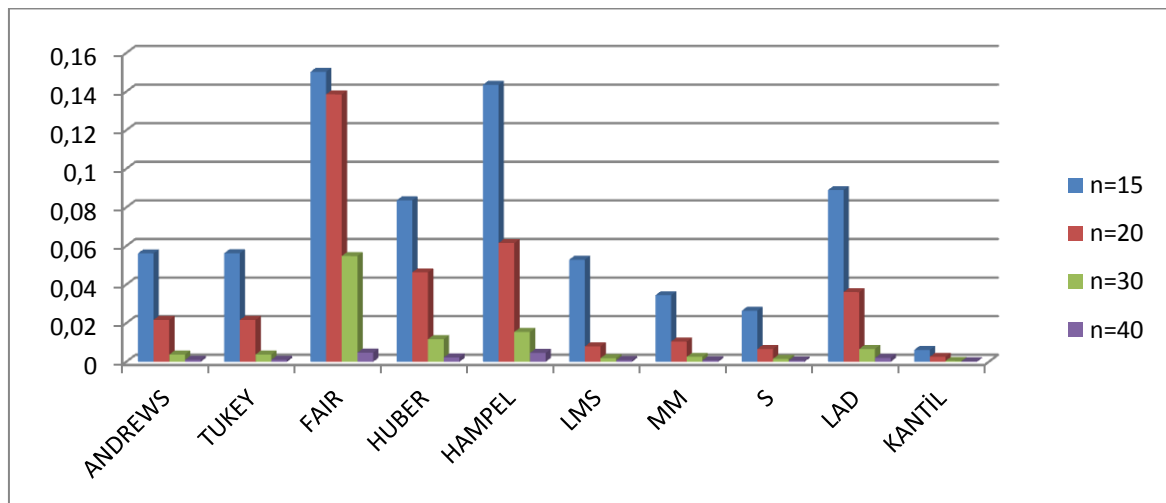
Şekil 7.13. Senaryo 5 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



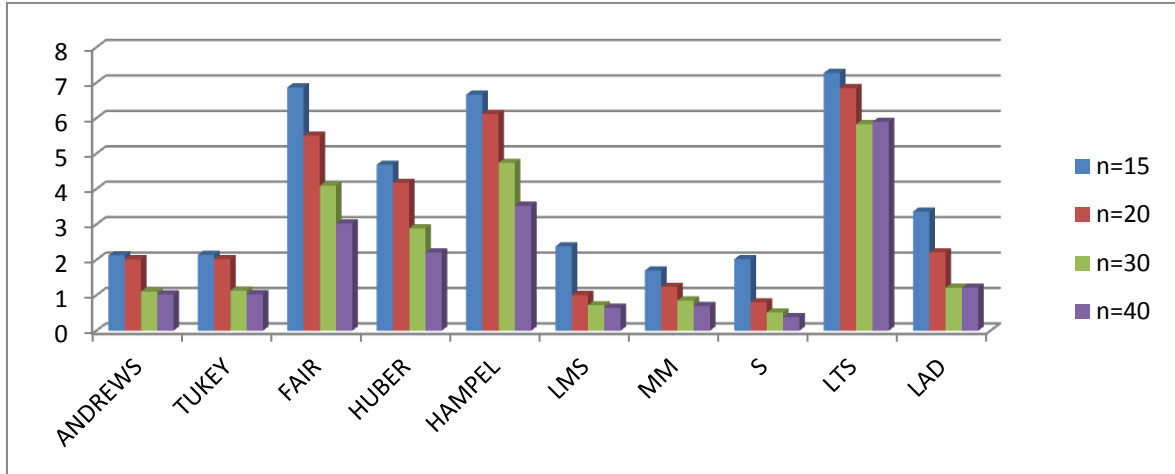
Şekil 7.14. Senaryo 5 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



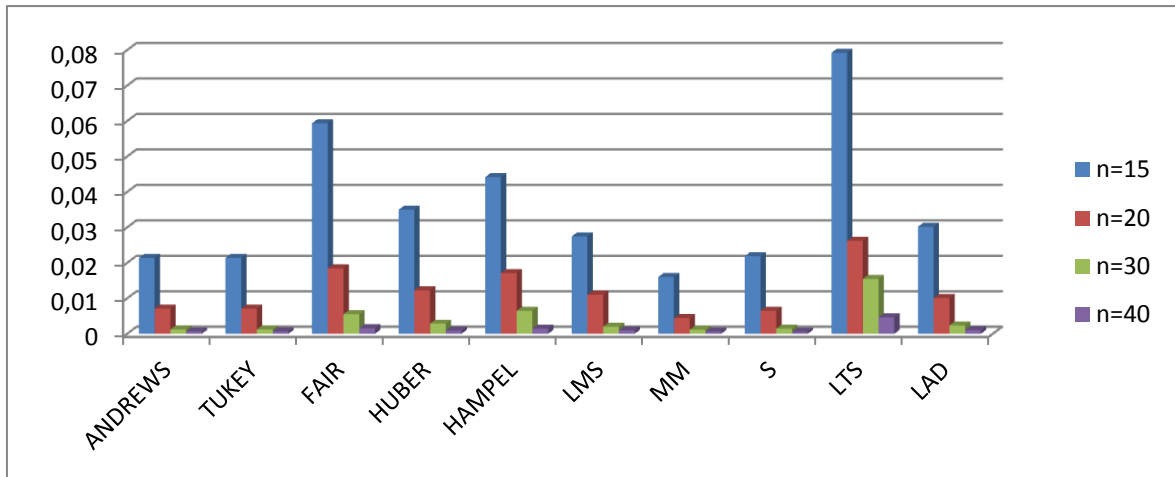
Şekil 7.15. Senaryo 6 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



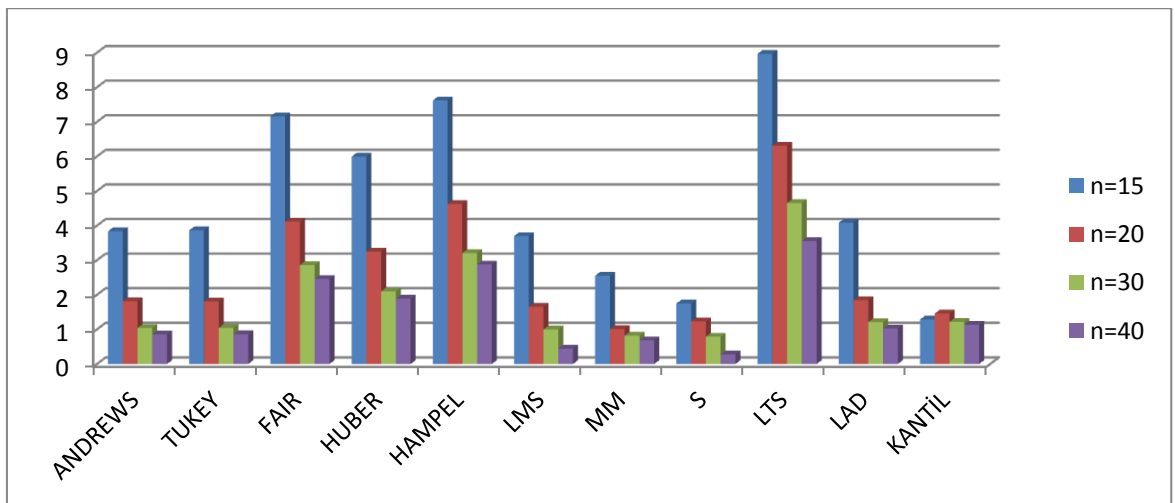
Şekil 7.16. Senaryo 6 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



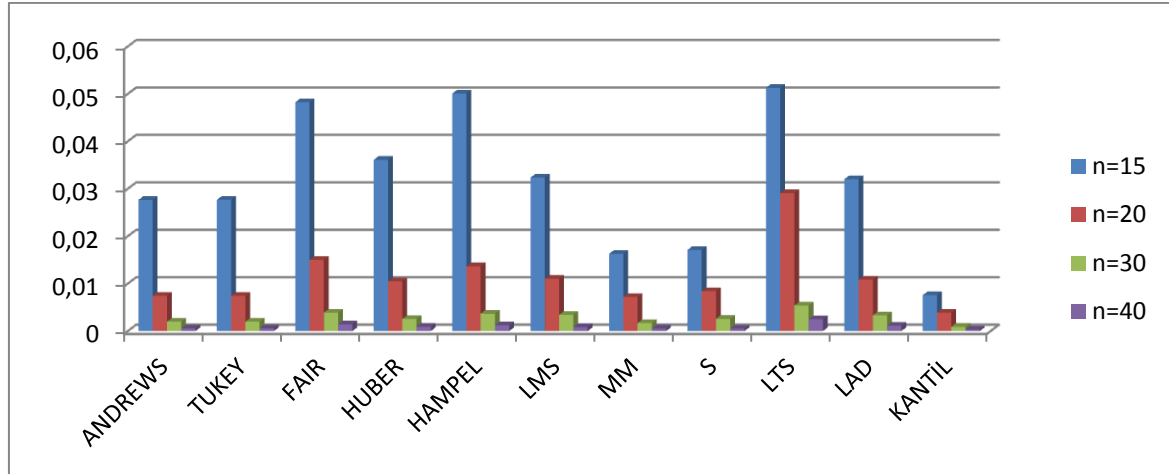
Şekil 7.17. Senaryo 7 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



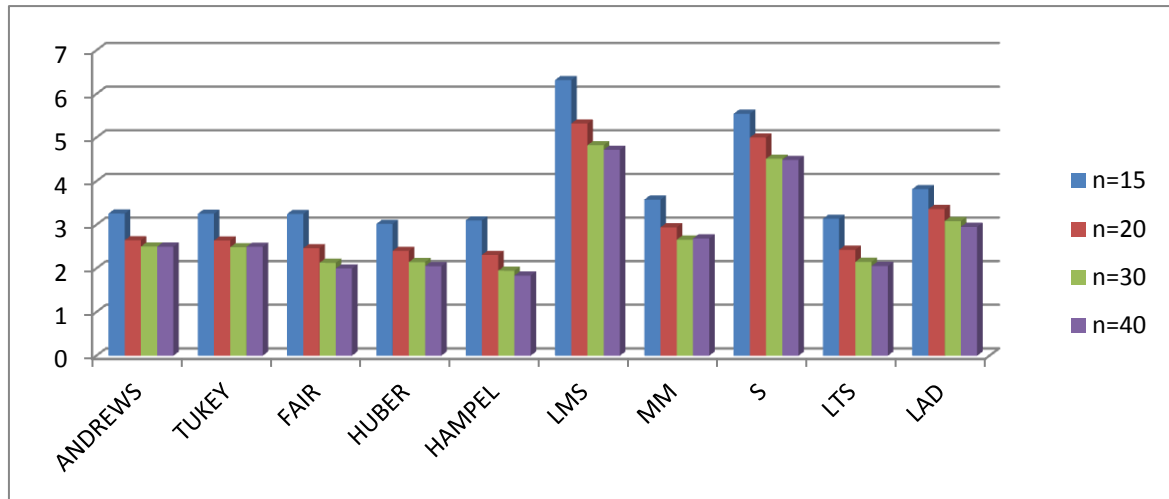
Şekil 7.18. Senaryo 7 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



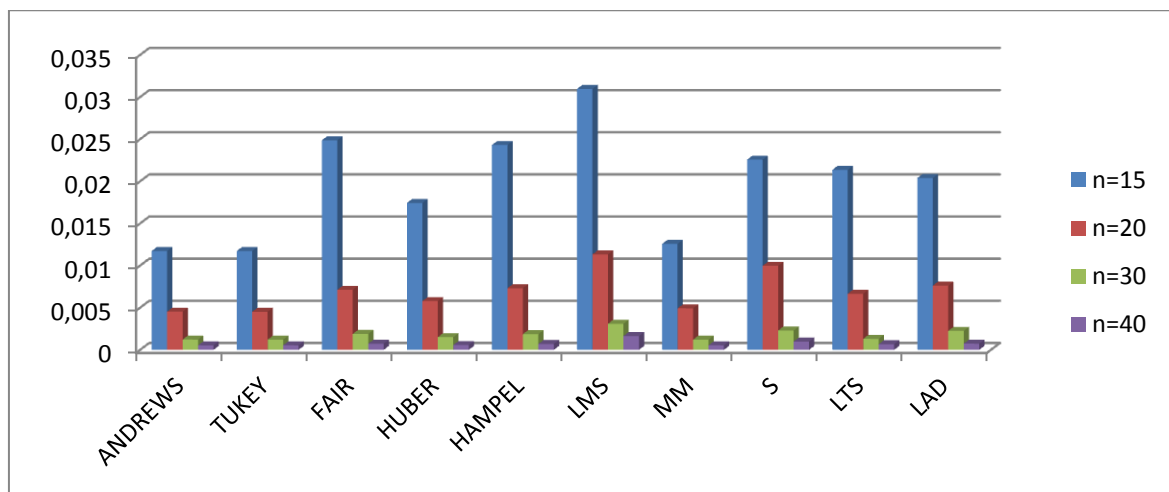
Şekil 7.19. Senaryo 8 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



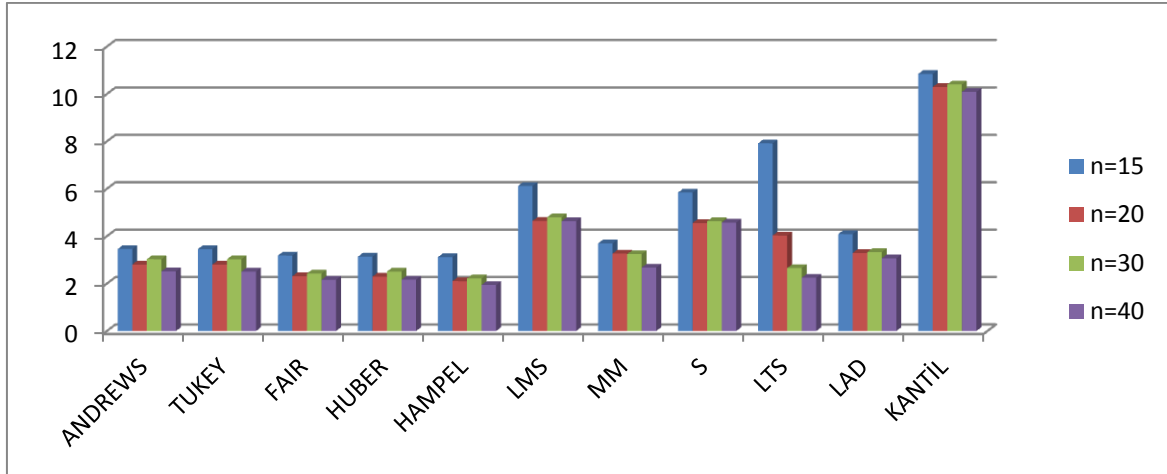
Şekil 7.20. Senaryo 8 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



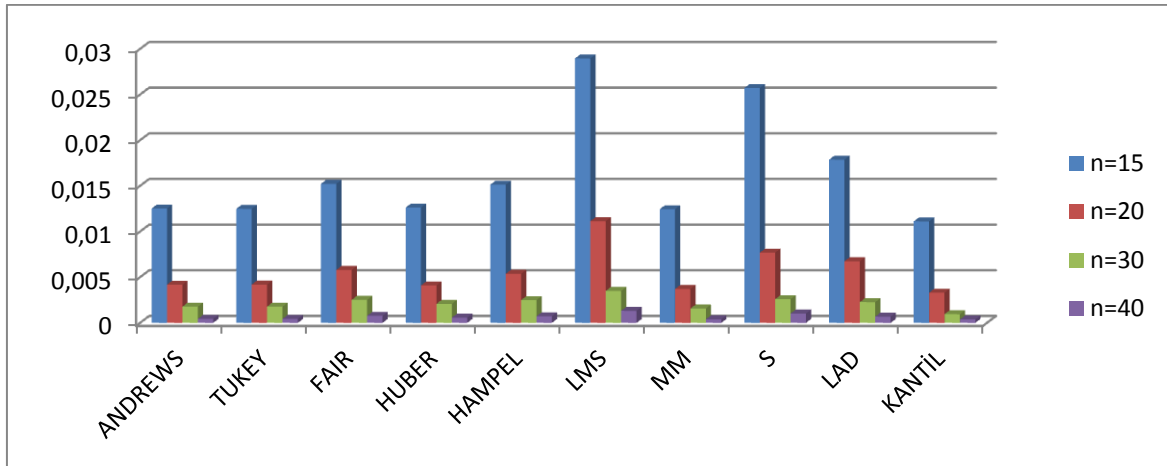
Şekil 7.21. Senaryo 9 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



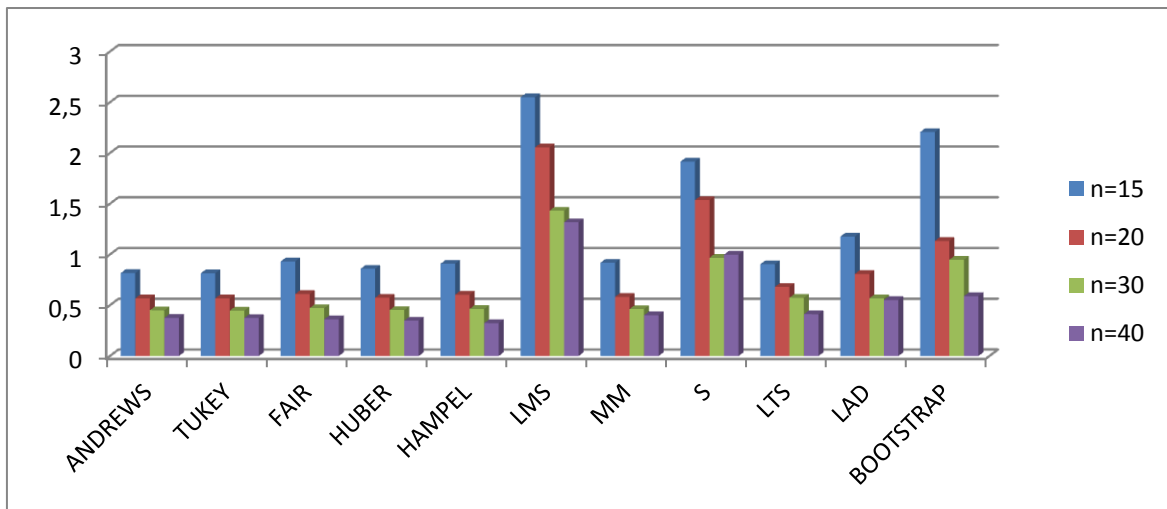
Şekil 7.22. Senaryo 9 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



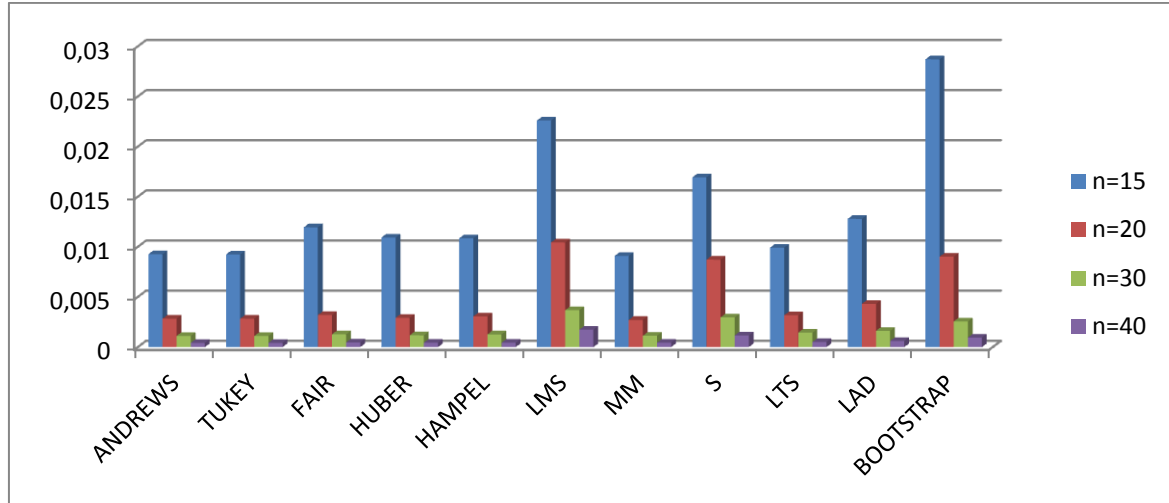
Şekil 7.23. Senaryo 10 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



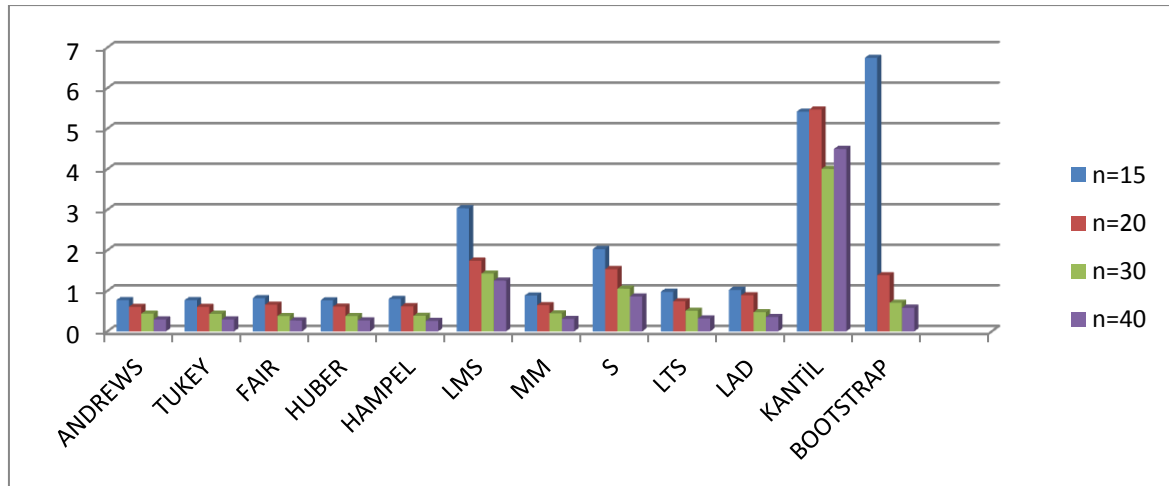
Şekil 7.24. Senaryo 10 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



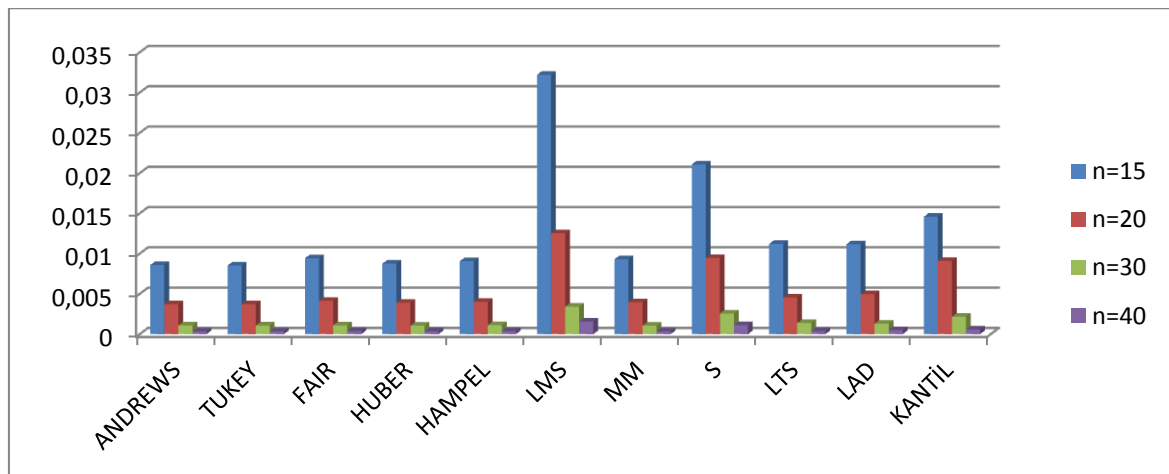
Şekil 7.25. Senaryo 11 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



Şekil 7.26. Senaryo 11 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi



Şekil 7.27. Senaryo 12 tüm örneklerde $\hat{\beta}_0$ için HKO sonuçlarının grafik gösterimi



Şekil 7.28. Senaryo 12 tüm örneklerde $\hat{\beta}_1$ için HKO sonuçlarının grafik gösterimi

8. SONUÇ ve ÖNERİLER

Bu çalışmada hata terimlerinin kararlı dağılıma sahip olduğu basit doğrusal regresyon modeline sağlam tahmin yöntemleri uygulanmıştır. Kararlı dağılımın farklı parametre değerleri için kuyruk kalınlığı ve çarpıklık yapısında değişiklikler yaparak kullanılan yöntemlerin bu durumdan nasıl etkilendiği araştırılmıştır. Yapılan simülasyon çalışmasından şu bulgular elde edilmiştir.

Simetrik yapılı ve $\alpha=0,7$ ile çalışmanın kuyruk kalınlığının en fazla olduğu durumda $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin edicilerinin ikisinde de S tahmin edicilerinin yüksek performans gösterdiği görülmüştür. Yüksek kırılma noktasına sahip olduğu bilinen bu tahminin bozulmalara karşı daha dirençli olduğu bu çalışma içinde geçerliliğini göstermiştir.

Simetrik yapılı ve $\alpha=1$ parametrelili kararlı dağılım (bu aynı zamanda Cauchy Dağılımını ifade eder) için S tahmini ile birlikte MM tahmininin de en iyi performansı gösterdiği gözlenmiştir.

Kuyruğun biraz daha incelendiği $\alpha=1,3$ kararlı dağılım için MM tahminin $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin edicilerinin ikisi içinde en yüksek performans gösterdiği görülmüştür.

Dağılım normale yaklaştıkça yani $\alpha=1,7$ olduğunda MM tahmini ile birlikte M tahminlerinin performansının arttığı gözlenmiştir.

Simetrik olmayan kararlı dağılımlar iki açıdan değerlendirilebilir. Sağa ve sola çarpık durumlarda yöntemlerin performanslarında farklılık gözlenmiştir.

Sağa çarpık ve kalın kuyruklu dağılımda yine S tahminleri en iyi performansı göstermiştir. Aynı kuyruk kalınlığında sola çarpık duruma bakıldığında Kantil regresyon modeli en iyi yöntem olarak karşımıza çıkar. Bu durum sola çarpık durumda negatif değer alan hata terimlerinin daha fazla olması ve dolayısıyla kantil regresyon yönteminin amaç fonksiyonunun minimum olması ile açıklanabilir. Simetrik olmayan senaryolarda LMS tahmininin de iyi olduğu görülmüştür. Bunun sebebi çarpık verilerde konum ölçüsü olarak medyanın ortalamadan daha uygun olması ile ilişkilendirilebilir. Kalın kuyruklu ve sağa

çarpık diğer bir durum $\alpha=1$, $\beta=1$ için küçük örnek çaplarında MM, örnek çapı büyüdükçe S tahmininin yüksek performans gösterdiği gözlenmiştir.

Sola çarpık ve kalın kuyruklu $\alpha=1$, $\beta=-1$ için küçük örnek çaplarında her iki tahmin edici için kantil regresyonun, $n=30$ ve 40 örnek çapları için $\hat{\beta}_0$ tahmin edicisi için S, $\hat{\beta}_1$ tahmin edicisi için ise Kantil regresyonun performansı yüksektir.

Sağa çarpık ve kuyruğun incelendiği $\alpha=1,3$ ve $\beta=1$ durumu için tüm örnek çaplarında MM ve M tahmin edicilerin en iyi performansa sahip olduğu görülmüştür. Aynı kuyruk kalınlığında sola çarpık dağılıma bakıldığında tüm örnek çaplarında $\hat{\beta}_0$ tahmin edicisi için Hampel olmak üzere M tahmininin, $\hat{\beta}_1$ tahmin edicisi için ise Kantil regresyon yönteminin performansı yüksektir.

Kuyruk kalınlığının daha da azaldığı $\alpha=1,7$ ve $\beta=1$ için M tahmin edicilerin performansı artmıştır. Örnek çapının küçük olduğu durumlarda $\hat{\beta}_1$ tahmin edicisi için MM tahminleri iyi iken, örnek büyüdükçe Tukey'in yaklaşımı ağırlıklı olmak üzere M tahminleri her iki tahmin edici içinde en iyi yöntemdir.

Normale yakın ve sola çarpık yapıda tüm örnek çapları ve her iki tahmin edici için M tahminleri en iyi performansa sahiptir.

Sonuç olarak kuyruk kalınlığının fazla olduğu kararlı yapılarda S tahminlerini, sola çarpık yapının söz konusu olduğu durumlarda Kantil regresyonu tercih etmek uygulayıcılara doğru tahminler sağlayacaktır. Bunun yanı sıra MM ve M tahminleri normallik ihlalinde diğer sağlam yöntemlere göre daha iyi bir alternatif olarak görülebilir.

KAYNAKLAR

- Abdi, H. (2010). The Least Squares. In N.J. Salkind, D.M. Dougherty and B.Frey (Eds.), *Encyclopedia of Research Design*, Thousand Oaks(CA), Sage Publications, pp.705-708.
- Akar, C. (2013). Gelişmekte Olan Piyasalarda Finansal Piyasa İstikrarının Kantil Regresyon Yöntemiyle Test Edilmesi. *Doğuş Üniversitesi Dergisi*, 14(1), 1-9.
- Aktükün, T.A. (2005). Asal Bileşenler Analizine Bootstrap Yaklaşımı. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik e-Dergisi*, (1), 1-10.
- Altın, A. ve Şenoğlu, B. (2008). Konum Parametresinin Bazı Sağlam Tahmin Edicilerinin Örnekleme Alanında Kullanılması ve Bir Tarım Uygulaması, *Afyon Kocatepe Üniversitesi Fen bilimleri Dergisi*, 8(1), 291-306.
- Arya, K.V., Gupta, P., Kalra P.K., Mitra P. (2007). Image Registration Using Robust M-Estimators. *Pattern Recognition Letters*, 28, 1957-1968.
- Birkes, D., Dodge,Y. (1993). *Alternative Methods of Regression*. (First Published), New York: John Wiley and Sons, Inc, 29, 85, 87-89, 99,100,191-192, 205
- Blattberg R., Sargent T. (1971). Regression with Non-Gaussian Stable Disturbances: Some Sampling Results. *Econometrica*, 39(3), 501-510.
- Burmaoğlu, S., Polat, M., Meydan C.H. (2013). Örgütsel Davranış Alanında İlişkisel Analiz Yöntemleri ve Türkçe Yazında Aracılık Modeli Kullanımı Üzerine Bir İnceleme. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 13(1), 13-26.
- Cameron A.C., Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*. (First Published) New York: Cambridge University Press, 361.
- Chen C., Wei Y. (2005). Computational Issues For Quantile Regression. *The Indian Journal of Statistics, Special Issue on Quantile Regression and Related Methods*, 67(2), 399-417.
- Çağlayan, E. ve Arıkan E. (2011). Determinants of House Prices in Istanbul: A Quantile Regression Approach. *Quality & Quantity*, 45(2), pp. 305-3017.
- Çolak, Ö.F. ve Öztürkler, H. (2012). Tasarrufun Belirleyicileri: Küresel Tasarruf Eğiliminde Değişim ve Türkiye’de Hanehalkı Tasarruf Eğiliminin Analizi, *Bankacılar Dergisi*, 23(82), 3-44.
- Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application* (First Published) New York: Cambridge University Press, ix.
- Davidson R. and MacKinnon J.G. (2006). Bootstrap Methods in Econometrics. In K. Patterson and T.C. Mills (Eds.), *Palgrave Handbook of Econometrics Volume 1 Theoretical Econometrics*, Basingstoke. Palgrave Macmillan, pp. 812-838.

Dielman, T.E. (1991). ***Applied Regression Analysis for Business and Economics***. (First Edition) Boston: PWS-KENT Publishing Company, 2, 62.

İnternet: Doornik, A.J. Robust Estimation Using Least Trimmed Squares.URL:http://www.webcitation.org/query?url=http%3A%2F%2Fcreates.au.dk%2Ffileadmin%2Fsite_files%2Ffiler_oekonomi%2Fsubsites%2Fcreates%2FSeminar_Papers%2F2011%2FELTS.pdf&date=2014-12-11, Son Erişim Tarihi: 11.12.2014.

Draper, R.N. and Smith, H. (1966). ***Applied Regression Analysis***. (Second Edition) New York: John Wiley and Sons, 342.

Efron, B. and Tibshirani, R.J. (1993) ***An Introduction to the Bootstrap***. (First Published), New York: Chapman and Hall, 1.

Fama, E.F. (1965). The Behavior of Stock-Market Prices. ***Journal of Business***, 38, 34-105.

Fama, E.F. and Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions. ***Journal of the American Statistical Association***, 66, 331-338.

İnternet: Fox, John. Robust Regression Appendix to An R and S-PLUS Companion to Applied Regression. URL: <http://www.webcitation.org/query?url=http%3A%2F%2Fcran.r-project.org%2Fdoc%2Fcontrib%2FFox-Companion%2Fappendix-robust-regression.pdf&date=2014-12-17>, Son Erişim Tarihi: 17.12.2014.

Fox, J. (2008). ***Applied Regression Analysis and Generalized Linear Models***. (Second Edition). Thousand Oaks, Sage Publications, 268, 597, 598.

İnternet: Frain, J.C. Total Returns on Equity Indices, Fat Tails and the α - Stable Distribution.URL:http://www.webcitation.org/query?url=http%3A%2F%2Fwww.tcd.ie%2FEconomics%2Fstaff%2Ffrainj%2FStable_Distribution%2FTotal%2520Returns.pdf&date=2014-12-11, Son Erişim Tarihi: 11.12.2014.

Ghahfarokhi, M.A.B. and Ghahfarokhi P.B. (2009). Applications of Stable Distributions in Time Series Analysis, Computer Sciences and Financial Markets. ***World Academy of Science ,Engineering and Technology***, 25, 1027-1031.

Gill, I.S., Montenegro, C.E. and Dömeland, D. (2002). ***Crafting Labor Policy: Techniques and Lessons from Latin America***. Washington: World Bank, 249.

Gschwandtner, M. and Filzmoser, P. (2012). Computing Robust Regression Estimators: Developments since Dutter (1977). ***Austrian Journal of Statistics***, 41(1), 41-58.

Gujarati D.N. (1999) ***Temel Ekonometri*** (çev. Ü. Şenesen ve G.Günlük Şenesen). İstanbul: Literatür Yayıncılık (Eserin orijinali 1995’de yayımlandı), 15, 16, 59-66.

Guo, Y. (2003). ***Application and Computation of Huber’s M estimator in the Global Positioning System***. Unpublished Master’s Thesis, McCill University, Montreal

- Gürünlü Alma., Ö. ve Vupa., Ö. (2008). Regresyon Analizinde Kullanılan En Küçük Kareler ve En Küçük Medyan Kareler Yöntemlerinin Karşılaştırılması, **Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi Fen Dergisi (E-Dergi)**, 3(2), 219-229
- Gürünlü Alma, Ö. (2011). Comparison of Robust Regression Methods in Linear Regression. **International Journal of Contemporary Mathematical Sciences**, 6(9), 409-421.
- Hampel, F.R. (1974). The Influence Curve and Its Role in Robust Estimation, **Journal of the American Statistical Association**, 69, pp.383-393.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A. (1986). **Robust Statistics: The Approach Based on Influence Functions**. (First Published). New York: John Wiley and Sons, 11.
- Hawkins, D.M., ve Olive, D.J. (2002). In consistency of resampling algorithms for high breakdown regression estimators and a new algorithm, **Journal of the American Statistical Association**, 97, 136-148.
- İnternet: Heikkilä, J. Robust Regression. URL: <http://www.webcitation.org/query?url=http%3A%2F%2Fwww.ee.oulu.fi%2F%7Ejth%2Frobust.pdf&date=2014-12-13>, Son Erişim Tarihi: 13.12.2014.
- İnternet: Hoffmann, J.P. Linear Regression Analysis: Applications and Assumptions. URL: <http://www.webcitation.org/query?url=https%3A%2F%2Fsociology.byu.edu%2FHoffmann%2FSiteAssets%2FHoffmann%2520%2520Linear%2520Regression%2520Analysis%2520second%2520edition.pdf&date=2014-12-11>, Son Erişim Tarihi: 11.12.2014.
- Hössjer, O. (1994). Rand-Based Estimates in the Linear Model with High Breakdown Point. **Journal of the American Statistical Association**, 89, 149-158.
- Huber, P.J. (1981). **Robust Statistics**, (First Published), New York: John Wiley and Sons, 1,4,5,7,43,54.
- Hubert, M., Rousseeuw, P.J. and S. Van Aelst. (2003). Robustness, In B. Sundt and J. Teugels (Eds.) **Encyclopedia of Actuarial Sciences**, New York: John Wiley, 6
- Hund E., Massart, D.L., Verbeke, J.S. (2002). Robust Regression and Outlier Detection in the Evaluation of Robustness Tests with Different Experimental Designs. **Analytica Chimica Acta**, 463(1), 53-73
- Janicki A. And Weron A. (1994). Can One See α - Stable Variables and Processes ?. **Statistical Science**, 9(1), 109-126.
- Jammalamadaka, S.R., Iyer, S.K. (2003). A Simple Estimate of the Index of Stability for Symmetric Stable Distributions. **Stochastic Modelling and Applications**, 6(1), 9-19.

- Koç, Y.S.(2007). ***Robust tahmin ediciler ve özellikleri***, Yüksek Lisans Tezi, Çukurova Üniversitesi Fen Bilimler Enstitüsü, Adana, 95-101.
- Koenker, R. and Hallock, K.F. (2001). Quantile Regression an Introduction. ***The Journal of Economics Perspectives***, 15(4), 143-156.
- Koenker, R. (2002). Quantile regression, In S. Fienberg, J. Kadane (Eds.), ***International Encyclopedia of the Social Sciences***, Statistics Volume.
- Koenker, R. (2005). ***Quantile Regression***. (First Published). New York: Cambridge University Press, 5-10,43.
- Leping, K.O. (2005). Public -Private Sector Wage Differential in Estonia: Evidence from Quantile Regression, ***Faculty of Economics and Business Administration Working Paper*** Tartu: Tartu Universty Press, Order No:431, 12.
- Levin J. (2002). For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement. In B. Fitzenberger, R. Koenker and J.A.F. Machado (Eds.), ***Economic Applications of Quantile Regression***, Heidelberg. Physica-Verlag, pp.221-246.
- Liu, H., Shah, S., Jiang, W. (2004). On-line Outlier Detection and Data Cleaning. ***Computer and Chemical Engineering***, 28, 1635-1647.
- Lin, S. W. (2006). A Comparative Study of Robust Estimates in Regression. ***Proceedings of the 11th Annual Conference of Asia Pacific Decision Sciences Institute***, pp.569-572.
- Maronna, A.R., Martin, D.R., Yohai, J.V. (2006). ***Robust Statistics Theory and Methods***. (First Edution). England: John Wiley and Sons Ltd, xv,xvi,xvii,1,24,29-30,124-125,130,131.
- Maru, P.A. and Modi, C.K. (2012). Selecting The Most Favorable Robust Estimator for Denoising of Gray Scale Image with Impulse Noise. ***Quest International Multidisciplinary Research Journal***, 1(1), pp.77-85.
- McCulloch, J.H. (1986). Simple Consistent Estimators of Stable Distribution Parameters. ***Communications in Statistics: Simulation and Computation***, 15, 1109-1136.
- McCulloch, J.H. (1998). Linear Regression with Stable Disturbances. In R.J. Adler, R.E. Feldman, M.S. Taqqu (Eds.). ***A Practical Guide to Heavy Tails; Statistical Techniques and Applications***, Boston, Birkhauser, 359-376.
- Mikosch, T. and Vries, C.G. (2013). Heavy Tails of OLS. ***Journal of Econometrics***, 172(2), 205-221.
- Momeni, M., Nayeri, M.D., Ghayoumi, A.F., Ghorbani H. (2010). Robust Regression and its Application in Financial Data Analysis, ***World Academy of Science, Engineering and Technology***, 71, 521-526.

- Montgomery, D.C., Peck, E.A. and Vining G.G. (2001). **Introduction to Linear Regression Analysis**. (Third Edition). New York: John Wiley and Sons, 3, 13, 14, 15, 401.
- Nevitt, J. and Tam, H. (1998). A Comparison of Robust and Nonparametric Estimators Under the Simple Linear Regression Model. **Multiple Linear Regression Viewpoints**, 25, 54-69.
- Neykov, N.M., Cizek, P., Filzmoser P., Neytchev, P.N. (2012). The Least Trimmed Quantile Regression. **Computational Statistic & Data Analysis**, 56(6) 1757-1770.
- Nolan, J.P. (1997). Numerical Calculation of Stable Densities and Distribution Functions. **Communications in Statistics-Stochastic Models**, 13, 759-774.
- İnternet: Nolan, J.P. Stable Distributions Models for Heavy Tailed Data. URL: <http://www.webcitation.org/query?url=http%3A%2F%2Facademic2.american.edu%2F%7Ejpnolan%2Fstable%2Fchap1.pdf&date=2014-12-13>, Son Erişim Tarihi: 13.12.2014.
- İnternet: Nolan, J.P. Modeling financial data with stable distributions. URL: <http://www.webcitation.org/query?url=http%3A%2F%2Facademic2.american.edu%2F%7Ejpnolan%2Fstable%2FStableFinance23Mar2005.pdf&date=2015-01-03>, Son Erişim Tarihi: 03.01.2015
- Nolan, J.P., Ojeda Revah, D. (2013). Linear and Nonlinear Regression with Stable Errors. **Journal of Econometrics**, 172, 186-194.
- Officer, R.R. (1972). The Distribution of Stock Returns. **Journal of American Statistical Association**, 67(340), 807-812.
- Özyurt D.B., Pike, R.W. (2004). Theory and Practice of Simultaneous Data Reconciliation and Gross Error Detection for Chemical Processes. **Computers and Chemical Engineering**, 28, 381-402.
- Press, S.J. (2005). **Applied Multivariate Analysis Using Bayesian and Frequentist Methods of Inference**. (Second Edition). New York: Dover Publications Inc., 150,151.
- Radhakrishna, C. R., Toutenberg, H. (1999). **Linear Models: Least Squares and Alternatives**. (Second Edition). New York: Springer-Verlag, 273.
- Rousseeuw, P.J. (1984). Least Median of Squares Regression, **Journal of the American Statistical Association**, 79(388), 871-880.
- Rousseeuw, P.J. and Leroy, A.M. (1987). **Robust Regression and Outlier Detection**. (First Edition). New York: John Wiley and Sons Inc., 9,10,14,15,136.
- Rousseeuw, P.J. and Van Driessen, K., (2006). Computing LTS Regression for Large Data Sets, **Data Mining and Knowledge Discovery**, (12), 29-45.

- Ryan, T.P. (2009). **Modern Regression Methods**. (Second Edition). New York: John Wiley and Sons Inc., 15,16,421,427-430,435.
- Saçaklı, İ. (2005). **Kantil Regresyon ve Alternatif Regresyon Modelleri ile Karşılaştırılması**, Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 93.
- Samorodnitsky, G. and Taqqu M.S. (1994). **Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance**, (First Edition), New York: Chapman and Hall, 2,5.
- Schumacker, R.E., Monahan, M.P., and Mount R.E. (2002). A Comparison of OLS to LTS and MM robust regression in S-PLUS. **Southwest Educational Research Association 25th Annual Meeting**, 3.
- Sınıksaran, E. ve Aktükün, A. (2005). Rastlantısal Şeritler ile En Küçük Medyan Kareler Doğrusunun Bulunması. **İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi**, 1, 10-20.
- Staudte, R.G. and Sheather, S.J. (1990). **Robust Estimation and Testing**, Canada: John Wiley and Sons, pp. 44.
- Stigler, M.S. (2010). The Changing History of Robustness. **American Statistical Association**, 64(4), 277-281.
- Street, J.O., Carroll, R.J. and Ruppert, D. (1988) A note on computing robust regression estimates via iteratively reweighted least squares. **The American Statistician**. 42(2), pp.152-154.
- Sufahani S.F. and Ahmad, A. (2012). A Comparison Between Normal and Non-Normal Data in Bootstrap. **Applied Mathematical Sciences**, 6(92), 4547-4560.
- Sundberg, R. (2002). Shrinkage Regression. In A.H. El- Shaarawi and W.W. Piegorisch (Eds.). **Encyclopedia of Environmetrics**, Chichester, John Wiley and Sons Ltd, pp. 1994-1998.
- Tharmarathan, K. and Claeskens, G. (2013). A comparison of Robust Versions of the AIC based on M, S and MM Estimators. **Statistics: A Journal of Theoretical and Applied Statistics**, 1(47), 216-235.
- Tyler, D.E. (1999). S Estimators, In S. Kotz, C. B. Read, D. L. Banks (Eds.), **Encyclopedia of Statistical Sciences**, York: John Wiley and Sons, Inc., 659-662.
- Türkay, H. ve Öztürk, L. (2005). Aşırı değer içeren veri kümelerinde hata terimlerinin binom dağılımına uyduğu durumda EKK ve robust LTS regresyon tahmin edicilerinin simülasyon çalışması ile karşılaştırılması. **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**. 10(1), 263-279.

- Üstünel, İ.E. (2000). *Durağan Portföy Analizi ve İMKB Verilerine Uygulanması*. (İlk Baskı), Ankara: İMKB Yayınları, 18-28.
- Weisberg, S. (2005). *Applied Linear Regression*. (Third Edition), New York: John Wiley and Sons, Inc., 21.
- Verardi, V. and Croux, C. (2009). Robust Regression in Stata. *The Stata Journal*, 9 (3), 439-453.
- İnternet: Yaffee, R.A. Robust Regression Analysis:Some Popular Statistical Package Options.URL:http://www.webcitation.org/query?url=http%3A%2F%2Fweb.ipac.caltech.edu%2Fstaff%2Ffmasci%2Fhome%2Fstatistics_refs%2FRobustRegAnalysis.pdf&date=2014-12-13 , Son Erişim Tarihi:13.12.2014
- Yohai, V.J. (1987) High Breakdown- Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistics*. 15(2), pp.642-656.
- Zhang, Z. (1995). Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting; Inria Sophia Antipolis 2676. *France*, 1-44.
- Zolotarev, V.M. (1986). *One-Dimensional Stable Distributions*. U.S.A, American Mathematical Society, 12-28.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : KORALTAN, Enise
 Uyruğu : T.C.
 Doğum tarihi ve yeri : 21.08.1987, Sivas
 Medeni hali : Bekâr
 Telefon : 0 (537) 4595776
 e-mail : koraltan.enise@gmail.com



Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Yüksek lisans	Gazi Üniversitesi / İstatistik Bölümü	2015
Lisans	19 Mayıs Üniversitesi/ İstatistik Bölümü	2010
Lise	Sivas Lisesi	2004

İş Deneyimi

Yıl	Yer	Görev
2011-Halen	Ziraat Bankası	Servis Yetkilisi

Yabancı Dil

İngilizce

Yayınlar

-

Hobiler

Tarih, Edebiyat, El Sanatları

DİZİN

A

Andrews'ın Yaklaşımı · 33
Asimptotik Etkinlik · 24
aykırı değer · 1, 7, 12, 15, 16, 17,
24, 37, 65

B

bağımlı değişken · 65
bağımsız değişken · 10, 16, 65, 66
Bootstrap · 8, 46, 47, 48, 50, 70,
73, 74, 75, 76, 77, 78, 79, 80,
81, 82, 83, 84, 85, 86, 87, 88,
89, 90, 91, 92, 102, 103, 108

C

Cauchy dağılımı · 52, 56, 59

Ç

çarpıklık · 100

E

En Küçük Budanmış Kareler · 3,
7, 38
En Küçük Kareler · 1, 2, 3, 7, 12,
13, 24, 25, 34, 43, 65, 104
En Küçük Medyan Kareler · 3, 8,
104, 107
En Küçük Mutlak Sapmalar · 3, 8,
26

F

Fair'ın Yaklaşımı · 33

G

Gauss Markov Teoremi · 12
German-McClure Yaklaşımı · 34

H

Hampel'in Yaklaşımı · 32
Huber M tahmini · 4, 30, 31

K

Kaldıraç noktaları · 16
kalın kuyruk · 65
Kantil Regresyon · 8, 42, 43, 102,
107
kararlı dağılımlar · 1, 8, 52, 53,
55, 56, 66, 101
kararlı regresyon · 8
kırılma noktası · 4, 5, 22, 23, 24,
36, 37, 38, 40, 42

L

Levy dağılımı · 56, 59

M

M tipi tahmin ediciler · 7
Merkezi Limit Teoremi · 12, 20,
52, 64, 65
MM Tahmini · 8, 37

N

Normal dağılım · 6, 20, 55

O

otokorelasyon · 11

R

regresyon · 2, 3, 4, 5, 6, 7, 10, 11,
13, 14, 16, 17, 18, 20, 21, 22,
23, 24, 25, 27, 28, 30, 37, 38,
39, 40, 41, 43, 44, 45, 46, 49,
50, 64, 67, 68, 100, 101, 108
Regresyon analizi · 9, 11, 13

S

S tahmini · 4, 8, 36, 37, 100
Sağlam yöntemler · 1, 2, 28

T

Tukey · 2, 4, 31, 32, 37, 70, 73,
74, 75, 76, 77, 78, 79, 80, 81,
82, 83, 84, 85, 86, 87, 88, 89,
90, 91, 101

V

varsayım · 12, 20



GAZİ GELECEKTİR..