

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

AMPİRİK OLABİLİRLİK YÖNTEMİ İLE
ROBUST REGRESYON ANALİZİ

Şenay ÖZDEMİR

İSTATİSTİK ANABİLİM DALI

ANKARA
2018

Her hakkı saklıdır.

TEZ ONAYI

Şenay ÖZDEMİR tarafından hazırlanan “Ampirik Olabilirlik Yöntemi ile Robust Regresyon Analizi ” adlı tez çalışması 01/03/2018 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı’nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Danışman : Prof. Dr. Olcay ARSLAN
Ankara Üniversitesi İstatistik Anabilim Dalı

Jüri Üyeleri :

Başkan : Prof. Dr. Cemal ATAKAN
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Prof. Dr. Olcay ARSLAN
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Doç. Dr. Meral EBEGİL
Gazi Üniversitesi İstatistik Anabilim Dalı

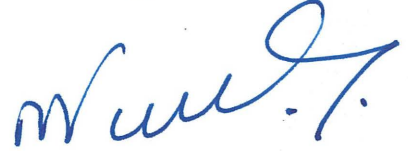
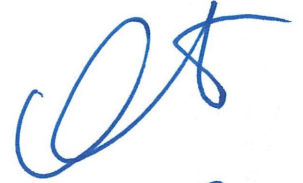
Üye : Doç. Dr. Esin KÖKSAL BABACAN
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Yrd. Doç. Dr. Engin TAŞ
Afyon Kocatepe Üniversitesi İstatistik Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. Atila YETİŞEMİYEN

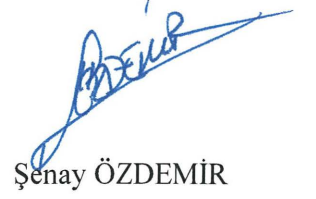
Enstitü Müdürü



ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

01/03/2018



Şenay ÖZDEMİR

ÖZET

Doktora Tezi

AMPIRİK OLABİLİRLİK YÖNTEMİ İLE ROBUST REGRESYON ANALİZİ

Şenay ÖZDEMİR

Ankara Üniversitesi
Fen Bilimleri ENSTİTÜSÜ
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Olcay ARSLAN

Doğrusal regresyon modellerin parametre tahmini için En Küçük Kareler (EKK) ve En Çok Olabilirlik (EÇO) yöntemleri sıklıkla kullanılmaktadır. EKK ve EÇO tahmin edicilerinin veri setindeki sapan gözlem, çarpıklık ve/veya kalın kuyrukluluk durumlarından etkilendiği bilinmekte ve bu durumlarda robust tahmin edicilere başvurulmaktadır. Bu tahmin yöntemleri, hatalar üzerinde bazı varsayımlar sağlandığında iyi sonuçlar vermektedir. Bu varsayımların sağlanmaması durumunda ise parametrik olmayan yöntemlerden yararlanılmaktadır. Parametrik olmayan yöntemlerden bir tanesi olan ampirik olabilirlik yöntemi, gözlemlere bilinmeyen olasılıksal ağırlıklar verir ve bu ağırlıkların bir fonksiyonu olarak tanımlanan ampirik olabilirlik fonksiyonunun bazı kısıtlar altında en büyük olmasını sağlayan parametre tahminlerini bulur. Bu süreçte tahminler açıkça yazılamadığı için nümerik yöntemlere başvurulur. Sözü geçen kısıtlardan bazılarının EKK tahmin yöntemindeki normal denklemlere benzerliği bu yöntemin, veri setindeki bozukluklarla karşılaştığında kullanılabilirliğine gölge düşürmektedir. Buradan yola çıkarak ampirik olabilirlik yönteminde kullanılan kısıtlar, robust yöntemler yardımıyla yeniden ele alınmış ve robust kısıtlara sahip ampirik olabilirlik tahmin edicisinin çalışabilirliği bazı simülasyon çalışmaları ve gerçek veri setleri yardımıyla gösterilmiştir. Buna ek olarak tahmin edicilerin elde edilişinde kullanılabilecek alternatif bir hesaplama yöntemi de bu çalışmada sunulmuştur.

Mart 2018, 74 sayfa

Anahtar Kelimeler: Ampirik Olabilirlik Yöntemi, Doğrusal regresyon modeli, Robust regresyon, M tahmini, MM tahmini.

ABSTRACT

Ph. D. Thesis

ROBUST REGRESSION ANALYSIS WITH EMPIRICAL LIKELIHOOD

Şenay ÖZDEMİR

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Prof. Dr. Olcay ARSLAN

Ordinal Least Square (OLS) and Maximum Likelihood (ML) estimation methods are frequently used to estimate the parameters of a linear regression model. It is known that OLS and ML estimators are affected by outliers, skewness and \ or heavy tailedness in the data sets. To deal with this problem robust estimation methods have been proposed. These estimation methods perform properly under some assumptions on error terms. If these assumptions are not met, non-parametric methods are used. Empirical likelihood (EL) method, one of these nonparametric methods, gives unknown probabilistic weights to the observations and finds parameter estimates by maximizing the empirical likelihood function defined as a multiplication of these probabilistic weights, under some constraints. Some constraints in EL method are similar to the normal equations in the OLS estimation method. However, it is well known that the OLS method has poor performance when there are some outliers in the data. In this study, the constraints used in the empirical likelihood method are combined with robust methods and the feasibility of the empirical likelihood estimator with robust constraints is shown with some simulation studies and real data examples. Concerning the computation of the EL estimators numerical methods are used because estimates can not be explicitly written. In literature this problem is solved by using Lagrange multipliers and duality approach. In this thesis, an alternative computation method is proposed to obtain the EL estimators.

March 2018, 74 pages

Key Words: Empirical Likelihood, Linear Regression Model, Robust Regression, M estimation, MM estimation

TEŞEKKÜR

Bu tez çalışmasına engin bilgi birikimi ile yön veren, akademik anlamda olduğu kadar hayata dair kıymetli öneri ve yardımlarını benden esirgemeyen değerli danışman hocam sayın Prof. Dr. Olcay ARSLAN'a (Ankara Üniversitesi İstatistik Bölümü) teşekkürü bir borç bilirim.

Danışman hocamla tanışmama vesile olan kıymetli hocam Prof. Dr. Birdal ŞENOĞLU'na (Ankara Üniversitesi İstatistik Bölümü), Tez İzleme Komitelerinde değerli görüşleriyle bu tez çalışmasına katkıda bulunan hocalarım Prof.Dr. Cemal ATAKAN (Ankara Üniversitesi İstatistik Bölümü) ve Doç Dr. Meral EBEBİL'e (Gazi Üniversitesi İstatistik Bölümü), bu tez çalışmasını özenle inceleyen Doç. Dr. Esin KÖKSAL BABACAN'a (Ankara Üniversitesi İstatistik Bölümü), bu tez çalışmasının oluşturulması sırasında kıymetli bilgilerini benimle paylaşan, yardım eden ve bu tez çalışmasını dikkatle inceleyen Yrd. Doç.Dr. Engin TAŞ'a (Afyon Kocatepe Üniversitesi İstatistik Bölümü), birer hoca ve arkadaş olarak verdikleri destek ve yardımları için Yrd. Doç.Dr. Ayça Hatice TÜRKAN (Afyon Kocatepe Üniversitesi İstatistik Bölümü) ve Doç.Dr. Sinan SARAÇLI'ya (Afyon Kocatepe Üniversitesi İstatistik Bölümü) teşekkürlerimi sunarım. Doktora çalışmalarım sırasında tanıştığım ve birlikte çalışmaktan keyif aldığım arkadaşlarım Araş. Gör. Yeşim GÜNEY (Ankara Üniversitesi İstatistik Bölümü) ve Araş. Gör. Yetkin TUAÇ'a (Ankara Üniversitesi İstatistik Bölümü) misafirperver ve yardımseverlikleri için teşekkür ederim.

Her daim olduğu gibi bu süreçte de yanımda olan, maddi manevi hiçbir desteğini esirgemeyen ve varlıklarıyla beni onurlandıran çok sevdiğim aileme en içten duygularla teşekkürlerimi sunarım.

Şenay ÖZDEMİR

Ankara, Mart 2018

İÇİNDEKİLER

TEZ ONAY SAYFASI

ETİK.....	i
ÖZET.....	ii
ABSTRACT.....	iii
TEŞEKKÜR	iv
KISALTMALAR DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	ix
1. GİRİŞ	1
2. AMPİRİK OLABİLİRLİK YÖNTEMİ.....	9
2.1 Ampirik Olabilirlik Yöntemi ile Konum Modelinde Parametre Tahmini	10
2.2 Ampirik Olabilirlik Yöntemi ile Konum-Ölçek Modelinde Parametre Tahmini.....	14
2.3 Ampirik Olabilirlik Yöntemi ile Doğrusal Regresyon Modelinde Parametre Tahmini.....	16
2.4 Ampirik Olabilirlik Yöntemi ile Doğrusal Regresyon Parametreleri ve Varyansın Eşanlı Tahmini	20
3. AMPİRİK OLABİLİRLİK YÖNTEMİ İLE REGRESYON PARAMETRELERİNİN TAHMİNİNDE ALTERNATİF BİR NÜMERİK HESAPLAMA ALGORİTMASI	23
3.1 Lagrange Çarpanları için Newton Algoritmasının Ampirik Olabilirlik Yöntemine Uyarlanması.....	25
3.2 Simülasyon Çalışması	28
4. REGRESYON PARAMETRELERİNİN ROBUST TAHMİNİ	34
4.1 Regresyon Parametrelerinin M Tahmini.....	34
4.2 Regresyon Parametrelerinin MM Tahmini	39
5. REGRESYON PARAMETRELERİNİN ROBUST KISITLARLA AMPİRİK OLABİLİRLİK TAHMİNİ.....	42
5.1 M Tahmininden Faydalanılarak Oluşturulan Kısıtlarla Regresyon Parametrelerinin Ampirik Olabilirlik Tahmini.....	42
5.2 MM Tahmininden Faydalanılarak Oluşturulan Kısıtlarla Regresyon Parametrelerinin Ampirik Olabilirlik Tahmini.....	46

5.3 Simülasyon Çalışmaları	48
6. GERÇEK DÜNYADAN ÖRNEKLER	58
6.1 Stackloss Verisi.....	58
6.2 Belçika'dan Uluslararası Telefon Aramaları Örneği	60
6.3 Yıldız Verisi	63
6.4 Coleman Verisi	65
7. TARTIŞMA VE SONUÇ	68
KAYNAKLAR	70
ÖZGEÇMİŞ.....	73



KISALTMALAR DİZİNİ

EÇO	En Çok Olabilirlik
EÇON	Hataların normal dağıldığı varsayımında EÇO
EÇOT	Hatalar ϑ serbestlik dereceli t dağıldığında EÇO
EKK	En Küçük Kareler
EL	Ampirik Olabilirlik (Empirical Likelihood)
EL _d	Dual problemle çözülen EL tahmini
EL _{HUB}	Kısıtlarda Huber fonksiyonundan faydalanılan EL tahmini
EL _{MM}	Kısıtlarda MM tahmin yönteminden faydalanılan EL tahmini
EL _t	Kısıtlarda t dağılımından faydalanılan EL tahmini
EL _{TUK}	Kısıtlarda Tukey fonksiyonundan faydalanılan EL tahmini
EL _p	Primal problemle çözülen EL tahmini
GE	Görelî Etkinlik
HKO	Hata Kareler Ortalaması

ŞEKİLLER DİZİNİ

Şekil 4.1 Bazı ρ fonksiyonları.....	38
Şekil 4.2 Bazı ψ fonksiyonları.....	38
Şekil 6.1 Stackloss verisi için tahmin edicilere ait hatalar.....	60
Şekil 6.2 Belçika'dan yapılan uluslararası telefon aramaları (milyon) için serpm diyagramı ve tahmin edicilere ilişkin regresyon doğruları	61
Şekil 6.3 Belçika'dan yapılan uluslararası telefon aramaları (milyon) için hata, p ve w değerleri.....	62
Şekil 6.4 Yıldız verisi için serpm diyagramı ve tahmin edicilere ilişkin regresyon doğruları.....	64
Şekil 6.5 Coleman verisi için tahmin edilen hata, p ve w değerleri.....	67

ÇİZELGELER DİZİNİ

Çizelge 3.1 Hatalar standart normal dağılıma sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	29
Çizelge 3.2 Hatalar 3 serbestlik dereceli t dağılımına sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	30
Çizelge 3.3 Hatalar basıklığı arttırılmış 3 serbestlik dereceli t dağılımına sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	30
Çizelge 3.4 Gözlem değerleri standart normal dağılıma sahip ve iki sapan gözlem olması durumunda tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri	31
Çizelge 3.5 Hatalar standart normal dağılıma sahip ve iki sapan gözlem olması durumunda tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	31
Çizelge 3.6 Hatalar standart normal dağılıma sahipken regresyon parametrelerinin tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	32
Çizelge 3.7 Hatalar standart normal dağılıma sahipken varyansa ilişkin tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	32
Çizelge 3.8 Hatalar 3 serbestlik dereceli t dağılımına sahipken regresyon parametrelerinin tahmin edicilerinin EKK tahmin edicisine göre görelî etkinlikleri	32
Çizelge 3.9 Hatalar 3 serbestlik dereceli t sahipken varyansa ilişkin tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri.....	33
Çizelge 5.1 $\varepsilon_i \sim N(0,1)$ iken tahmin edicilerin HKO değerleri.....	50
Çizelge 5.2 $\varepsilon_i \sim N(0,1)$ iken tahmin edicilerin GE değerleri	50
Çizelge 5.3 $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken tahmin edicilerin HKO değerleri.....	51
Çizelge 5.4 $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken tahmin edicilerin GE değerleri.....	51
Çizelge 5.5 $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken bir örneğe ilişkin ε_i , p_i ve w_i değerleri.....	52
Çizelge 5.6 $X \sim (0,90)N(0,1) + (0,10)N(10,1)$ iken tahmin edicilerin HKO değerleri	53
Çizelge 5.7 $X \sim (0,90)N(0,1) + (0,10)N(10,1)$ iken tahmin edicilerin GE değerleri.....	54
Çizelge 5.8 $X \sim (0,90)N(0,1) + (0,10)N(10,1)$ iken bir örneğe ilişkin X , p_i ve w_i değerleri.....	55
Çizelge 5.9 Tahmin değerleri.....	56
Çizelge 5.10 $X \sim (0,95)N(0,1) + (0,05)N(10,1)$ ve $\varepsilon_i \sim (0,95)N(0,1) + (0,05)N(10,1)$ iken tahmin edicilerin HKO değerleri	56
Çizelge 5.11 $X \sim (0,95)N(0,1) + (0,05)N(10,1)$ ve $\varepsilon_i \sim (0,95)N(0,1) + (0,05)N(10,1)$ iken tahmin edicilerin GE değerleri	57
Çizelge 6.1 Stackloss verisi	58
Çizelge 6.2 Model parametrelerinin tahminleri	59
Çizelge 6.3 Tahminlere ilişkin HKT değerleri.....	59
Çizelge 6.4 Belçika'dan yapılan uluslararası telefon aramaları (milyon).....	60
Çizelge 6.6 Yıldız verisi (Rousseeuw ve Leory, 1987)	63
Çizelge 6.7 Yıldız verisi için parametre tahminleri	64
Çizelge 6.8 Coleman verisi (Rousseeuw ve Leory, 1987)	65
Çizelge 6.9 Coleman verisi için parametre tahminleri.....	66

1. GİRİŞ

Regresyon analizi, değişkenler arasındaki ilişkilerin yapısını matematiksel modeller yardımıyla inceleyen istatistiksel bir tekniktir. Regresyon analizi mühendislik, fizik, biyoloji, ekonomi ve sosyal bilimler gibi birçok alanda veri yapısını tanımlamak, tahmin ve kestirim yapmak gibi birçok amaç için kullanılmaktadır.

Doğrusal bir regresyon modeli,

$$y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

olarak ifade edilebilir. Burada k parametre sayısı ve n gözlem sayısı olmak üzere, y_i i-inci yanıt değişkenine ilişkin değeri, $\mathbf{X}_i \in R^k$ i-inci gözlemin açıklayıcı değişkenler bakımından aldığı değerleri gösteren vektörü, $\boldsymbol{\beta} \in R^k$ bilinmeyen regresyon katsayılarına ilişkin $k \times 1$ boyutlu vektörü ve ε_i rasgele hataları göstermektedir. Regresyon modelinde birbirinden ve açıklayıcı değişkenlerden bağımsız olan rasgele hataların beklenen değerinin sıfır ve varyansının σ^2 olduğu varsayılır.

Regresyon parametre vektörü $\boldsymbol{\beta}$ 'nin tahmininde yaygın olarak kullanılan yöntemlerden bir tanesi en küçük kareler (EKK) yöntemidir ve

$$\frac{1}{n} \sum_{i=1}^n (y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2$$

ifadesinin $\boldsymbol{\beta}$ parametre vektörüne göre en küçüklenmesiyle elde edilir. Yani

$$\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i (y_i - \mathbf{X}_i^T \boldsymbol{\beta}) = 0$$

normal denklemlerinin çözümü regresyon parametrelerinin EKK tahminlerini verir. Normal denklemlerden görüldüğü üzere EKK tahmin yöntemi tüm gözlemlere eşit olasılık atamaktadır. Yani tüm gözlemlerin katkıları eşit ve $1/n$ 'dir. Genel olarak EKK

tahmin edicileri bulunurken $1/n$ çarpanı tahmin sonucunu etkilemediği için ihmal edilmektedir. Öte yandan en çok olabilirlik (EÇO) tahmin yöntemi kullanıldığında log olabilirlik fonksiyonunun düzenlenmesi ile bu çarpan kolayca görülebilir. Bilinmeyen regresyon parametresi ve rasgele hatalara ilişkin varyansın EKK tahmin edicileri sırasıyla

$$\hat{\beta} = (\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T)^{-1} \sum_{i=1}^n \mathbf{x}_i y_i \quad (1.2)$$

ve

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n (y_i - \mathbf{x}_i^T \hat{\beta})^2$$

olarak tanımlanır.

Hatalar beklenen değeri sıfır ve sabit varyanslı normal dağılıma sahip ise EKK tahmin edicileri yansız ve en küçük varyanslıdır. Hataların dağılım bilgisi mevcut ise EÇO tahmin yöntemi regresyon katsayılarının tahmininde alternatif bir yöntem olarak kullanılabilir. Örneğin hatalar normal dağılıyorsa regresyon parametrelerinin EÇO tahmini

$$L(y_i, \mathbf{x}_i, \beta, \sigma^2) = \prod_{i=1}^n (2\pi\sigma^2)^{-1/2} \exp \left[-\frac{1}{2\sigma^2} (y_i - \mathbf{x}_i^T \beta)^2 \right]$$

olabilirlik fonksiyonu en büyüleyecek şekilde bulunur. İşlem kolaylığı sağlamak adına L fonksiyonu ile aynı değerlerde maksimum noktasına ulaşan $\ln(L)$, yani L fonksiyonunun doğal logaritması kullanılır. $\ln(L)$ fonksiyonunun ilgili parametrelere göre türevi alınıp sıfıra eşitlenmesiyle β ve σ^2 parametrelerinin EÇO tahmin edicileri

$$\hat{\beta} = \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T \right)^{-1} \sum_{i=1}^n \mathbf{x}_i y_i$$

ve

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \mathbf{x}_i^T \hat{\beta})^2$$

olarak elde edilir.

Hataların kalın kuyruklu olması durumunda farklı dağılımlar kullanılabilir. Örneğin hataların ϑ serbestlik dereceli t dağılımına sahip olduğu varsayılırsa hatalara ilişkin olasılık yoğunluk fonksiyonu

$$f(y_i, \mathbf{X}_i, \boldsymbol{\beta}, \sigma^2) = \frac{\Gamma[(\vartheta + 1)/2]}{\Gamma((\vartheta/2))} \frac{1}{\sqrt{\pi\vartheta\sigma^2}} \left[1 + \frac{(y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2}{\vartheta\sigma^2} \right]^{-(\vartheta+1)/2} \quad (1.3)$$

olduğu için parametrelerin en çok olabilirlik tahminleri

$$L(y_i, \mathbf{X}_i, \boldsymbol{\beta}, \sigma^2) = \prod_{i=1}^n \frac{\Gamma[(\vartheta + 1)/2]}{\Gamma((\vartheta/2))} \frac{1}{\sqrt{\pi\vartheta\sigma^2}} \left[1 + \frac{(y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2}{\vartheta\sigma^2} \right]^{-(\vartheta+1)/2}$$

olabilirlik fonksiyonun en büyüklenmesiyle bulunur. $L(y_i, \mathbf{X}_i, \boldsymbol{\beta}, \sigma^2)$ fonksiyonuyla aynı noktada en büyük değerini alan logaritmasının $\boldsymbol{\beta}$ parametresine göre türevi alınarak elde edilen

$$\frac{\partial L(y_i, \mathbf{X}_i, \boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n (\vartheta + 1) \frac{\mathbf{X}_i (y_i - \mathbf{X}_i^T \hat{\boldsymbol{\beta}})}{\vartheta\sigma^2 + (y_i - \mathbf{X}_i^T \hat{\boldsymbol{\beta}})^2} = 0$$

eşitliğinin çözümü $\boldsymbol{\beta}$ parametresinin EÇO tahmin edicisini verir. Burada

$$w_i = \frac{(\vartheta + 1)}{\vartheta\sigma^2 + (y_i - \mathbf{X}_i^T \hat{\boldsymbol{\beta}})^2}$$

alınarak $\hat{\boldsymbol{\beta}}$

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^n w_i \mathbf{X}_i \mathbf{X}_i^T \right)^{-1} \sum_{i=1}^n w_i \mathbf{X}_i y_i$$

olarak bulunur. Benzer şekilde hatalara ilişkin varyansın EÇO tahmin edicisi

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n w_i (y_i - \mathbf{X}_i^T \hat{\boldsymbol{\beta}})^2$$

olarak elde edilir.

Dağılım bilgisi kullanıldığı için EÇO tahmin yöntemi EKK'den daha iyi istatistiksel özelliklere sahiptir. EÇO tahmin edicileri yansız (veya asimptotik yansız), en küçük varyanslı, tutarlı ve yeterli istatistiklerdir. Görüldüğü üzere EÇO yöntemi ile hataların varyansı için yanlı bir tahmin elde edilir. Ancak varyansın EÇO tahmin edicisinin asimptotik olarak yansız olduğu bilinmektedir.

EÇO ve EKK tahmin edicilerinin analitik olarak hesaplanma kolaylığı beraberinde sağlanması gereken, hataların sıfır ortalamalı, sabit varyanslı olması veya bilinen bir dağılıma sahip olması gibi bazı varsayımları getirmektedir. Ancak birçok alanda söz konusu varsayımları sağlamayan veri setleri ile sıklıkla karşılaşmaktadır.

Açıklayıcı değişkenle yanıt değişkeni arasındaki ilişkiyi önceden belirlenmiş bir kalıpta inceleyen parametrik yöntemler yerine, veriyle uyumlu yapılar sunan yöntemlere ihtiyaç duyulması, parametrik olmayan yöntemlerin ortaya çıkmasına neden olmuştur. Regresyon parametrelerinin tahmininde düzleştirici metotlar (smoothing methods), çekirdek tahmini (kernel estimation), jackknife ve bootstrap gibi birçok parametrik olmayan yöntem önerilmiştir.

Gün geçtikçe gelişen istatistiksel yazılımlar ve öne sürülen nümerik hesaplama yöntemleri araştırmacılara parametrik yöntemlerden daha esnek, veri setinin yapısına uygun çıkarım yöntemleri sunmaktadır ve parametrik olmayan yöntemlerin kullanımı giderek yaygınlaşmaktadır. Bu yöntemlerden bir tanesi de ampirik olabilirlik yöntemidir. Bu yöntem, veri setinin bilinen bir dağılım ailesinden geldiği varsayımı olmadan olabilirlik yönteminin kullanılmasını mümkün kılmaktadır.

Tahmin problemlerinin çözümü için bir çok çalışmada ampirik olabilirlik yöntemi kullanılmıştır. Bunlardan bazıları şu şekilde özetlenebilir: Owen (1990), ampirik olabilirlik oranı yardımıyla güven bölgesi oluşturulması üzerinde çalışmıştır. Yine Owen (1991), ampirik olabilirlik yöntemi ile doğrusal regresyon parametrelerinin tahmin problemini ele almıştır. Kolaczyk (1994) ise, ampirik olabilirlik yöntemini genelleştirilmiş doğrusal regresyon modeline uyarlamıştır. Qin ve Lawless (1994),

moment kısıtları ile parametre tahminine ampirik olabilirlik yaklaşımında bulunmuşlardır. Chen (1993, 1994, 1996), ampirik olabilirlik yöntemiyle güven bölgesi oluşturmada çeşitli çalışmalar yapmıştır. Kitamura (2001), moment kısıtlarını test etmede ampirik olabilirlik yönteminin bazı asimptotik özellikleri üzerine bir çalışma yapmıştır. Imbens (2002) çalışmasında, genelleştirilmiş momentler yöntemi ile ampirik olabilirlik yöntemini ele almıştır. Chen ve Cui (2003), yeni kısıtlar ekleyerek genelleştirilmiş doğrusal regresyon modeli için ampirik yöntemlerle daha etkin tahmin ediciler bulmayı amaçlamıştır. Newey ve Smith (2004), küçük örneklemeler için genelleştirilmiş moment yöntemi ile genelleştirilmiş ampirik olabilirlik yöntemini karşılaştırmıştır. Kitamura (2006), ampirik olabilirlik yöntemi ile ilişkilendirilen genelleştirilmiş momentler yöntemi ve bazı tahmin yöntemleri arasındaki bağlantılardan söz etmiştir ve ampirik olabilirlik yönteminin diğerlerinden daha avantajlı yönleri olduğunu belirtmiştir. Chen ve Keilegom (2009), ampirik olabilirlik yönteminin doğrusal regresyona uyarlanması hakkında özet bilgi sağlayan bir çalışma yapmıştır. Bondell ve Stefanski (2013) yaptıkları çalışmada, genelleştirilmiş ampirik olabilirlik yöntemi ile regresyon parametreleri için üstel olarak tanımladıkları olasılıksal ağırlıklar yardımıyla robust bir tahmin edici önermişlerdir. Bu tahmin edicinin diğer robust tahmin edicilerden daha etkin olduğunu belirtmişlerdir.

Ampirik olabilirlik yöntemleri gözlemlere dayalı olarak elde ettiği olasılıksal ağırlıklardan oluşan bir olabilirlik fonksiyonunu, bazı kısıtlar altında en büyük yapan değerleri belirleyerek parametre tahmini yapar. Owen (1988, 1990, 1991), ampirik olabilirlik yöntemi ile parametre tahmini problemini çözmek için Lagrange çarpanları yöntemini kullanmıştır. Çözüm sırasında olasılıksal ağırlıkları, bilinmeyen parametrelerin ve Lagrange çarpanlarının bir fonksiyonu olarak ifade etmiş ve problemi bilinmeyen parametrelerin ve Lagrange çarpanlarının bulunmasına indirgemıştır. Üzerinde düzenlemeler yaptığı Lagrange fonksiyonunun Taylor açılımından faydalanmıştır. Ancak parametre tahminleri analitik olarak açıkça yazılamadığı için nümerik yöntemlerden faydalanılır.

Son zamanlarda sıklıkla başvuru alan parametre tahmin yöntemleri arasında robust yöntemler de yer almaktadır. Robust tahmin yöntemleri ile modeldeki varsayımların

sağlanmaması durumunda da güvenilir ve etkin tahmin edicilerin bulunması amaçlanmaktadır. Günümüzde kullanılan robust istatistik yöntemlerinin temelini Huber (1964)'in yaptığı çalışma oluşturmaktadır. Huber (1964), hata kareler yerine başka bir fonksiyonun en küçüklenmesi ile EKK yöntemine göre daha robust bir tahmin edicinin elde edilebilirliği üzerinde çalışmıştır.

Bir tahmin edicinin robust olması etki fonksiyonu ve kırılma noktası yardımıyla ölçülür. Bir tahmin edicinin etki fonksiyonu (IF: Influence function), söz konusu tahmin edicinin hassas eğrisinin (SC: sensitivity curve) asimptotik yorumu olarak tanımlanabilir. Etki fonksiyonu örneklem büyüklüğü sonsuz ve örneklem özdeş aykırı değerlerin ε kadar küçük bir kısmını içerdiğinde tahmin edicinin davranışını belirlemede kullanılan bir yaklaşımdır. δ_{x_0} , $P(X = x_0) = 1$ olan dağılım fonksiyonu (yani aykırı değere ait dağılım fonksiyonunu) ve “ $\downarrow$ ” simgesi “sağdan limit”i göstermek üzere bir θ parametresinin $\hat{\theta}_\infty$ tahmin edicisinin etki fonksiyonu

$$\begin{aligned} IF_{\hat{\theta}}(x_0, F) &= \lim_{\varepsilon \downarrow 0} \frac{\hat{\theta}_\infty((1 - \varepsilon)F + \varepsilon\delta_{x_0}) - \hat{\theta}_\infty(F)}{\varepsilon} \\ &= \left. \frac{\partial}{\partial \varepsilon} \hat{\theta}_\infty((1 - \varepsilon)F + \varepsilon\delta_{x_0}) \right|_{\varepsilon \downarrow 0} \end{aligned}$$

olarak tanımlanır (Maronna vd. 2006).

Bir regresyon probleminde, bir tahmin edicinin kırılma noktası $\varepsilon^* = \frac{m^*}{n}$ olarak tanımlanır. Burada $m^* = \max\{m \geq 0: \hat{\beta}(X, y_m), \forall y_m \in Y_m\}$ ve Y_m , en az $n - m$ elemanı y ile ortak olan n -vektörler kümesidir. Veri setinde bir nokta değiştirildiğinde EKK tahmin edicisinin değerinin de değiştiği açıkça görülmektedir. Bu nedenle EKK tahmin edicilerinin kırılma noktası $\varepsilon^* = \frac{1}{n}$ 'dir. Dolayısıyla n örneklem büyüklüğü arttıkça, ε^* ifadesinin 0'a yaklaştığı görülmektedir. Diğer bir ifadeyle EKK tahmin edicilerinin kırılma noktası sıfırdır (Maronna vd. 2006).

Etki fonksiyonları sınırlı olan tahmin ediciler yerel (local) anlamda robust tahmin ediciler olarak adlandırılır. Huber M tahmin edicileri (mutlak değer ve Huber ρ fonksiyonu) hatalara göre sınırlı etki fonksiyonuna sahip olduğu için y yönündeki sapan gözlemlere karşı dayanıklıdır. Diğer robustluk ölçüsü olan kırılma noktasına göre, veri setinde y yönünde sapan gözlem olması durumunda EKK tahmin edicilerinin aksine LAD (Least Absolute Deviation: En Küçük Mutlak Sapma) ve Huber M tahmin edicilerinin sapan gözlemlere karşı dayanıklı olduğu görülmektedir. Fakat M tahmin edicileri x yönündeki sapan gözlemlere karşı dayanıklı değildir. x yönündeki sapmalara karşı dayanıklı olan tahmin ediciler Genelleştirilmiş M (GM) tahmin edicileridir. Bu tahmin ediciler hem hatalar hem de x yönünde sınırlı etki fonksiyonuna sahiptirler. Fakat x yönünde sapan gözlem olduğunda bu tahmin ediciler kırılma noktası bakımından çok dayanıklı değildir. $k - 1$ bağımsız değişken sayısı olmak üzere kırılma noktası en fazla $1/k$ 'dir.

M ve GM tahmin edicileri sınırlı etki fonksiyonlarına sahip oldukları için yerel anlamda robust tahmin edicilerdir. Robust istatistikteki diğer bir gelişme ise yüksek kırılma noktasına sahip robust regresyon tahmin edicilerinin elde edilmesi olmuştur. Bu doğrultuda EKK'da olduğu gibi hataların kareler toplamı yerine hataların karelerinin medyanını en küçüklemeyi amaçlayan LMS (Least Median of Squares: Karelerin En Küçük Medyanı) tahmin edicileri önerilmiştir (Rousseeuw 1984). LMS tahmin edicileri, regresyonda eş değişim özelliği olan, sınırlı etki fonksiyonuna ve yüksek kırılma noktasına sahip olan tahmin edicilerdir. Ancak bu tahmin ediciler $\sqrt{n}$ - tutarlı değildir ($\sqrt{n}\|\hat{\beta} - \beta\| \xrightarrow{k} \infty$). Rousseeuw ve Yohai (1984), $\sqrt{n}$ - tutarlı ve asimptotik normal olan S tahmin edicilerini önermiştir. S tahmin edicileri yüksek kırılma noktasına sahip olduklarında etkin tahmin ediciler değildir. Bu nedenle yüksek kırılma noktasına ve yüksek etkinliğe sahip MM ve τ tahmin edicileri tanımlanmıştır (Yohai 1987, Yohai ve Zamar 1988). Bunlara ek olarak yüksek kırılma noktasına ve yüksek etkinliğe sahip diğer bir tahmin edici CM (Constrained M: Kısıtlı M) tahmin edicisidir (Mendes ve Tyler 1996).

Veri setlerinde sapan gözlemlerle, kalın kuyruklu ve/veya çarpıklıkla karşılaşıldığında doğru çıkarımlara ulaşabilmek için robust tahmin yöntemleri kullanılmaktadır. Regresyon modeli oluşturmada hatalara ilişkin dağılım bilgisi olmadığında kullanılan

parametrik olmayan yöntemlerden bir tanesinin ampirik olabilirlik yöntemi olduğu daha önce belirtilmişti. Ampirik olabilirlik yönteminde yer alan olasılıksal ağırlıklar sapan gözlemlere karşı dayanıklı olan robust tahmin edicilerin ağırlıklarından yapısal olarak farklıdır. Robust yöntemlerde de hataların genelini uyduğu düşünülen bir dağılım varsayımı yapılmaktadır. Hatalara ilişkin dağılım bilgisi olmadığında ve verilerde verinin çoğunluğundan farklılık gösteren gözlemler olması durumunda bu gözlemlerin etkisini azaltacak bir tahmin ediciye, yani ampirik olabilirlik yöntemi ile robust yöntemlerin birleştirilmesine ihtiyaç duyulmaktadır.

Bu tez çalışmasının iki amacı vardır. Bunlardan ilki ampirik olabilirlik yöntemi ile parametre tahmininde kullanılan sayısal yöntem yerine kullanılabilecek alternatif bir hesaplama yöntemini ve yöntemin çalışabilirliğini göstermek için yapılan simülasyon çalışmalarını sunmaktır. Tezin ikinci ve esas amacı ise ampirik olabilirlik yönteminde kullanılan kısıtları robust hale getirerek, sapan gözlemlerden, kalın kuyruklu olma ve/veya çarpıklıktan daha az etkilenen ve dağılım varsayımı gerektirmeyen tahmin ediciler tanımlamaktır.

Bu tez çalışması 7 bölümden oluşmaktadır. İkinci bölümde konum, konum-ölçek ve doğrusal regresyon modellerinde ampirik olabilirlik yöntemi ile parametre tahmini verilmektedir. Üçüncü bölümde ise ampirik olabilirlik yönteminde yaygın olarak kullanılan hesaplama algoritması yerine önerilen alternatif hesaplama algoritması ve ilgili simülasyon çalışması yer almaktadır. Dördüncü bölümde ampirik olabilirlik yönteminde dayanıklı kısıtlar elde edebilmek için kullanılacak robust yöntemler kısaca özetlenmiştir. Beşinci bölümde ilk olarak y yönünde sapan gözlemlere karşı dayanıklı olan kısıtlarla elde edilen ampirik olabilirlik tahminleri, sonrasında ise x yönünde sapan gözlem olması durumunda da dayanıklı olan MM tahmin edicileri yardımıyla elde edilen kısıtlarla ampirik olabilirlik tahmin edicileri ve bazı simülasyon çalışmaları verilmektedir. Gerçek dünyadan alınan bazı örneklere ilişkin model parametrelerinin robust kısıtlarla ampirik olabilirlik tahminleri bölüm altıdadır. Bu tez çalışmasının son bölümünü tartışma ve sonuç bölümü oluşturmaktadır.

2. AMPİRİK OLABİLİRLİK YÖNTEMİ

Olabilirlik yöntemleri, etkin tahmin ediciler bulmada ve güçlü test fonksiyonları oluşturmada sıklıkla kullanılmaktadır. Ancak klasik olabilirlik yöntemleri veri setinin bilinen bir dağılımdan geldiğini varsaymaktadır. Veri setinin dağılımı bilinen bir dağılım ailesinden gelmiyorsa klasik olabilirlik yöntemi kullanılamaz. Bunun yerine dağılım varsayımlarına ihtiyaç duyulmadan analiz yapılabilen parametrik olmayan yöntemlere başvurulabilir. Ampirik olabilirlik yöntemi de dağılım bilgisi olmadan, örnekleme dayalı çıkarımlarda bulunan istatistiksel bir yöntemdir. Genel olarak ampirik olabilirlik yönteminin parametrik olmayan yöntemlerin güvenilirliği ile olabilirlik fonksiyonunun etkinliğini birleştirdiği söylenebilir. Bu yöntem, rasgele örneklem modelleri, regresyon modelleri ve otoregresif modeller gibi birçok modelin parametre tahmin probleminde uygulanabilir (Owen 1988, Hall ve La Scala 1990). Bu bölümde ilk olarak ampirik olabilirlik oranı ve sonrasında konum, konum-ölçek ve regresyon modelleri için ampirik olabilirlik tahminleri ele alınacaktır.

Ampirik dağılım fonksiyonu, $X_1, X_2, \dots, X_n$, F_0 ortak dağılımına sahip $\mathbb{R}^k, k \geq 1$ de bağımsız rasgele değişkenlerin vektörleri olmak üzere

$$F_n = \frac{1}{n} \sum_{i=1}^n I_{x_i \leq x}, -\infty < x < \infty \quad (2.1)$$

olarak tanımlanır. Burada $I_{x_i \leq x}; x_i \leq x$ doğru iken 1, aksi durumda 0 değerini alan gösterge fonksiyonudur.

$x \in \mathbb{R}$ için dağılım fonksiyonu $F(x) = P(X \leq x)$ olarak tanımlanmaktadır. $F(x^-) = P(X < x)$ olarak tanımlanırsa $P(X = x) = F(x) - F(x^-)$ olacaktır. F dağılım fonksiyonuna ilişkin parametrik olmayan olabilirlik fonksiyonu

$$L(F) = \prod_{i=1}^n (F(X_i) - F(X_i^-)) \quad (2.2)$$

şeklinde tanımlanır.

F_n ile ifade edilen ampirik dağılım fonksiyonu, F fonksiyonunun parametrik olmayan EÇO tahmin edicisidir ve (2.2) eşitliğinde tanımlanan olabilirlik fonksiyonunu maksimize eder (Owen 2001). Parametrik olmayan olabilirlik yöntemi biraz daha geliştirildiğinde ampirik olabilirlik oranı elde edilmektedir. $R(F)$ ampirik olabilirlik oranı

$$R(F) = \frac{L(F)}{L(F_n)} = \prod_{i=1}^n np_i \quad (2.3)$$

olarak tanımlanmaktadır. Burada n , örnek çapı ve p_i, x_i noktasındaki olasılık değeri olup $\sum_{i=1}^n p_i = 1$ 'dir. $\mu_0 = E(X_i)$ ve $0 < Var(X_i) < \infty$ olduğu varsayıldığında $-2\log R(\mu_0)$ dağılımında $n \rightarrow \infty$ iken $\chi^2_{(1)}$ dağılımına yakınsamaktadır (Owen 2001).

Bu bölümde ampirik olabilirlik yöntemi ile parametre tahmini üzerinde durulacaktır. Sırasıyla konum, konum-ölçek ve doğrusal regresyon model parametrelerinin ampirik olabilirlik yöntemi ile nasıl tahmin edildiği aşağıda verilmiştir.

2.1 Ampirik Olabilirlik Yöntemi ile Konum Modelinde Parametre Tahmini

$X_1, X_2, \dots, X_n$ rasgele değişkenlerinin, bilinmeyen μ parametresi ve rasgele hatalara (U_i) bağımlı olarak

$$X_i = \mu + U_i, \quad i = 1, 2, \dots, n$$

şeklinde tanımlanmasıyla oluşturulan model, *konum modeli* olarak adlandırılır. U_i rasgele değişkenlerinin bağımsız ve aynı dağılımlı olduğu varsayımı altında X_i 'ler de bağımsız ve aynı dağılımlı olur (Maronna vd. 2006).

μ bilinmeyen konum parametresi için tahmin denklemleri (estimating equations) yani $E(X_i - \mu) = 0$ ifadesi yardımıyla bir tahmin edici $(1/n) \sum_{i=1}^n (x_i - \hat{\mu}) = 0$ eşitliğinin çözümü ile elde edilebilir. Ampirik olabilirlik yönteminde ise p_i değerlerine ilişkin

$$\sum_{i=1}^n p_i = 1 \quad (2.4)$$

ve tahmin denklemleri yardımıyla tanımlanan

$$\sum_{i=1}^n p_i(x_i - \mu) = 0 \quad (2.5)$$

kısıtları altında ampirik olabilirlik oranı

$$R(\mu) = \prod_{i=1}^n np_i \quad (2.6)$$

ifadesini maksimum yapacak şekilde tanımlanan

$$R(\mu) = \max \left\{ \prod_{i=1}^n np_i \mid p_i > 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i(x_i - \mu) = 0 \right\} \quad (2.7)$$

optimizasyon probleminin çözümü ile μ parametresi için bir tahmin edici elde edilir. Bu tahmin edici (2.5) eşitliğinde verilen kısıttan

$$\hat{\mu} = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}$$

olarak ifade edilebilir. Dolayısıyla (2.4) eşitliğine göre $\sum_{i=1}^n p_i = 1$ olduğu için

$$\hat{\mu} = \sum_{i=1}^n p_i x_i$$

biçiminde yazılabilir. Böylece tahmin edicinin elde edilebilmesi için $R(\mu)$ ampirik olabilirlik oranını maksimum yapan p_i olasılıksal ağırlıklarının belirlenmesi gerekmektedir.

$R(\mu)$ ve $\log R(\mu)$ ifadelerinin maksimum değerini aldığı noktalar aynı olduğu için $\sum_{i=1}^n \log(np_i)$ ifadesinin maksimizasyon problemi ele alınabilir.

Bu kısıtlı optimizasyon probleminin çözümünde Lagrange çarpanları yönteminden faydalanılabilir. $\lambda_0, \lambda_1 \in \mathbb{R}$ Lagrange çarpanları olmak üzere amaç fonksiyonu

$$G = \sum_{i=1}^n \log(np_i) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) - n\lambda_1 \sum_{i=1}^n p_i(x_i - \mu) \quad (2.8)$$

olarak tanımlanır. G fonksiyonunun p_i 'lere göre türevi alınıp sıfıra eşitlenirse

$$\frac{\partial G}{\partial p_i} = \frac{1}{p_i} - n\lambda_1(x_i - \mu) + \lambda_0 = 0 \quad (2.9)$$

ifadesine ulaşılır. (2.9) eşitliği p_i 'lerle çarpılarak i üzerinden toplamı alındığında

$$p_i \frac{\partial G}{\partial p_i} = 1 - n\lambda_1 p_i(x_i - \mu) + p_i \lambda_0 = 0$$

$$\sum_{i=1}^n p_i \frac{\partial G}{\partial p_i} = \sum_{i=1}^n 1 - n\lambda_1 \sum_{i=1}^n p_i(x_i - \mu) + \lambda_0 \sum_{i=1}^n p_i = 0 \quad (2.10)$$

elde edilir. (2.10) eşitliğinde (2.4) ve (2.5) eşitliklerinde tanımlanan kısıtlardaki bilgiler kullanılarak $n + \lambda_0 = 0$, yani $\lambda_0 = -n$ olarak bulunur. Bu Lagrange çarpanının değeri (2.9) eşitliğinde yerine yazıldığında

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda_1(x_i - \mu)} \quad (2.11)$$

olarak elde edilir. Eşitlik (2.11)'de yer alan p_i değerinin, (2.5) eşitliğinde verilen $\sum_{i=1}^n p_i(x_i - \mu) = 0$ kısıtında yerine yazılmasıyla λ_1 çarpanına bağlı

$$\sum_{i=1}^n \frac{(x_i - \mu)}{1 + \lambda_1(x_i - \mu)} = 0 \quad (2.12)$$

eşitliği elde edilir. Bu eşitlik $u_i = x_i - \mu$ olmak üzere $f(\lambda_1) = -\sum_{i=1}^n \log(1 + \lambda_1 u_i)$ olarak tanımlanan $-f$ fonksiyonunun λ_1 'e göre birinci türevini vermektedir. f fonksiyonu $\{\lambda_1 \in \mathbb{R} | 1 + \lambda_1 u_i > 0, i = 1, 2, \dots, n\}$ üzerinde konvektir. f fonksiyonunu minimum yapan λ_1 yardımıyla, ampirik log olabilirlik oranını maksimum yapan p_i olasılıksal ağırlıkları belirlenir.

μ bir iç nokta (interior point) olduğunda $p_i > 0$, $\sum_{i=1}^n p_i = 1$ olacak şekilde $\sum_{i=1}^n p_i(x_i - \mu) = 0$ ifadesinin bir çözümü vardır. Bununla birlikte $0 < p_i < 1$ olduğu için $1 + \lambda_1 u_i > 1/n$ eşitsizliği de sağlanmaktadır.

Bu noktada optimizasyon probleminin çözümünde Owen (2001) logaritma fonksiyonunun $1/n$ civarında ikinci dereceden Taylor açılımını kullanarak

$$\log_*(x) = \begin{cases} \log(x) & , x > 1/n \\ \log(1/n) - \frac{3}{2} + 2nx - (nx)^2/2 & , x \leq 1/n \end{cases} \quad (2.13)$$

fonksiyonunu tanımlamıştır. (2.13) eşitliğiyle tanımlanan yapı kullanılarak elde edilen

$$f_*(\lambda_1) = -\sum_{i=1}^n \log_*(1 + \lambda_1 u_i) \quad (2.14)$$

fonksiyonu tüm $\lambda_1 \in \mathbb{R}$ değerleri için konvektir ve μ bir iç nokta olduğunda f fonksiyonu ile aynı minimum noktaya sahiptir. Bu nedenle μ 'nün bir iç nokta olup olmadığına bakılmaksızın f_* fonksiyonu minimum yapılmalıdır.

Bu aşamada Newton yöntemine dayalı bir algoritma kullanılabilir. Deneysel olarak, $\|\lambda_1\| \rightarrow \infty$ iken $\|\nabla f_*(\lambda_1)\| \rightarrow 0$ olacaktır. Algoritma adım sayısı üst sınıra ulaştığında veya birinci türev vektörünün normu alt sınıra ulaştığında ya da ilk olarak gerçekleşen durumda ve $R(\mu) = -\infty$ olduğunda durdurulabilir.

2.2 Ampirik Olabilirlik Yöntemi ile Konum-Ölçek Modelinde Parametre Tahmini

$X_1, X_2, \dots, X_n$ rasgele değişkenlerinin, bilinmeyen μ konum parametresi, σ ölçek parametresi ve rasgele hatalara (U_i) bağımlı olarak

$$X_i = \mu + \sigma U_i, i = 1, 2, \dots, n$$

şeklinde tanımlanmasıyla oluşturulan model, *konum-ölçek modeli* olarak adlandırılır. U_i rasgele değişkenlerinin bağımsız ve aynı dağılımlı olduğu varsayımı altında konum modelinde olduğu gibi X_i 'ler de bağımsız ve aynı dağılımlıdır (Maronna vd. 2006).

Konum modelinde olduğu gibi p_i değerlerinin yapısından ve tahmin denklemlerinden faydalanarak

$$\sum_{i=1}^n p_i = 1 \quad (2.15)$$

$$\sum_{i=1}^n p_i \frac{1}{\sigma} (x_i - \mu) = 0 \quad (2.16)$$

ve

$$\sum_{i=1}^n p_i (x_i - \mu)^2 - \sigma^2 = 0 \quad (2.17)$$

kısıtları altında $\sum_{i=1}^n \log(np_i)$ ifadesini maksimum yapacak şekilde μ ve σ parametreleri için tahmin ediciler bulunabilir. $\lambda_0, \lambda_1, \lambda_2 \in \mathbb{R}$ Lagrange çarpanları olmak üzere amaç fonksiyonu

$$G = \sum_{i=1}^n \log(np_i) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) - n\lambda_1 \sum_{i=1}^n p_i \frac{1}{\sigma} (x_i - \mu) - n\lambda_2 \left(\sum_{i=1}^n p_i (x_i - \mu)^2 - \sigma^2 \right) \quad (2.18)$$

olarak tanımlanır. (2.18) eşitliğinde verilen G fonksiyonunun p_i 'lere göre türevi alınıp sıfıra eşitlenirse

$$\frac{\partial G}{\partial p_i} = \frac{1}{p_i} + \lambda_0 - n\lambda_1 \frac{1}{\sigma} (x_i - \mu) - n\lambda_2 ((x_i - \mu)^2 - \sigma^2) = 0 \quad (2.19)$$

ifadesine ulaşılır. (2.19) eşitliğinin her iki tarafı p_i ile çarpılıp i üzerinden toplamının alınmasıyla

$$\begin{aligned} p_i \frac{\partial G}{\partial p_i} &= 1 + p_i \lambda_0 - n\lambda_1 p_i \frac{1}{\sigma} (x_i - \mu) - n\lambda_2 (p_i (x_i - \mu)^2 - \sigma^2) = 0 \\ \sum_{i=1}^n p_i \frac{\partial G}{\partial p_i} &= \sum_{i=1}^n 1 + \lambda_0 \sum_{i=1}^n p_i - n\lambda_1 \sum_{i=1}^n p_i \frac{1}{\sigma} (x_i - \mu) \\ &\quad - \lambda_2 \left(\sum_{i=1}^n p_i (x_i - \mu)^2 - \sigma^2 \right) \\ &= 0 \end{aligned} \quad (2.20)$$

elde edilir. (2.20) eşitliğinde (2.15), (2.16) ve (2.17) eşitliklerinde tanımlanan kısıtlardaki bilgi kullanılarak $n + \lambda_0 = 0$, yani $\lambda_0 = -n$ olarak bulunur. Bu Lagrange çarpanının değeri (2.19) eşitliğinde yerine yazıldığında

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda_1 (x_i - \mu)/\sigma + \lambda_2 ((x_i - \mu)^2 - \sigma^2)} \quad (2.21)$$

olarak elde edilir. p_i değeri (2.16) eşitliğinde verilen $\sum_{i=1}^n p_i \frac{1}{\sigma} (x_i - \mu) = 0$ ve (2.17) eşitliğinde verilen $\sum_{i=1}^n p_i (x_i - \mu)^2 - \sigma^2 = 0$ kısıtlarında yerine yazıldığında λ_1 ve λ_2 ye bağlı

$$\sum_{i=1}^n \frac{(x_i - \mu)}{1 + \lambda_1 \frac{1}{\sigma} (x_i - \mu) + \lambda_2 ((x_i - \mu)^2 - \sigma^2)} = 0 \quad (2.22)$$

ve

$$\sum_{i=1}^n \frac{(x_i - \mu)^2}{1 + \lambda_1 \frac{1}{\sigma} (x_i - \mu) + \lambda_2 ((x_i - \mu)^2 - \sigma^2)} - \sigma^2 = 0 \quad (2.23)$$

eşitlikleri elde edilir. Bu iki denklemin çözümü olan ve ampirik olabilirlik oranını maksimum yapan λ_1 ve λ_2 çarpanları nümerik yöntemlerle elde edilebilir. Bu çarpanlar yardımıyla elde edilen olasılıksal ağırlıklar kullanılarak μ ve σ^2 parametreleri için tahmin ediciler

$$\hat{\mu} = \sum_{i=1}^n p_i x_i = 0$$

ve

$$\hat{\sigma}^2 = \sum_{i=1}^n p_i (x_i - \hat{\mu})^2$$

olarak elde edilebilir. Görüldüğü üzere konum ve ölçek parametrelerinin ampirik olabilirlik tahminleri, yaygın olarak kullanılan örneklem ortalaması $\bar{x} = 1/n \sum_{i=1}^n x_i$ ve örneklem varyansı $s^2 = 1/n \sum_{i=1}^n (x_i - \bar{x})^2$ tahmin edicilerinin aksine her gözleme eşit ağırlık vermemektedir. Bunun yerine gözlemlere dayalı olarak elde ettiği p_i ($i = 1, 2, \dots, n$) olasılıksal ağırlıklarını vermektedir.

2.3 Ampirik Olabilirlik Yöntemi ile Doğrusal Regresyon Modelinde Parametre Tahmini

Ampirik olabilirlik yöntemi ile regresyon parametrelerinin tahmini, tahmin denklemlerinde bir başka deyiş ile EKK (En Küçük Kareler) yönteminde oluşturulan normal denklemlerde gözlemlere eşit ağırlık vermek yerine her gözlemin olasılıksal ağırlıklara sahip olduğu düşünülerek, bu ağırlıkların çarpımını maksimize etmeyi amaçlayan kısıtlı bir optimizasyon problemidir. Burada kısıtlardan ilki bu ağırlıkların, olasılık fonksiyonu ifade edebilmesi için toplamalarının 1 olmasıdır. İkinci kısıt ise söz konusu ağırlıklı normal denklemleri ifade etmektedir.

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \varepsilon_i, i = 1, 2, \dots, n$$

doğrusal regresyon modeli ele alınsın. Burada Y_i i -inci yanıt değişken değeri, $\mathbf{X}_i$ 1 sabiti ve i -inci gözlemin açıklayıcı değişken değerlerinden oluşan k -boyutlu vektör ($k < n$), $\boldsymbol{\beta} \in \mathbb{R}^k$ bilinmeyen parametre vektörü ve ε_i ; $E(\varepsilon_i | \mathbf{X}_i) = 0$ ve $Var(\varepsilon_i | \mathbf{X}_i) = \sigma^2$ olacak şekilde tanımlanan bağımsız rasgele hata terimleridir. Regresyon parametrelerinin EKK tahmin edicilerinin

$$\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = \mathbf{0}$$

normal denklemlerinin çözümü ile elde edildiği bilinmektedir. p_i veriye paylaştırılan olasılık ağırlıklarının bir kümesi olmak üzere $\boldsymbol{\beta}$ parametresi için ampirik olabilirlik fonksiyonu

$$L(\boldsymbol{\beta}) = \max \prod_{i=1}^n p_i$$

olarak tanımlanır. Bu maksimizasyon probleminin kısıtları

$$\sum_{i=1}^n p_i = 1$$

ve

$$\sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = \mathbf{0} \quad (2.24)$$

şeklinde verilmektedir. Eşitlik (2.24) yardımıyla $\boldsymbol{\beta}$ parametre vektörü için bir tahmin edici,

$$\hat{\boldsymbol{\beta}}_{EL} = \sum_{i=1}^n (\hat{p}_i \mathbf{X}_i \mathbf{X}_i^T)^{-1} \sum_{i=1}^n \hat{p}_i \mathbf{X}_i Y_i$$

olarak yazılabilir. $\hat{\boldsymbol{\beta}}_{EL}$ tahmin edicinin elde edilebilmesi için ampirik olabilirlik oranını maksimum yapan p_i olasılıksal ağırlıkları belirlenmelidir. p_i olasılıksal ağırlıklarını

belirlemek için yukarıda tanımlanan kısıtlı optimizasyon problemi çözümlenmelidir. Bu kısıtlı optimizasyon probleminin çözümü, konum ve konum-ölçek modellerinde olduğu gibi Lagrange çarpanları kullanılarak elde edilebilir. $\lambda_0 \in \mathbb{R}$, $\lambda_1^T \in \mathbb{R}^k$ Lagrange çarpanları olmak üzere amaç fonksiyonu

$$G = \sum_{i=1}^n \log(np_i) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) - \lambda_1^T \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \quad (2.25)$$

olarak tanımlanır. (2.25) eşitliğinde verilen G fonksiyonunun p_i 'lere göre türevi alınıp sıfıra eşitlenirse,

$$\frac{\partial G}{\partial p_i} = \frac{1}{p_i} + \lambda_0 - \lambda_1^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = 0 \quad (2.26)$$

ifadesine ulaşılır. (2.26) eşitliği p_i 'lerle çarpılarak i üzerinden toplamı alındığında,

$$p_i \frac{\partial G}{\partial p_i} = 1 + p_i \lambda_0 - \lambda_1^T p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = 0$$

$$\sum_{i=1}^n p_i \frac{\partial G}{\partial p_i} = \sum_{i=1}^n 1 + \lambda_0 \sum_{i=1}^n p_i - \lambda_1^T \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = 0 \quad (2.27)$$

ifadesi elde edilir. (2.27) eşitliğinde kısıtlardaki bilgi kullanılarak $n + \lambda_0 = 0$, yani $\lambda_0 = -n$ olarak bulunur. $\lambda_1 = -n\lambda$ olarak alınıp Lagrange çarpanlarının değerleri (2.26) eşitliğinde yerine yazıldığında

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}} \quad (2.28)$$

olarak elde edilir. p_i değeri (2.24) eşitliğinde tanımlanan ikinci kısıtta yani $\sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = 0$ ifadesinde yerine yazıldığında λ 'ya bağlı

$$\sum_{i=1}^n \frac{\mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}}{1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}} = 0 \quad (2.29)$$

eşitliği elde edilir. Bununla birlikte ampirik olabilirlik fonksiyonu da

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \frac{1}{n} \frac{1}{1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}} \quad (2.30)$$

olacaktır. (2.30) eşitliğinin logaritması alınarak

$$\log(L(\boldsymbol{\beta})) = - \sum_{i=1}^n \log \{1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}\} - n \log(n) \quad (2.31)$$

ifadesi elde edilir.

p_i değerleri olasılıksal ağırlıklar olup $0 \leq p_i \leq 1$ olduğu için

$$1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \geq \frac{1}{n} ; i = 1, 2, \dots, n$$

dir. Böylece λ için bir tanım kümesi

$$D = \left\{ \lambda \mid 1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \geq \frac{1}{n} ; i = 1, 2, \dots, n \right\}$$

olarak oluşturulabilir.

$$Q(\lambda) = - \sum_{i=1}^n \log \{1 + \lambda^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}\} - n \log(n)$$

olmak üzere problem D tanım kümesi üzerinde $Q(\lambda)$ fonksiyonunun minimize edilmesine dönüşecektir. Bu noktada logaritma fonksiyonunun $1/n$ civarında ikinci dereceden Taylor açılımını kullanarak

$$\log_*(x) = \begin{cases} \log(x) & , x > 1/n \\ \log(1/n) - \frac{3}{2} + 2nx - (nx)^2/2 & , x \leq 1/n \end{cases}$$

fonksiyonundan faydalanılabilir. Böylece nümerik yöntemler aracılığıyla Lagrange çarpanları hesaplanarak parametreler için tahmin değerlerine ulaşılır.

Bununla birlikte bu optimizasyon problemin çözümünde Owen (2013), Newton yöntemine benzer bir yöntem (Damped Newton iterasyonu) ve Bondell ve Stefanski (2013) ise, eyer noktası yaklaşımını kullanmışlardır. Özdemir ve Arslan (2018) ise, problemi bu aşamaya taşımadan, p_i olasılıksal ağırlıklarını da parametrelerle birlikte bulmayı hedefleyen alternatif bir hesaplama algoritması önermiştir.

2.4 Ampirik Olabilirlik Yöntemi ile Doğrusal Regresyon Parametreleri ve Varyansın Eşanlı Tahmini

Önceki bölümde varyansın bilindiği varsayımı altında regresyon parametrelerinin tahmin problemi ele alınmıştı. Varyansın bilinmediği durumlarda regresyon parametrelerinin ampirik olabilirlik yöntemi ile tahmininde

$$\begin{aligned} \max_{p_i \in (0,1)} \sum_{i=1}^n \log(np_i) \\ \sum_{i=1}^n p_i - 1 &= 0 \\ \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} &= 0 \end{aligned}$$

olarak tanımlanan optimizasyon probleminde kısıtlara

$$\sum_{i=1}^n p_i \left[\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2 \right] = 0$$

kısıtı eklenir. Böylece varyans için de bir tahmin edici veren bir kısıtlı optimizasyon problemi oluşturulur. Bu kısıtlı optimizasyon problemi için $\lambda_0, \lambda_2 \in \mathbb{R}$, $\lambda_1^T \in \mathbb{R}^k$ Lagrange çarpanları olmak üzere Lagrange fonksiyonu

$$L(\mathbf{p}, \boldsymbol{\beta}, \sigma^2, \lambda_0, \lambda_1, \lambda_2) = F(\mathbf{p}) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) + \lambda_1^T \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} + \lambda_2 \sum_{i=1}^n p_i \left[\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2 \right] \quad (2.32)$$

olarak tanımlanmaktadır. Burada $F(\mathbf{p}) = \sum_{i=1}^n \log(np_i)$ dir.

Bölüm 2.3’de yer alan probleme, varyansa ilişkin kısıt ilave edilerek regresyon parametreleri ile birlikte varyans için de bir tahmin elde edilebilir. Eşitlik (2.32)’de verilen Lagrange fonksiyonunun p_i ’lere göre türevi alınıp gerekli işlemler yapılarak

$$p_i = \frac{1}{n - \lambda_1^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} - \lambda_2 [\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2]}$$

eşitliği elde edilir. Bu problemde $n + 2k + 2$ tane bilinmeyen vardır. Bunlardan $n + k + 1$ tanesi problemin doğasındaki bilinmeyen sayısı, $k + 1$ ise Lagrange çarpanlarının sayısıdır. p_i ($i = 1, 2, \dots, n$) değerleri amaç fonksiyonunda ve kısıtlarda yerine yazılarak problemdeki bilinmeyen sayısını n boyut azaltan dual probleme geçilebilir. Oluşan dual problemde amaç fonksiyonu

$$Q(\lambda) = - \sum_{i=1}^n \log \left\{ n - \lambda_1^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} - \lambda_2 [\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2] \right\}$$

olacaktır. Optimizasyon problemine son eklenen kısıt

$$\sum_{i=1}^n p_i \left[\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2 \right] = 0$$

yardımla

$$\hat{\boldsymbol{\beta}}_{EL} = \sum_{i=1}^n (\hat{p}_i \mathbf{X}_i \mathbf{X}_i^T)^{-1} \sum_{i=1}^n \hat{p}_i \mathbf{X}_i Y_i$$

tahmin edicisine ek olarak σ^2 için bir tahmin edici

$$\hat{\sigma}_{EL}^2 = \sum_{i=1}^n \hat{p}_i \{Y_i - \mathbf{X}_i^T \tilde{\boldsymbol{\beta}}\}^2$$

yani

$$\tilde{\sigma}^2 = \sum_{i=1}^n \frac{1}{n - \hat{\lambda}_1^T \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \tilde{\boldsymbol{\beta}}\} - \hat{\lambda}_2 [\{Y_i - \mathbf{X}_i^T \tilde{\boldsymbol{\beta}}\}^2 - \tilde{\sigma}^2]} \{Y_i - \mathbf{X}_i^T \tilde{\boldsymbol{\beta}}\}^2$$

olarak bulunabilir. Ancak bu tahmin edicideki $\boldsymbol{\beta}$ ve λ_1^T, λ_2 Lagrange çarpanları analitik olarak bulunamadığı için, önceki bölümde de belirtildiği üzere belirlenen başlangıç değerleri yardımıyla $Q(\lambda)$ fonksiyonunu minimize eden ve λ_1^T, λ_2 Lagrange çarpanları bulunur. Bulunan λ_1^T, λ_2 Lagrange çarpanları yerine yazılarak $Q(\lambda)$ fonksiyonunu maksimize eden $\boldsymbol{\beta}$ ve σ^2 değerleri yani bu parametrelerin tahmin edicileri bir yakınsama sağlanana kadar nümerik yöntemler yardımıyla hesaplanır.

3. AMPİRİK OLABİLİRLİK YÖNTEMİ İLE REGRESYON PARAMETRELERİNİN TAHMİNİNDE ALTERNATİF BİR NÜMERİK HESAPLAMA ALGORİTMASI

Optimizasyon yöntemleri birçok alanda yaygın kullanıma sahiptir. Ancak bazı optimizasyon problemlerinin çözümü matematiksel olarak açık bir şekilde elde edilememektedir. Bu durumlarda sıklıkla başvurulan yöntemlerden biri Taylor açılımından faydalanılarak oluşturulan Newton Algoritmasıdır.

$F(x)$ amaç fonksiyonu ve $G_i(x) = 0; i = 1, 2, \dots, m$ kısıt fonksiyonları olmak üzere

$$\begin{aligned} \min_{x \in \mathbb{R}^n} F(x) \\ G_i(x) = 0; i = 1, 2, \dots, m \end{aligned} \quad (3.1)$$

biçiminde kısıtlı bir optimizasyon problemi tanımlansın. Bu kısıtlı optimizasyon problemlerinin çözümünde Lagrange çarpanları yönteminin kullanılabileceği önceki bölümde belirtilmişti. Bilindiği üzere $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanları ve $L: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ olmak üzere

$$L(x, \lambda_1, \lambda_2, \dots, \lambda_m) = F(x) + \lambda_1 G_1(x) + \lambda_2 G_2(x) + \dots + G_m(x) \lambda_m \quad (3.2)$$

fonksiyonu Lagrange fonksiyonu olarak adlandırılır. Böylece (3.1) de tanımlanan optimizasyon problemi, (3.2) ile verilen $L(x, \lambda_1, \lambda_2, \dots, \lambda_m)$ fonksiyonunun $x, \lambda_1, \lambda_2, \dots, \lambda_m$ değişkenlerine göre minimize edilmesi problemine dönüşür.

$$\nabla L(x, \lambda_1, \lambda_2, \dots, \lambda_m) \equiv \begin{bmatrix} \frac{\partial L}{\partial x_1} & \frac{\partial L}{\partial x_2} & \dots & \frac{\partial L}{\partial x_n} & \frac{\partial L}{\partial \lambda_1} & \frac{\partial L}{\partial \lambda_2} & \dots & \frac{\partial L}{\partial \lambda_m} \end{bmatrix}^T = 0$$

denkleminin çözümü ile $F(x)$ fonksiyonunu minimize eden $x, \lambda_1, \lambda_2, \dots, \lambda_m$ değerleri bulunur. Bu probleme Newton algoritmasını uygulayabilmek için

$$G(x) = [G_1(x), G_2(x), \dots, G_m(x)]^T$$

ve

$$\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_m]^T$$

vektörleri tanımlansın. L fonksiyonunun birinci türev (gradient) vektörü

$$\nabla L(\mathbf{x}, \Lambda) = \begin{bmatrix} \nabla F(\mathbf{x}) + (D\mathbf{G}(\mathbf{x}))^T \Lambda \\ \mathbf{G}(\mathbf{x}) \end{bmatrix}$$

olarak yazılabilir. Burada

$$D\mathbf{G}(\mathbf{x}) = \begin{bmatrix} \frac{\partial G_1(\mathbf{x})}{\partial x_1} & \frac{\partial G_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial G_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial G_2(\mathbf{x})}{\partial x_1} & \frac{\partial G_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial G_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_m(\mathbf{x})}{\partial x_1} & \frac{\partial G_m(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial G_m(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

$\mathbf{G}(\mathbf{x})$ vektörünün Jacobian matrisidir. $H F(\mathbf{x})$, $F(\mathbf{x})$ fonksiyonunun ikinci türevi ve $H G_i(\mathbf{x})$, i -inci kısıtın ikinci türevi olmak üzere $n \times n$ boyutlu

$$B(\mathbf{x}, \Lambda) = H F(\mathbf{x}) + \sum_{i=1}^m \lambda_i H G_i(\mathbf{x})$$

matrisi tanımlansın ve $\mathbb{O}$ $m \times m$ boyutlu sıfır matrisi olarak alınsın. L fonksiyonunun ikinci türev (Hessian) matrisi

$$H L(\mathbf{x}, \Lambda) = \begin{bmatrix} B(\mathbf{x}, \Lambda) & [D\mathbf{G}(\mathbf{x})]^T \\ D\mathbf{G}(\mathbf{x}) & \mathbb{O} \end{bmatrix}$$

olarak yazılabilir.

Böylece $L(\mathbf{x}, \Lambda)$ için Newton algoritması aşağıdaki gibi özetlenebilir:

1. $(\mathbf{x}_0, \Lambda_0) \in \mathbb{R}^{n+m}$ için bir başlangıç değeri belirlenir.
2. $t = 1, 2, \dots$ adımı için $\mathbf{G}_t = \mathbf{G}(\mathbf{x}_t)$, $\mathbf{W}_t = \nabla F(\mathbf{x}_t) + [D\mathbf{G}(\mathbf{x})]^T \Lambda_t$,
 $B_t = B(\mathbf{x}_t, \Lambda_t)$, $D_t = D\mathbf{G}(\mathbf{x}_t)$ olmak üzere

$$\begin{bmatrix} B_t & D_t^T \\ D_t & \mathbb{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_t \\ \Lambda_t \end{bmatrix} = - \begin{bmatrix} \mathbf{W}_t \\ \mathbf{G}_t \end{bmatrix} \Leftrightarrow \begin{cases} B_t \mathbf{X}_t + D_t^T \Lambda_t = -\mathbf{W}_t \\ D_t \mathbf{X}_t = -\mathbf{G}_t \end{cases}$$

lineer denklem sisteminin çözümü bulunur.

3. $\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{X}_t$ olarak güncellenir.
 $\Lambda_{t+1} = \Lambda_t + \Lambda_t$
4. $\|\mathbf{X}_t\| + \|\Lambda_t\| < \varepsilon$ olduğunda algoritma durdurulur (Khounder 2012).

3.1 Lagrange Çarpanları için Newton Algoritmasının Ampirik Olabilirlik Yöntemine Uyarlanması

Ampirik olabilirlik yönteminde parametre tahminlerini elde etmek için

$$\min_{p_i \in (0,1)} \sum_{i=1}^n -\log(np_i)$$

optimizasyon probleminin

$$\begin{aligned} \sum_{i=1}^n p_i - 1 &= 0 \\ \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} &= 0 \end{aligned}$$

kısıtları altında çözümü aranmaktadır. Bu kısıtlı optimizasyon problemi için $\lambda_0 \in \mathbb{R}$, $\boldsymbol{\lambda}_1^T \in \mathbb{R}^k$ Lagrange çarpanları olmak üzere Lagrange fonksiyonunun

$$L(\mathbf{p}, \boldsymbol{\beta}, \lambda_0, \boldsymbol{\lambda}_1) = F(\mathbf{x}) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) + \boldsymbol{\lambda}_1^T \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}$$

olduğu açıkça görülmektedir. Kısıtlardan oluşan

$$\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}) = \left[\sum_{i=1}^n p_i - 1, \left[\sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \right]^T \right]^T$$

vektörünün Jacobian matrisi

$$D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}) = \begin{bmatrix} 1 & 1 & \cdots & 1 & \mathbf{0} \\ \mathbf{X}_1\{Y_1 - \mathbf{X}_1^T \boldsymbol{\beta}\} & \mathbf{X}_2\{Y_2 - \mathbf{X}_2^T \boldsymbol{\beta}\} & \cdots & \mathbf{X}_n\{Y_n - \mathbf{X}_n^T \boldsymbol{\beta}\} & \left[\sum_{i=1}^n p_i \mathbf{X}_i \mathbf{X}_i^T \right]^T \end{bmatrix}$$

olarak yazılabilir. Burada $\mathbf{0}$ ve $\left[\sum_{i=1}^n p_i \mathbf{X}_i \mathbf{X}_i^T \right]^T$ $1 \times k$ boyutlu vektörlerdir.

$$\Lambda = \begin{bmatrix} \lambda_0 \\ \boldsymbol{\lambda}_1^T \end{bmatrix}$$

olmak üzere

$$\nabla L(\mathbf{p}, \boldsymbol{\beta}, \Lambda) = \begin{bmatrix} \nabla F(\mathbf{p}, \boldsymbol{\beta}) + (D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}))^T \Lambda \\ \mathbf{G}(\mathbf{p}, \boldsymbol{\beta}) \end{bmatrix}$$

olarak hesaplanabilen Lagrange fonksiyonunun birinci türevi

$$\nabla L(\mathbf{p}, \boldsymbol{\beta}, \Lambda) = \begin{bmatrix} -\frac{1}{p_1} + \lambda_0 + \boldsymbol{\lambda}_1^T \mathbf{X}_1 \{Y_1 - \mathbf{X}_1^T \boldsymbol{\beta}\} \\ -\frac{1}{p_2} + \lambda_0 + \boldsymbol{\lambda}_1^T \mathbf{X}_2 \{Y_2 - \mathbf{X}_2^T \boldsymbol{\beta}\} \\ \vdots \\ -\frac{1}{p_n} + \lambda_0 + \boldsymbol{\lambda}_1^T \mathbf{X}_n \{Y_n - \mathbf{X}_n^T \boldsymbol{\beta}\} \\ \boldsymbol{\lambda}_1^T \sum_{i=1}^n p_i \mathbf{X}_i \mathbf{X}_i^T \\ \sum_{i=1}^n p_i - 1 \\ \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \end{bmatrix}$$

olarak elde edilir.

$$\begin{aligned}
& B(\mathbf{p}, \boldsymbol{\beta}, \Lambda) \\
&= \begin{bmatrix} \frac{1}{p_1^2} & 0 & 0 & \cdots & 0 & -\lambda_1^T \mathbf{X}_1 \mathbf{X}_1^T \\ 0 & \frac{1}{p_2^2} & 0 & \cdots & 0 & -\lambda_1^T \mathbf{X}_2 \mathbf{X}_2^T \\ 0 & 0 & \frac{1}{p_3^2} & \cdots & 0 & -\lambda_1^T \mathbf{X}_3 \mathbf{X}_3^T \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{p_n^2} & -\lambda_1^T \mathbf{X}_n \mathbf{X}_n^T \\ -[\lambda_1^T \mathbf{X}_1 \mathbf{X}_1^T]^T & -[\lambda_1^T \mathbf{X}_2 \mathbf{X}_2^T]^T & -[\lambda_1^T \mathbf{X}_3 \mathbf{X}_3^T]^T & \cdots & -[\lambda_1^T \mathbf{X}_n \mathbf{X}_n^T]^T & \mathbb{O}_{k \times k} \end{bmatrix}
\end{aligned}$$

olmak üzere L Lagrange fonksiyonunun ikinci türev matrisi ise

$$HL(\mathbf{x}, \Lambda) = \begin{bmatrix} B(\mathbf{p}, \boldsymbol{\beta}, \Lambda) & [D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta})]^T \\ D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}) & \mathbb{O} \end{bmatrix}$$

olarak elde edilir. Böylece

$$\begin{aligned}
\mathbf{W}_t &= \nabla F(\mathbf{p}, \boldsymbol{\beta}) + [D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta})]^T \Lambda_t \\
&= \begin{bmatrix} -\frac{1}{p_1} + \lambda_0 + \lambda_1^T \mathbf{X}_1 \{Y_1 - \mathbf{X}_1^T \boldsymbol{\beta}\} \\ -\frac{1}{p_2} + \lambda_0 + \lambda_1^T \mathbf{X}_2 \{Y_2 - \mathbf{X}_2^T \boldsymbol{\beta}\} \\ \vdots \\ -\frac{1}{p_n} + \lambda_0 + \lambda_1^T \mathbf{X}_n \{Y_n - \mathbf{X}_n^T \boldsymbol{\beta}\} \\ -\lambda_1^T \sum_{i=1}^n p_i \mathbf{X}_i \mathbf{X}_i^T \end{bmatrix}^T
\end{aligned}$$

ve

$$\begin{aligned}
G_t &= \mathbf{G}(\mathbf{p}, \boldsymbol{\beta}), \\
B_t &= B(\mathbf{p}, \boldsymbol{\beta}, \Lambda), \\
D_t &= D\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}), \\
\mathbf{A}_t &= \begin{bmatrix} \mathbf{p} \\ \boldsymbol{\beta} \end{bmatrix}_{(n+k) \times 1}
\end{aligned}$$

olmak üzere

$$\begin{bmatrix} B_t & D_t^T \\ D_t & \mathbb{O} \end{bmatrix} \begin{bmatrix} \mathbf{A}_t \\ \Lambda_t \end{bmatrix} = - \begin{bmatrix} \mathbf{W}_t \\ \mathbf{G}_t \end{bmatrix} \Leftrightarrow \begin{cases} B_t \mathbf{A}_t + D_t^T \Lambda_t = -\mathbf{W}_t \\ D_t \mathbf{A}_t = -\mathbf{G}_t \end{cases}$$

lineer denklem sisteminin Newton algoritması aracılığıyla çözümü ile p_i ($i = 1, 2, \dots, n$), $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \dots \ \beta_{k-1}]^T$ ve Λ_t Lagrange çarpanlarına ilişkin tahmin değerlerine ulaşılabilir.

Benzer şekilde varyansın regresyon parametreleri ile eşanlı tahmini için

$$\sum_{i=1}^n p_i [\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2] = 0$$

kısıtı probleme eklendiğinde

$$\begin{aligned} L(\mathbf{p}, \boldsymbol{\beta}, \lambda_0, \lambda_1) = & F(x) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) + \lambda_1^T \sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \\ & + \lambda_2 \sum_{i=1}^n p_i [\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2] \end{aligned}$$

ve

$$\mathbf{G}(\mathbf{p}, \boldsymbol{\beta}) = \left[\sum_{i=1}^n p_i - 1, \left[\sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \right]^T, \sum_{i=1}^n p_i [\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2 - \sigma^2] \right]^T$$

olacaktır ve bu fonksiyonların birinci ve ikinci türev matrisleri sadece $\boldsymbol{\beta}$ parametre vektörünün tahminine benzer şekilde elde edilerek tahmin işlemi gerçekleştirilir.

3.2 Simülasyon Çalışması

Bu bölümde, regresyon parametrelerinin ampirik olabilirlik yöntemiyle elde edilen tahmin edicilerinin, EÇO ve EKK tahmin edicileri ile 1000 denemeden oluşan simülasyon çalışması yardımıyla karşılaştırılması amaçlanmıştır. Bu amaç doğrultusunda farklı örnek çapı ve parametre sayılarında, beklenen değeri sıfır ve varyansı 1 olan normal dağılımdan veri setleri üretilmiştir. Bununla birlikte regresyon modelindeki β parametreleri de rassal olarak belirlenmiştir. Parametre tahminlerini elde etmede kullanılan algorithmada başlangıç değeri olarak üretilen parametre değerleri kullanılmıştır.

Ampirik olabilirlik yöntemi ile elde edilen tahmin edicileri ve EÇO tahmin edicisinin EKK tahmin edicisine göre görelî etkinlik (GE) değeri

$$GE = \frac{\sum_{j=1}^{1000} \|\hat{\beta}_j - \beta\|^2}{\sum_{j=1}^{1000} \|\hat{\beta}_j^{EKK} - \beta\|^2}$$

ifadesi kullanılarak belirlenmiştir. Regresyon modelindeki farklı parametre sayıları ($k = 2, 5, 15$) ve farklı örnek çapları ($n = 30, 50, 100$) için görelî etkinlik değeri hesaplanmıştır. İlk olarak hatalara ilişkin $\varepsilon_i \sim N(0, 1)$ varsayımı altında bulunan görelî etkinlik değeri hesaplanmış ve çizelge 3.1’de özetlenmiştir. Burada, EL_p Lagrange-Newton metodu yardımıyla, EL_d ise dual optimizasyon problemi ile elde edilen ampirik olabilirlik tahmin edicisini ifade etmektedir. Bununla birlikte EÇON hataların normal dağıldığı varsayımında elde edilen en çok olabilirlik tahmin edicisini, EÇOT hataların ϑ serbestlik dereceli t dağılımına sahip olduğu varsayımında elde edilen en çok olabilirlik tahmin edicisini belirtmektedir.

Çizelge 3.1 Hatalar standart normal dağılıma sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$			$k = 5$			$k = 15$		
		EL_p	EL_d	EÇON	EL_p	EL_d	EÇON	EL_p	EL_d	EÇON
$\varepsilon_i \sim N(0, 1)$	$n = 30$	0,0015	0,6348	1	0,0027	0,6712	1	0,0168	0,5541	1
	$n = 50$	0,0005	0,3662	1	0,0007	0,6514	1	0,0028	0,5626	1
	$n = 100$	0,0001	0,0244	1	0,0001	0,0213	1	0,0002	0,6905	1

Hataların dağılımı normal ve varyans bilindiğinde EKK ve EÇO tahmin edicileri aynı etkinlik düzeyine sahiptir. Ancak EL tahmin edicilerinin ele alınan tüm k ve n değeri için hesaplanan görelî etkinlik değeri bu iki tahmin ediciye göre oldukça küçük olduğu görülmektedir.

Çizelge 3.2 Hatalar 3 serbestlik dereceli t dağılımına sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$			$k = 5$			$k = 15$		
		EL _p	EL _d	EÇO _T	EL _p	EL _d	EÇO _T	EL _p	EL _d	EÇO _T
$\varepsilon_i \sim t_3$	$n = 30$	0,0003	0,6948	0,5431	0,0006	0,6447	0,4404	0,0059	0,4693	0,8404
	$n = 50$	0,0001	0,2153	0,5647	0,0002	0,2941	0,1780	0,0009	0,5610	0,8699
	$n = 100$	0,0001	0,0862	0,4024	0,0000	0,3971	0,4191	0,0001	0,5077	0,5891

Çizelge 3.2'ye bakıldığında, hatalar 3 serbestlik dereceli t dağılımdan geldiğinde EÇO tahmin edicisinin EKK tahmin edicisinden daha etkin olduğu görülmektedir. Bununla birlikte EL tahmin edicileri söz konusu iki tahmin ediciden de daha etkindir.

Hataların dağılımı 3 serbestlik dereceli t dağılımına sahip iken dağılımın basıklığı artırıldığında ise görelî etkinlik değerleri çizelge 3.3'de verildiği gibi elde edilmektedir.

Çizelge 3.3 Hatalar basıklığı arttırılmış 3 serbestlik dereceli t dağılımına sahipken tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$			$k = 5$			$k = 15$		
		EL _p	EL _d	EÇO _T	EL _p	EL _d	EÇO _T	EL _p	EL _d	EÇO _T
$\varepsilon_i \sim t_3$	$n = 30$	0,0001	0,6622	0,3079	0,0001	0,6713	0,5469	0,0001	0,4999	0,5659
	$n = 50$	0,0001	0,3940	0,6920	0,0001	0,6182	0,5871	0,0001	0,4326	0,6047
	$n = 100$	0,0001	0,0294	0,5509	0,0001	0,4953	0,4518	0,0001	0,5085	0,5623

Çizelge 3.3 yardımıyla dağılımın basıklığı arttırıldığında da EL tahmin edicisinin EKK ve EÇO tahmin edicilerinden daha etkin olduğu görülmektedir.

Gözlem değerleri normal dağılıma sahip olduğunda ikişer sapan gözlem eklenerek hesaplanan görelî etkinlik değerleri çizelge 3.4'de verildiği gibidir. Verilerin $(n - 2) \times k$ tanesi, 0 ortalamalı ve 1 varyanslı normal dağılımdan, diğer $2 \times k$ tanesi 10 ortalamalı ve 1 varyanslı normal dağılımdan üretilmiştir. Hatalar ise 3 serbestlik dereceli t dağılımından üretilmiştir.

Çizelge 3.4 Gözlem değerleri standart normal dağılıma sahip ve iki sapan gözlem olması durumunda tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$			$k = 5$			$k = 15$		
		EL _p	EL _d	EÇÖ _T	EL _p	EL _d	EÇÖ _T	EL _p	EL _d	EÇÖ _T
$\varepsilon_i \sim N(0,1)$	$n = 30$	0,0034	0,4134	0,6109	0,0017	0,4267	0,6495	0,0062	0,2962	0,6965
	$n = 50$	0,0011	0,4389	0,7201	0,0009	0,4178	0,7061	0,0011	0,3835	0,7134
	$n = 100$	0,0001	0,4059	0,5708	0,0001	0,3316	0,5158	0,0002	0,4223	0,5668

Çizelge 3.4'e bakıldığında da farklı parametre sayıları ve örnek çapları için her iki yöntemle elde edilen EL tahmin edicilerinin EKK ve EÇÖ tahmin edicilerinden daha etkin oldukları görülmektedir.

Benzer olarak veri seti üretilirken hatalarda sapan gözlemler olması durumu ele alınabilir. Hatalar normal dağılıma sahipken ikişer sapan gözlem eklenerek hesaplanan görelî etkinlik değerleri çizelge 3.5'de verildiği gibidir.

Çizelge 3.5 Hatalar standart normal dağılıma sahip ve iki sapan gözlem olması durumunda tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$			$k = 5$			$k = 15$		
		EL _p	EL _d	EÇÖ _T	EL _p	EL _d	EÇÖ _T	EL _p	EL _d	EÇÖ _T
$\varepsilon_i \sim N(0,1)$	$n = 30$	0,0002	0,7510	0,1025	0,0004	0,4591	0,1587	0,0019	0,2946	0,2395
	$n = 50$	0,0001	0,3811	0,1377	0,0001	0,5467	0,2312	0,0004	0,2577	0,2372
	$n = 100$	0,0001	0,2470	0,2444	0,0001	0,4301	0,3241	0,0001	0,3449	0,3470

Çizelge 3.4 - 3.5 oluşturulurken kullanılan EÇÖ tahmin edicisi, hataların dağılımı t olduğu varsayımı altında hesaplanmıştır. İkinci durumda hataların çoğu normal dağılımdan üretilmiş olmasına rağmen yapılan simülasyon çalışmalarına göre EL_p ve EL_d tahmin edicileri gibi EÇÖ tahmin edicisinin de farklı örnek çapları ve parametre sayıları için EKK tahmin edicisinden daha etkin olduğu görülmektedir.

Varyansların regresyon parametreleri ile eş anlı tahmininde ampirik olabilirlik tahminlerinin Lagrange-Newton metodu yardımıyla da elde edilebilir olduğunu

göstermek için yapılan simülasyon çalışması sonuçları aşağıdaki çizelgelerde özetlenmiştir. Çizelge 3.6 - 3.7 hataların standart normal dağılımdan üretildiğinde elde edilen sonuçları göstermektedir. Çizelge 3.8 - 3.9 ise 3 serbestlik t dağılımından üretilen hatalarla elde edilen görelî etkinlik değeri verilmektedir. Hatalar standart normal dağılıma sahip olduğunda beklenildiği gibi EKK tahmin edicileri ampirik olabilirlik yönteminden daha etkindir. Ancak alternatif nümerik hesaplama yöntemi ile elde edilen ve EL_p ile gösterilen ampirik olabilirlik tahminlerinin bu durumda bile görelî etkinliklere göre üstün olduğu görülmektedir.

Çizelge 3.6 Hatalar standart normal dağılıma sahipken regresyon parametrelerinin tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$		$k = 5$		$k = 15$	
		EL_p	EL_d	EL_p	EL_d	EL_p	EL_d
$\varepsilon_i \sim N(0,1)$	$n = 30$	0,0001	1,3775	0,0001	1,0901	0,0001	0,9725
	$n = 50$	0,0001	1,6053	0,0001	1,1749	0,0001	1,1034
	$n = 100$	0,0001	1,4840	0,0001	1,0602	0,0001	0,8692

Çizelge 3.7 Hatalar standart normal dağılıma sahipken varyansa ilişkin tahmin edicilerin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$		$k = 5$		$k = 15$	
		EL_p	EL_d	EL_p	EL_d	EL_p	EL_d
$\varepsilon_i \sim N(0,1)$	$n = 30$	0,0015	0,0742	0,0013	0,0346	0,0013	0,1376
	$n = 50$	0,0022	0,2299	0,0013	0,0909	0,0013	0,1659
	$n = 100$	0,0018	0,1950	0,0011	0,0659	0,0012	0,1483

Çizelge 3.8 Hatalar 3 serbestlik dereceli t dağılımına sahipken regresyon parametrelerinin tahmin edicilerinin EKK tahmin edicisine göre görelî etkinlikleri

		$k = 2$		$k = 5$		$k = 15$	
		EL_p	EL_d	EL_p	EL_d	EL_p	EL_d
$\varepsilon_i \sim t_3$	$n = 30$	0,0005	1,2295	0,0018	0,1949	0,0017	0,9983
	$n = 50$	0,0001	1,5874	0,0003	0,1108	0,0015	0,9770
	$n = 100$	0,0001	1,5878	0,0001	0,1125	0,0001	0,9670

Benzer sonuçlar hatalar 3 serbestlik dereceli dağılıma sahip olduğunda da çizelge 3.8-3.9 yardımıyla görülebilir.

Çizelge 3.9 Hatalar 3 serbestlik dereceli t sahipken varyansa ilişkin tahmin edicilerin EKK tahmin edicisine göre görel etkinlikleri

		$k = 2$		$k = 5$		$k = 15$	
		EL _p	EL _d	EL _p	EL _d	EL _p	EL _d
$\varepsilon_t \sim t_3$	$n = 30$	0,0010	1,1955	0,0015	0,0310	0,0001	0,0104
	$n = 50$	0,0003	0,6880	0,0010	0,0840	0,0001	0,0068
	$n = 100$	0,0004	0,6792	0,0010	0,0180	0,0001	0,0049

4. REGRESYON PARAMETRELERİNİN ROBUST TAHMİNİ

EÇO tahmin edicilerinin genelleştirilmiş hali olan M tahmin yönteminin hatalarda kalın kuyrukluluk veya sapan gözlem olması durumunda, bu gözlemlere daha küçük ağırlıklar atayarak tüm gözlemlere eşit ağırlıklar veren EKK ve EÇO tahmin edicilerine göre daha küçük hata kareler toplamı sağladığı bilinmektedir. Ancak veri setinde kaldıraç noktası olması durumunda M tahmin yöntemi yerine MM tahmin yöntemi kullanılarak daha sağlıklı sonuçlar elde edilebilir. Bu doğrultuda bu bölümde öncelikle hatalarda kalın kuyrukluluk ve/veya sapan gözlemler olması durumunda sağlıklı sonuçlar veren M tahmin yöntemi sonrasında ise kaldıraç noktalarına karşı dayanıklı olan MM tahmin yöntemi ile regresyon parametrelerinin tahmini hakkında kısa bilgiler verilecektir.

4.1 Regresyon Parametrelerinin M Tahmini

Hatalarda sapan gözlem, kalın kuyrukluluk ve/veya çarpıklık olması durumunda sıklıkla başvurulanan M tahmin edicileri en çok olabilirlik tahmin edicilerinin genelleştirilmiş hali olarak ifade edilebilir. Regresyon parametrelerinin M tahmin edicileri hatalara ilişkin varyansın bilindiği varsayımı altında

$$\frac{1}{n} \sum_{i=1}^n \rho(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) \quad (4.1)$$

amaç fonksiyonunun minimize edilmesiyle bulunur. (4.1) eşitliğinde yer alan ρ fonksiyonu negatif olmayan, azalmayan ve $\rho(0) = 0$ olacak şekilde seçilebilir.

f hatalara ilişkin olasılık yoğunluk fonksiyonunu, r artıkları göstermek üzere $\rho(r) = -\log(f(r))$ olarak alındığında (4.1)'de verilen amaç fonksiyonunun minimum yapılması ile, $\log(f(r))$ fonksiyonunu maksimum yapan regresyon parametrelerinin EÇO tahmin edicilerine ulaşılacağı açıkça görülmektedir. Bununla birlikte $\rho(r) = |r|$ alındığında hataların Laplace dağılıma sahip olması durumunda elde edilen en küçük mutlak sapma

(LAD) tahminlerine ulařılabilir. $\rho(r) = r^2$ olduėunda ise ama fonksiyonunu en kk yapan deėerlerin, parametrelerin EKK tahmin edicileri olduėu kolayca grlmektedir. (4.1) ile verilen ama fonksiyonunu minimize etmek iin, fonksiyonun trevi alınıp sıfıra eřitlenir. Trevlenebilir ρ fonksiyonları iin, $\rho' = \psi$ olmak zere

$$\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \psi(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) = 0 \quad (4.2)$$

fonksiyonunun czm ile regresyon parametrelerinin M tahminine ulařılır. Bu eřitlik

$$\frac{1}{n} \sum_{i=1}^n w_i \mathbf{X}_i (Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) = 0 \quad (4.3)$$

olarak yazılabilir. Burada $w(r) = \psi(r)/r$ 'dir. Bylece $\boldsymbol{\beta}$ parametre vektr iin M tahmin edicileri

$$\hat{\boldsymbol{\beta}}_M = \left(\sum_{i=1}^n w_i \mathbf{X}_i \mathbf{X}_i^T \right)^{-1} \left(\sum_{i=1}^n w_i \mathbf{X}_i Y_i \right)$$

olarak elde edilir.

rneėin, hataların serbestlik derecesi ϑ olan t daėılımına sahip olduėu varsayılarak olasılık yoėunluk fonksiyonu (1.3) eřitliėinde verilen f olarak alındıėında, $\rho = -\log(f)$ kabul edilerek $\psi = \rho' = -f'/f$ olduėu grlebilir. Bylece

$$w_i = \frac{\vartheta + 1}{\vartheta + \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2} \quad (4.4)$$

olmak zere (4.3) eřitliėinin czm ile elde edilen M tahminlerinin, hatalar ϑ serbestlik dereceli t daėılımlı iken regresyon parametrelerinin EO tahmini olduėu grlmektedir.

ρ fonksiyonu olarak Huber'in tanımlamıř olduėu

$$\rho = \begin{cases} \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ 2\kappa |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| - \kappa^2, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

fonksiyonu kullanılarak Huber M-tahmin edicileri elde edilir. Huber ρ fonksiyonu kullanıldığında $\psi = \rho'$ fonksiyonu

$$\psi = \begin{cases} \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ \text{sgn}(Y_i - \mathbf{X}_i^T \boldsymbol{\beta})\kappa, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

ve bu ψ fonksiyonu kullanılarak eşitlik (4.2) düzenlendiğinde, w_i ler

$$w_i = \begin{cases} 1, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ \text{sgn}\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \frac{1}{\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}} \kappa, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases} \quad (4.5)$$

olmak üzere (4.3) eşitliğinin çözümü ile regresyon parametrelerinin Huber M tahmini elde edilir. Burada $\text{sgn}\{x\} = \frac{x}{|x|}$ ve κ ; $\kappa \rightarrow \infty$ iken EKK, $\kappa \rightarrow 0$ iken LAD tahmin edicisini veren bir sınır değeridir (Huber 1981).

Benzer şekilde Tukey tarafından önerilen ρ fonksiyonu

$$\rho = \begin{cases} \frac{\kappa^2}{6} \left(1 - \left(1 - \left\{ \frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\kappa} \right\}^2 \right)^3 \right), & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ \frac{\kappa^2}{6}, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

kullanılırsa ψ fonksiyonu

$$\psi = \begin{cases} \left(1 - \left\{ \frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\kappa} \right\}^2 \right)^2, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ 0, & |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

ve ağırlıklar

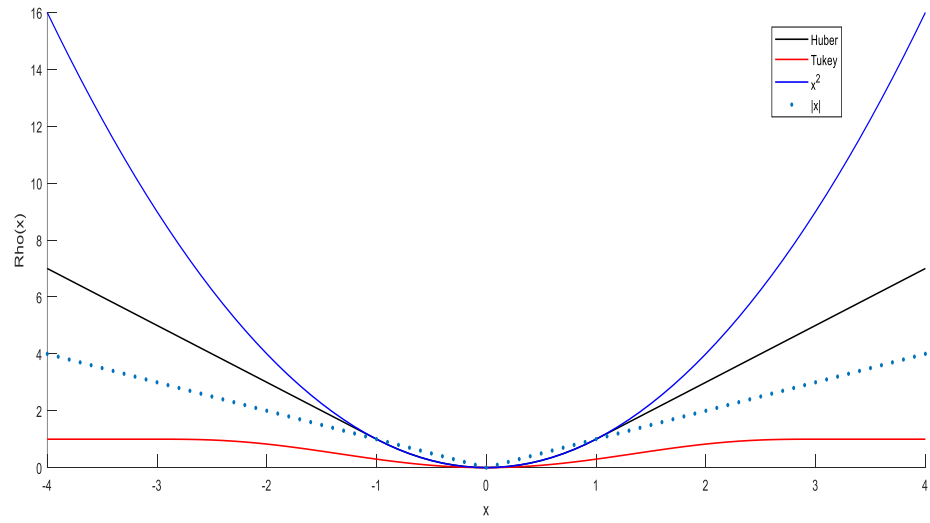
$$w_i = \begin{cases} \left(1 - \left\{\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\kappa}\right\}^2\right)^2 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ 0 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases} \quad (4.6)$$

olarak elde edilir.

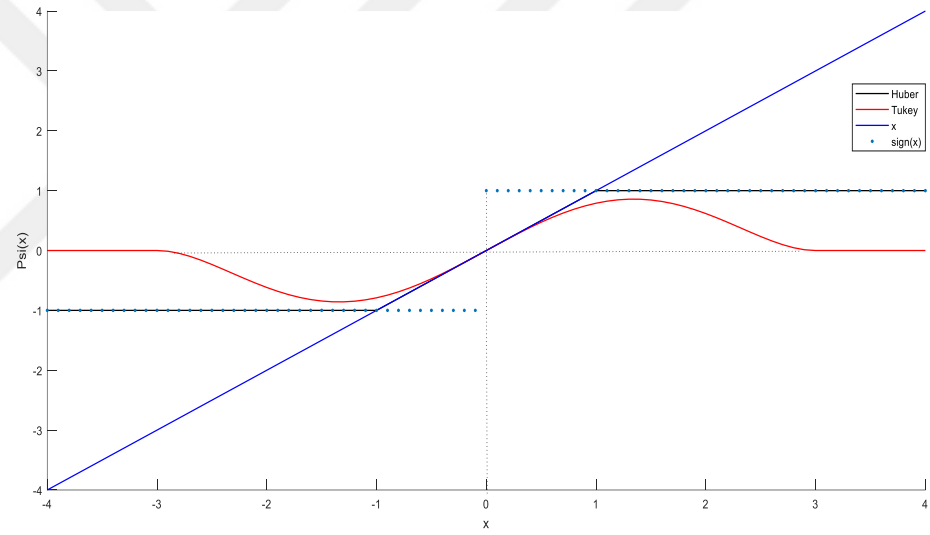
Genelde w_i ağırlıkları, hataların artmayan bir fonksiyonu olduğu için sapan gözlemler daha küçük ağırlıklar almaktadır. Böylece veri setindeki y yönünde sapan gözlemlerden daha az etkilenen tahminler elde edilmektedir (Maronna vd. 2006). Bir robustluk ölçüsü olan etki fonksiyonu, ψ 'ın bir fonksiyonu olduğu için Huber ve Tukey bi-square fonksiyonları gibi sınırlı veya sabit ρ fonksiyonları ile y yönündeki sapan gözlemlerin etkisi azaltılarak yerel anlamda robust tahmin ediciler elde edilebilir.

Şekil 4.1, Huber, Tukey bi-square, karesel ve mutlak değer fonksiyonları için ρ fonksiyonunu gösterirken, şekil 4.2 ψ fonksiyonlarını göstermektedir.

Huber ρ fonksiyonunun merkezde karesel, uçlarda ise mutlak değer fonksiyonu gibi davrandığı şekil 4.1'de görülmektedir. Böylece şekil 4.2'de olduğu gibi sınırlı ψ , dolayısıyla sınırlı etki fonksiyonu elde edildiği açıktır. Şekil 4.1'e göre Tukey bisquare ρ fonksiyonu uçlara doğru sabit değer almaktadır. Uçlarda sabit değer alan bu ρ fonksiyonu kullanıldığında ψ fonksiyonunun sıfır olduğu şekil 4.2'de gösterilmiştir. Dayanıklılığın bir ölçüsü olan etki fonksiyonu $\psi = \rho'$ 'nin bir fonksiyonu olduğu için sabit veya sınırlı ρ fonksiyonları ile yerel anlamda dayanıklı tahmin ediciler elde edilebilir.



Şekil 4.1 Bazı ρ fonksiyonları



Şekil 4.2 Bazı ψ fonksiyonları

M tahmin edicilerini elde etmede kullanılan ağırlıklar hatalardaki sapan gözlem, çarpıklık ve/veya kalın kuyrukluluk durumlarında iyi sonuçlar verirken veri setinde kaldırma noktaları, yani x yönündeki sapan gözlemler ile mücadele etmekte yetersiz kalmaktadır. Bu noktada parametre tahmininde kullanılması önerilen tahmin edicilerden bir tanesi MM tahmin edicileridir.

Regresyon parametrelerinin yanı sıra hataların varyansı da M tahmin edicileri kullanılarak tahmin edilebilir.

$$\frac{1}{n} \sum_{i=1}^n \rho \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\sigma} \right) \sigma + a\sigma$$

amaç fonksiyonunun minimize edilmesi ile $\boldsymbol{\beta}$ ve σ^2 parametrelerinin M tahminlerine ulaşılabilir. Bu amaç fonksiyonunun $\boldsymbol{\beta}$ ve σ^2 ye göre türevlerinin anılıp sıfıra eşitlenmesiyle

$$\frac{1}{n} \sum_{i=1}^n \psi \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\sigma} \right) \mathbf{X}_i = 0$$

ve $\rho_0(r) = r\psi(r) - \rho(r)$ olmak üzere

$$\frac{1}{n} \sum_{i=1}^n \rho_0 \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\sigma} \right) = a$$

denklemlerinin çözümü $\boldsymbol{\beta}$ ve σ^2 parametrelerinin M tahminlerini verir (Huber ve Ronchetti 2009).

M tahmin edicileri ile hatalardaki sapan gözlemlere karşı dayanıklı tahminler elde edilirken, x yönündeki sapan gözlemlere karşı robust tahminler elde etmek için MM tahminleri kullanılabilir.

4.2 Regresyon Parametrelerinin MM Tahmini

MM tahmin edicileri üç aşamada elde edilmektedir: Birinci aşamada başlangıç değeri olarak kullanılmak üzere regresyon parametrelerini yüksek kırılma noktasına sahip tahmin edicileri elde edilir. İkinci aşamada ise yüksek kırılma noktalı bu tahmin ediciler kullanılarak $r_1, r_2, \dots, r_n$ artıkları hesaplanır. Bu artıklar kullanılarak, hatalara ilişkin ölçek

parametresinin M tahmini bulunur. Bu tahmin ediciler S tahmin edicileri olarak bilinir. Bu aşamada kullanılan ρ fonksiyonu ρ_0 ile gösterilir ve ρ_0 yüksek kırılma noktasına sahip tahmin edici elde edilecek şekilde seçilir. Hesaplanan $r_1, r_2, \dots, r_n$ artıkları ve ρ_0 fonksiyonu ile ölçek ve regresyon parametrelerinin S tahmin edicileri elde edildikten sonra azalmayan, simetrik, sürekli ve S tahmininde kullanılan daha küçük hata kareler toplamı veren bir ρ fonksiyonu yardımı ile regresyon parametrelerinin M tahmini bulunur. Bu şekilde parametrelerin art arda iki kez M tahmini bulunarak, yüksek kırılma noktasına sahip ve etkin MM tahminlerine ulaşılır (Yohai 1987).

Bu doğrultuda başlangıç değeri olacak yüksek kırılma noktalı dayanıklı tahminlerin bulunmasından sonra

$$\min \frac{1}{n} \sum_{i=1}^n \rho \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right)$$

ifadesini sağlayan S tahmini bulunur. Burada ρ türevlenebilir bir fonksiyon ise tahmin denklemi $\rho' = \psi$ olmak üzere,

$$\frac{1}{n} \sum_{i=1}^n \psi \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right) \mathbf{X}_i = 0$$

olacaktır. Tahmin denkleminin M'den farklı olarak ölçek parametresini içerdiği görülmektedir. MM tahminleri regresyon katsayılarının yanı sıra hatalara ilişkin ölçek parametresini de tahmin etmektedir.

Ölçek parametresi için S tahmini,

$$\hat{\sigma}_s = \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2}$$

ve $K = 0,199$ olmakla beraber birinci iterasyonda w_i 'ler henüz hesaplanmamış olduğu için $\hat{\sigma}_s$ MAD olarak bilinen

$$\hat{\sigma}_s = \frac{\text{medyan}|e_i - \text{medyan}(e_i)|}{0,645}$$

tahmin edici yardımıyla hesaplanabilir. Belirlenen ρ fonksiyonu ile w_i ağırlıkları elde edilir ve köşegen elemanları bu ağırlıklardan oluşan $n \times n$ boyutlu $\mathbf{W}_i$ matrisi kullanılarak $\hat{\beta}_s = (\mathbf{X}^T \mathbf{W}_i \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{W}_i \mathbf{Y})$ hesaplanır. Bu işleme yakınsama olana kadar devam edilir.

Yohai (1987), regresyon parametrelerinin bu MM tahmin edicilerinin, hatalar normal dağıldığında yüksek kırılma noktası ve yüksek etkinliğe sahip olduğunu göstermiştir. MM tahmin edicilerinin hesaplanmasında ilk aşamada tahmin edicinin dayanıklı olması sağlanırken, ikinci aşamada belirlenen ρ fonksiyonu ile tahmin edici etkin hale getirilmektedir. Bu doğrultuda MM tahmin edicilerinin elde edilmesinde genellikle Tukey bisquare ρ fonksiyonu kullanılır.

5. REGRESYON PARAMETRELERİNİN ROBUST KISITLARLA AMPİRİK OLABİLİRLİK TAHMİNİ

Bir doğrusal regresyon modelinde hataların dağılımının normal olduğu varsayımı altında diğerlerinden üstün olan EKK tahmin edicilerinin, veri setinde sapan gözlem bulunması halinde bozulduğu bilinmektedir. Çünkü EKK yönteminde sapan gözlemler veri setinin geneliyle aynı ağırlığa sahiptir. Bu nedenle regresyon doğrusu sapan gözlem tarafından etkilenmektedir. Ampirik olabilirlik yönteminde kullanılan kısıtların EKK tahmin edicilerinin elde edilmesinde kullanılan normal denklemlerle benzerliği aşıkardır. EKK tüm gözlemlere eşit ağırlık verirken, ampirik olabilirlik yöntemi her bir gözlem için olasılıksal ağırlık belirleyerek parametre tahmini yapmaktadır. Ancak bu ağırlıklar ne kadar küçük olursa olsun, sapan gözlem olması durumunda hataların toplamı büyük olacaktır. Bu bölümde hataların toplamını küçültmek için söz konusu olasılıksal ağırlıkları destekleyerek hatalarda (y yönünde) ve tasarım matrisinde (x yönünde) sapan gözlem veya dağılımın kalın kuyruklu olması durumunda klasik tahmin edicilerden daha iyi sonuçlar veren dayanıklı tahmin edicilere dayalı kısıtlar alınmasıyla dayanıklı ampirik olabilirlik tahmin edicilerinin elde edilmesi amaçlanmaktadır. Bu doğrultuda kısıtlar ilk olarak, M tahminlerinden faydalanılarak oluşturulacaktır. İkinci olarak, M tahmin edicileri x yönündeki sapan gözlemlere karşı dayanıklı olmadığı için yüksek kırılma noktalı ve etkin MM tahmin edicileri yardımıyla oluşturulan kısıtlara sahip ampirik olabilirlik tahmin edicileri verilecektir.

5.1 M Tahmininden Faydalanılarak Oluşturulan Kısıtlarla Regresyon Parametrelerinin Ampirik Olabilirlik Tahmini

Ampirik olabilirlik yönteminde regresyon parametrelerinin

$$L(\beta) = \prod_{i=1}^n np_i \quad (5.1)$$

olarak tanımlanan ampirik olabilirlik fonksiyonunun

$$\sum_{i=1}^n p_i = 1 \quad (5.2)$$

ve

$$\sum_{i=1}^n p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} = 0 \quad (5.3)$$

kısıtları altında maksimize edilmesiyle elde edildiği önceki bölümlerde belirtilmişti. (5.3) eşitliği ile verilen kısıtın EKK daki normal denklemlerin bilinmeyen p_i lerle ağırlıklandırılmış hali olduğu görülmektedir. Ancak bu ağırlıklar parametre tahmininde, sapan gözlemlerin, kalın kuyrukluluğun ve\veya çarpıklığın etkisini azaltmakta yetersiz kalmaktadır. Bu nedenle (5.3) kısıtı yerine M tahmin edicisi elde etmede kullanılan ve eşitlik (4.3)'te verilen

$$\frac{1}{n} \sum_{i=1}^n \psi(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) \mathbf{X}_i = 0$$

denkleminde $1/n$ yerine tüm gözlemlere ayrı olasılıksal ağırlık yani p_i değerleri vererek oluşturulan

$$\sum_{i=1}^n p_i \psi(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) \mathbf{X}_i = 0 \quad (5.4)$$

kısıtı kullanılarak uçlarda sınırlı veya sabit olan ψ fonksiyonları ile etki fonksiyonu bakımından dayanıklı ve dağılım varsayımına ihtiyaç duymayan etkin tahmin ediciler bulunması amaçlanmıştır.

Hesaplama kolaylığı sağlamak adına (5.4) eşitliği, M tahmin edicilerinde olduğu gibi $w(r) = \psi(r)/r$ olmak üzere

$$\sum_{i=1}^n p_i w_i(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}) \mathbf{X}_i = 0 \quad (5.5)$$

olarak yazılabilir.

Bu çalışmada üç farklı w_i ağırlığı ele alınmıştır. İlk olarak hataların ϑ serbestlik dereceli t dağıldığı varsayımı altında eşitlik (4.4)'te verildiği gibi

$$w_i = \frac{\vartheta + 1}{\vartheta + \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}^2}$$

olduğu düşünülmüştür.

Sonrasında w_i ler $\boldsymbol{\beta}$ parametre vektörünün Huber M-tahmin edicilerinde olduğu gibi eşitlik (4.5) de verilen

$$w_i = \begin{cases} 1 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ \text{sgn}\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \frac{1}{\{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \kappa} & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

ve Tukey bisquare fonksiyonu ile elde edilen ve eşitlik(4,6) da yer alan

$$w_i = \begin{cases} \left(1 - \left\{\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\kappa}\right\}^2\right)^2 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq \kappa \\ 0 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > \kappa \end{cases}$$

şeklinde alınacaktır.

Eşitlik'(5.1) de verilen ampirik olabilirlik fonksiyonu yerine, işlem kolaylığı sağlamak adına kendisiyle aynı noktada en büyük değerini alan logaritması kullanılarak, eşitlik (5.2)'de verilen ve (5.5)'te tanımlanan kısıtlar altında en büyüklenerek regresyon parametrelerinin tahminlerine ulaşılması için klasik ampirik olabilirlik yönteminde olduğu gibi Lagrange fonksiyonundan yararlanılabilir. λ_0 ve λ_1 Lagrange çarpanları olmak üzere Lagrange fonksiyonu

$$L = \sum_{i=1}^n \log(np_i) + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) - \lambda_1^T \sum_{i=1}^n w_i p_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\} \quad (5.6)$$

olarak yazılır. Eşitlik (5.6) ile verilen L Lagrange fonksiyonunun p_i lere göre türevi alınıp gerekli düzenlemelerin yapılmasıyla, $\lambda_0 = -n$ olduğu için $\lambda = -n\lambda_1$ alınarak

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda^T w_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}} \quad (5.7)$$

olarak bulunur. Eşitlik (5.7)'de verilen p_i değerlerinin ampirik olabilirlik fonksiyonu ve kısıtlarda yerine yazılmasıyla problem

$$\log(L(\boldsymbol{\beta}, \lambda)) = - \sum_{i=1}^n \log \{1 + \lambda^T w_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}\} - n \log n \quad (5.8)$$

fonksiyonunun λ ve $\boldsymbol{\beta}$ parametrelerine göre en büyüklenmesine indirgenir. Ancak çözüm aşamasında bu iki değer birbirinin fonksiyonu olarak bulunduğu için nümerik yöntemlerden faydalanılır. Bu noktada öncelikle λ ve $\boldsymbol{\beta}$ parametreleri için başlangıç değerleri belirlenerek w_i değerleri hesaplanır. Hesaplanan w_i değerleri ve $\boldsymbol{\beta}$ parametre vektörü için belirlenen başlangıç değeri kullanılarak eşitlik (5.8)'de verilen fonksiyonun en küçüklenmesi ile λ değerleri

$$\lambda(\boldsymbol{\beta}) = \underset{\lambda}{\operatorname{argmin}} - \sum_{i=1}^n \log \{1 + \lambda^T w_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}\} - n \log n$$

olarak hesaplanır. Bu λ değerleri yerine yazılarak eşitlik (5.8) i en büyükleyen $\boldsymbol{\beta}$ parametre vektörü bulunur. Yani regresyon parametreleri

$$\hat{\boldsymbol{\beta}}_{REL} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} - \sum_{i=1}^n \log \{1 + \lambda^T w_i \mathbf{X}_i \{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}\}\} - n \log n$$

olacak şekilde hesaplanır. Bulunan λ ve $\boldsymbol{\beta}$ değerleri kullanılarak w_i değerleri güncellenir. Güncellenen w_i değerleri kullanılarak yeniden eşitlik (5.8) i optimum yapan λ ve $\boldsymbol{\beta}$ sırasıyla bulunur. Parametrelerde yakınsama oluncaya kadar bu işleme devam edilir. Hatalara ilişkin varyansın bilinmediği durumda

$$\sum_{i=1}^n p_i \psi \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\sigma} \right) \mathbf{X}_i = 0$$

ve $\rho_0(r) = r\psi(r) - \rho(r)$ olmak üzere

$$\sum_{i=1}^n p_i \rho_0 \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\sigma} \right) = a$$

kısıtları ile tahmin problemi yeniden düzenlenebilir. Bu çalışma kapsamında hatalara ilişkin varyansın bilindiği durum ele alınmıştır.

M-tahmin edicilerinde ve t-dağılımı varsayımı altında elde edilen EÇO tahmin edicilerinde kullanılan ağırlıklara bakıldığında hataların birer fonksiyonu olduğu görülmektedir. M-tahmin edicileri ve t-dağılımı varsayımı altında elde edilen EÇO tahmin edicisi hatalardaki, yani y yönündeki sapan gözlemlerin etkisini azaltmada başarılı olduğu simülasyon çalışmaları yardımıyla görülmektedir.

M-tahmin edicileri ve t-dağılımı varsayımı altında elde edilen EÇO tahmin edicisi hatalardaki, yani y yönündeki sapan gözlemlerin etkisini azaltmayı başarırken, veri setinde kaldıraç noktası olması durumunda, yani x yönünde sapan gözlem varken bozulmakta ve uygun çözümü bulmada yetersiz kalmaktadır. Bu nedenle veri setinde kaldıraç noktası olması durumunda bu noktaların etkisini azaltan başka bir dayanıklı tahmin ediciye ihtiyaç duyulmaktadır. Bu nedenle bir sonraki bölümde ampirik olabilirlik tahmin yönteminde kısıtlar oluşturulurken MM tahmin edicilerinden faydalanılmıştır.

5.2 MM Tahmininden Faydalanılarak Oluşturulan Kısıtlarla Regresyon Parametrelerinin Ampirik Olabilirlik Tahmini

Bu bölümde, veri setinde x yönünde sapan gözlem olması durumunda, yüksek kırılma noktasına sahip S tahmin edicilerinin etkin olmaması, genelleştirilmiş M (GM) tahmin edicilerinin yüksek kırılma noktasına sahip olmaması ve τ tahmin edicilerinin hesaplanma zorluğu nedeniyle yüksek kırılma noktası ve yüksek etkinliğe sahip ayrıca

hesaplanması bakımından M tahmin edicilerine benzeyen MM tahmin edicileri kullanılarak oluşturulan kısıtlarla ampirik olabilirlik tahmini ele alınmıştır.

M tahmin edicileri kullanılarak elde edilen (5.4) kısıtına benzer şekilde

$$\sum_{i=1}^n p_i \psi \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right) \mathbf{X}_i = 0 \quad (5.9)$$

kısıtı, MM tahmin edicisini bulmak için en küçüklenmesi gereken

$$\frac{1}{n} \sum_{i=1}^n \rho \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right)$$

fonksiyonunda bütün gözlemlere eşit, yani $1/n$ ağırlık vermek yerine her bir gözlemin p_i olasılıksal ağırlığına sahip olduğu düşüncesi ile x yönünde sapan gözlemlere karşı dayanıklı ampirik olabilirlik tahmin edicisi önerilmektedir.

M tahmin edicilerinden farklı olarak eşitlik (5.9) da hatalara ilişkin ölçek parametresinin de tahmin edildiği görülmektedir. Ancak bu bölüm kapsamında regresyon parametrelerinin kaldırma noktalarına karşı dayanıklı ampirik olabilirlik tahmin edicileri üzerine yoğunlaşıldığı için, ölçek parametresi için ayrı bir kısıt yazılmamıştır. MM tahmin edicilerinde olduğu gibi ölçek parametresinin S tahmini $\hat{\sigma}_s$, regresyon parametrelerinin x yönündeki sapan gözlemlere karşı dayanıklı ve etkin tahminlerinin elde edilmesi için hesaplanmıştır.

M tahmin edicilerinden farklı olarak ağırlıklar

$$w_i = \frac{\psi \left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right)}{\left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s} \right)}$$

olarak tanımlanabilir. MM tahmin edicilerinde sıklıkla Tukey bisquare ρ fonksiyonu kullanılmaktadır. Bu çalışmada da Tukey bisquare ρ fonksiyonu yardımıyla

$$w_{0i} = \begin{cases} \left(1 - \left\{ \frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s 1,547} \right\}^2 \right)^2 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq 1,547 \\ 0 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > 1,547 \end{cases}$$

ağırlığı yardımıyla bulunan S tahminleri kullanılarak hatalar güncellenir. Sonrasında bu güncel değerlerle

$$w_{1i} = \begin{cases} \left(1 - \left\{ \frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta}}{\hat{\sigma}_s 4,865} \right\}^2 \right)^2 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| \leq 4,685 \\ 0 & , \quad |Y_i - \mathbf{X}_i^T \boldsymbol{\beta}| > 4,685 \end{cases}$$

ağırlıkları hesaplanır ve bu ağırlıklar kullanılarak elde edilen parametre tahminleri ile yakınsama oluncaya kadar bu işlem devam ettirilir. Burada w_{0i} ağırlığında eşik değerinin 1,547 alınmasının nedeni, Tukey bisquare ρ fonksiyonu için yüksek kırılma noktasına sahip tahmin edici elde etmektir. w_{1i} ağırlığında ise hatalar normal dağılımlı iken %95 etkinlik sağlayan 4,685 değeri kullanılmıştır. Nihai w_i 'lerin kullanılması ile kaldıraç noktalarının etkisini azaltan ampirik olabilirlik tahmin edicileri Bölüm 4'te verilen tahmin edicilere benzer hesaplamalarla bulunabilir.

5.3 Simülasyon Çalışmaları

Bu bölümde regresyon parametrelerinin M ve MM tahmininden faydalanılarak tanımlanan dayanıklı kısıtlara sahip ampirik olabilirlik tahmin edicilerinin klasik ampirik olabilirlik tahmini ve EKK (En Küçük Kareler) tahmin edicileri ile 100 denemeden oluşan simülasyon çalışması yardımıyla karşılaştırılması amaçlanmıştır. Bu amaç doğrultusunda farklı örnek çapı ve parametre sayılarında veri setleri üretilmiştir. Bununla birlikte β regresyon modelindeki parametreler de 5 ortalamalı, 1 varyanslı normal dağılım kullanılarak rassal olarak belirlenmiştir. Parametre tahminlerini elde etmede kullanılan

algoritmada başlangıç değerleri olarak, üretilen parametre değerlerinin civarında yine rassal olarak üretilen değerler kullanılmıştır.

Hatalarda sapan gözlem olması durumunda normal dağılıma göre daha kalın kuyruklu bir dağılım olan 3 serbestlik dereceli t-dağılımı kullanılarak elde edilen EÇO tahmin edicilerinin, EKK tahmin edicisine göre daha etkin olduğu yapılan simülasyon çalışmalarında gözlemlenmiş ve ampirik olabilirlik tahmin edicilerin elde edilişinde serbestlik derecesi 3 olarak alınmıştır.

Regresyon modelindeki farklı parametre sayıları ($k = 2, 5, 15$) ve farklı örnek çapları ($n = 30, 50, 100$) için hata kare ortalamaları (HKO) değerleri hesaplanmıştır. Hatalara ilişkin $\varepsilon_i \sim N(0,1)$ varsayımı altında bulunan HKO değerleri çizelge 5.1’de, $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ varsayımı altında bulunan HKO değerleri ise çizelge 5.3’te özetlenmiştir. $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ varsayımı ile hatalarda %10 sapan gözlem olması durumunu temsil edilmektedir.

Ampirik olabilirlik yöntemi ile elde edilen tahmin edicilerin EKK tahmin edicisine göre GE değerleri $\hat{\beta}_j$ ampirik olabilirlik tahmin edicilerini simgelemek üzere

$$GE = \frac{\sum_{j=1}^{1000} \|\hat{\beta}_j - \beta\|^2}{\sum_{j=1}^{1000} \|\hat{\beta}_j^{EKK} - \beta\|^2}$$

ifadesi kullanılarak belirlenmiştir. $\varepsilon_i \sim N(0,1)$ olması durumunda bulunan GE değerleri çizelge 5.2’de, $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ olması durumunda bulunan GE değerleri ise çizelge 5.4’te verilmiştir.

Hatalar normal dağılıma sahip olduğunda beklenildiği üzere, EKK tahmin edicileri hatalar standart normal dağılımlı iken EÇO tahmin edicilerine denk olarak, diğer tahmin edicilerden daha küçük HKO değerlerine sahiptir, dolayısıyla diğerlerinden daha etkindir.

Çizelge 5.1 $\varepsilon_i \sim N(0,1)$ iken tahmin edicilerin HKO değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}	EKK
2	30	0,1032	0,0964	0,1089	0,1113	0,0776	0,0698
	50	0,0965	0,0962	0,1119	0,1111	0,0553	0,0377
	100	0,0670	0,0614	0,0718	0,0688	0,0254	0,0207
5	30	0,2126	0,2182	0,2090	0,2107	0,2567	0,2083
	50	0,1323	0,1254	0,1283	0,1266	0,1329	0,1008
	100	0,1041	0,0924	0,1037	0,0996	0,0612	0,0605
15	30	0,8651	0,7051	0,8643	0,8502	0,6261	1,1643
	50	0,4201	0,4267	0,4310	0,4342	0,4179	0,4123
	100	0,2272	0,2259	0,2269	0,2214	0,2065	0,1724

Çizelge 5.2 $\varepsilon_i \sim N(0,1)$ iken tahmin edicilerin GE değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}
2	30	1,4784	1,3811	1,5599	1,5941	1,1117
	50	2,5622	2,5517	2,9695	2,9470	1,4668
	100	3,2329	0,2966	3,4628	3,3193	1,2271
5	30	1,0203	1,0475	1,0032	1,0113	1,2324
	50	1,3117	1,244	1,2726	1,2557	1,3185
	100	1,7208	1,5273	1,7137	1,6469	1,0116
15	30	0,7430	0,6056	0,7423	0,7302	0,5377
	50	1,0190	1,0349	1,0455	1,0532	1,0135
	100	1,3175	1,3103	1,3165	1,2839	1,1978

Çizelge 5.1 -5.2'ye göre, parametre sayısı ile örnek çapı birbirine yakın olduğu durumda yani bu simülasyon çalışması için $k = 15$ ve $n = 30$ iken hatalar normal dağılımlı olmasına rağmen, ampirik olabilirlik tahminlerinin EKK tahmin edicisine göre daha etkin olduğu görülmektedir.

Hatalarda sapan gözlem olması durumunda yani $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken, HKO değerlerine bakıldığında robust kısıtlar kullanma işleminin EL tahmin

edicilerinde hata toplamını küçültmede oldukça etkili olduğu söylenebilir. Ampirik olabilirlik yöntemi ile elde edilen tahminlere ilişkin HKO değerlerinin ele alınan durumlarda EKK tahmin edicisinin HKO değerinden küçük olduğu çizelge 5.4 verilen GE değerlerine bakılarak da görülebilir. Bununla birlikte robust kısıtlara sahip ampirik olabilirlik tahminleri, klasik ampirik olabilirlik tahmininden de daha küçük HKO değerlerine sahiptir.

Çizelge 5.3 $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken tahmin edicilerin HKO değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}	EKK
2	30	5,1604	2,3008	3,7829	1,6945	0,0792	5,6693
	50	4,1271	1,7492	3,0051	1,4651	0,0463	4,7171
	100	4,5244	1,5278	2,7594	1,2009	0,0293	4,6046
5	30	5,3300	2,3537	3,4541	2,0210	0,3354	10,9145
	50	5,2281	2,2606	3,3950	1,8708	0,1329	7,7431
	100	4,7561	2,1662	3,3600	1,6773	0,0727	5,8599
15	30	18,8479	4,7446	12,7101	12,9662	1,8888	48,4243
	50	5,7769	2,0780	2,1413	1,4580	0,4929	20,1266
	100	4,2138	1,5360	2,0062	1,1010	0,5908	10,1531

Çizelge 5.4 $\varepsilon_i \sim (0,90)N(0,1) + (0,10)N(20,1)$ iken tahmin edicilerin GE değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}
2	30	0,9102	0,4058	0,6673	0,2989	0,0139
	50	0,8749	0,3708	0,6371	0,3106	0,0098
	100	0,9826	0,3318	0,5993	0,2608	0,0728
5	30	0,4883	0,2156	0,3165	0,1852	0,0307
	50	0,6752	0,2920	0,4385	0,2416	0,0171
	100	0,8116	0,3697	0,5734	0,2862	0,0124
15	30	0,3892	0,0980	0,2625	0,2678	0,0390
	50	0,2870	0,1032	0,1064	0,0724	0,0245
	100	0,4150	0,1513	0,1976	0,1084	0,0581

Durumu daha detaylı irdelemek için üretilen bir deneysel veri seti yarımıyla oluşturulan ve çizelge 5.5'te verilen hatalara ilişkin p_i ve w_i değerleri incelenebilir.

Çizelge 5.5 $\varepsilon_i \sim (0,90)N(0, 1) + (0,10)N(20, 1)$ iken bir örneğe ε_i , p_i ve w_i değerleri

	EL		EL _t				EL _{HUB}			
ε_i	p_i	$p_i \varepsilon_i$	p_i	w_i	$p_i w_i$	$p_i w_i \varepsilon_i$	p_i	w_i	$p_i w_i$	$p_i w_i \varepsilon_i$
-0,192	0,083	-0,0160	0,098	1,317	0,130	-0,025	0,096	1	0,096	-0,018
0,888	0,089	0,0798	0,103	1,055	0,108	0,096	0,100	1	0,100	0,088
-0,764	0,115	-0,0885	0,098	1,115	0,109	-0,083	0,097	0,643	0,062	-0,048
-1,402	0,078	-0,1102	0,093	0,805	0,075	-0,105	0,093	0,609	0,057	-0,080
-1,422	0,089	-0,1267	0,088	0,796	0,070	-0,100	0,095	0,808	0,077	-0,109
0,488	0,087	0,0425	0,105	1,235	0,130	0,063	0,109	1	0,109	0,053
-0,177	0,089	-0,0158	0,097	1,319	0,129	-0,022	0,105	1	0,105	-0,018
-0,196	0,214	-0,0421	0,100	1,316	0,131	-0,025	0,101	0,742	0,075	-0,014
1,419	0,089	0,1269	0,111	0,797	0,088	0,125	0,119	1	0,119	0,170
20,291	0,063	1,2823	0,100	0,009	0,001	0,019	0,106	0,069	0,007	0,151

Çizelge 5.5'e bakıldığında sonuncu sütunda yer alan sapan gözleme karşılık gelen, EKK tahmin edicisinin normal denklemlerine benzer kısıtlar kullanılarak elde edilen p_i ağırlığının, EL-t ve EL-HUB tahmin edicileri ile elde edilen p_i ve w_i ağırlıklarının çarpımından büyük olduğu görülmektedir. p_i olasılıksal ağırlıklarına ek olarak w_i ağırlıklarının kullanılmasıyla, sapan gözlemlere ampirik olabilirlik yöntemine göre daha küçük ağırlıklar hesaplandığı, dolayısıyla daha robust tahminlerin elde edildiği söylenebilir. Çizelge 5.5'in sonunda yer alan hata değeri, diğer hatalara göre aykırı bir değerdir ve bu hataya karşılık gelen eşitlik (5.3) kısıtı ile elde edilen $p_i \varepsilon_i$ ve eşitlik (5.5) kısıtı kullanılarak elde edilen $p_i w_i \varepsilon_i$ değerleri karşılaştırıldığında $p_i w_i \varepsilon_i$ değerinin daha küçük olduğu görülebilir, yani sapan gözleme daha küçük ağırlık verilerek hatalara ilişkin kareler toplamı küçültülebilir. Yapılan simülasyon çalışmaları sonucu ele alınan veri setlerinde söz konusu durumlarda robust tahmin edicilerden faydalanılarak oluşturulan

kısıtlarla elde edilen ampirik olabilirlik tahminlerinin klasik olana göre daha etkin olduğu gözlemlenmiştir.

Hatalarda sapan gözlem olması durumuna ek olarak x yönünde sapan gözlem olması durumu düşünülerek tasarım matrisi için $X \sim (0,90)N(0, 1) + (0,10)N(10, 1)$ varsayımı altında bulunan HKO değerleri hesaplanmış ve çizelge 5.6'da verilmiştir.

Çizelge 5.6 $X \sim (0,90)N(0, 1) + (0,10)N(10, 1)$ iken tahmin edicilerin HKO değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}	EKK
2	30	20,7020	10,6065	19,2812	6,6349	5,3055	21,6246
	50	21,0308	21,0209	11,9046	6,2906	5,6458	21,1715
	100	19,9857	19,9937	13,8593	7,7800	5,7460	20,1021
5	30	99,7476	91,7638	91,5987	95,5031	95,5325	107,4313
	50	101,6562	87,3627	91,4646	96,3055	95,4331	111,5400
	100	96,2866	87,8779	88,6356	90,1862	88,6007	99,2554
15	30	456,9	420,9	449,2	416,6	107,1	715,7
	50	410,5	332,5	407,6	382,2	104,6	484,5
	100	377,0036	326,8103	306,6279	334,3113	97,8496	403,5136

Veri setinde kaldıraç noktası olması durumunda MM tahmin edicisinin ağırlıkları yardımıyla elde edilen ampirik olabilirlik tahmininin HKO değerlerinin diğer tahmin edicilere göre oldukça küçük olduğu görülmektedir.

$X \sim (0,90)N(0, 1) + (0,10)N(10, 1)$ varsayımı altında hesaplanan GE değerleri çizelge 5.7'de verilmiştir. EKK tahmin edicisine göre hesaplanan GE değerleri yardımıyla $X \sim (0,90)N(0, 1) + (0,10)N(10, 1)$ olması durumunda en etkin tahmin edicinin EL_{MM} olduğu, EL_t, EL_{HUB} ve EL_{TUK} tahmin edicilerin de EL tahmini gibi kaldıraç noktalarından etkilendiği görülmektedir.

Çizelge 5.7 $X \sim (0,90)N(0, 1) + (0,10)N(10, 1)$ iken tahmin edicilerin GE değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}
2	30	0,9573	0,5044	0,8916	0,3068	0,2453
	50	0,9934	0,9929	0,5623	0,2971	0,2666
	100	0,9942	0,9946	0,6894	0,3870	0,2858
5	30	0,9284	0,8541	0,8525	0,8889	0,8891
	50	0,9114	0,7832	0,8201	0,8634	0,8556
	100	0,9701	0,8854	0,8930	0,9086	0,8927
15	30	0,6384	0,5952	0,6276	0,5820	0,1496
	50	0,8474	0,6862	0,8413	0,7888	0,2159
	100	0,9343	0,8099	0,7599	0,8285	0,2425

Durumu daha detaylı irdelemek için üretilen bir deneysel veri seti yarımıyla oluşturulan ve çizelge 5.8’de verilen p_i ve w_i değerleri incelenebilir. Kaldıraç nokrası olan son gözlem değerine daha küçük bir ağırlık atanması istenir. Ancak p_i ve w_i değerlerine bakıldığında klasik yaklaşımda bu gözleme atanan olasılıksal ağırlığın robust kısıt kullanma işlemi ile elde edilen $p_i w_i$ çarpımından daha büyük olduğu görülmektedir. EL_t ve EL_{HUB} tahmin edicilerinin kaldıraç noktasına karşılık gelen $p_i w_i$ çarpımlarının, EL tahmin edicisinin p_i değerinden küçük olmasına rağmen, diğer gözlemlerin $p_i w_i$ çarpımlarından çok da farklı olmadığı çizelge 5.8’den görülebilmektedir. Ancak, Tukey-M ve MM tahmin edicilerinde kullanılan ρ fonksiyonun yapısı gereği, robust kısıtlara sahip EL tahmin edicilerinin, kaldıraç noktasına karşılık gelen w_i değerlerinin 0 olduğu görülmektedir. Bununla birlikte EL_{TUK} tahmin edicisi ile hesaplanan p_i değerlerinin kendi içlerinde tutarlı olmaması nedeniyle, veride kaldıraç noktası olması durumunda mevcut tahmin ediciler içinden EL_{MM} tahmin edicisinin kullanılması daha uygun olacaktır.

Çizelge 5.8 $X \sim (0,90)N(0,1) + (0,10)N(10,1)$ iken bir örneğe ilişkin x_i , p_i ve w_i değerleri

x_i	EL	EL _t			EL _{HUB}			EL _{TUK}			EL _{MM}		
	p_i	p_i	w_i	$p_i w_i$	p_i	w_i	$p_i w_i$	p_i	w_i	$p_i w_i$	p_i	w_i	$p_i w_i$
-1,4275	0,0892	0,0971	0,4232	0,0411	0,0916	0,4274	0,03917	0,11236	0,9199	0,10336	0,1024	0,9818	0,1006
-0,7902	0,0922	0,0996	0,7534	0,0751	0,0951	0,6489	0,06171	0,15457	0,9997	0,15454	0,1002	0,9929	0,0994
-0,1488	0,1090	0,1278	0,7868	0,1006	0,1665	1	0,16650	0,09384	0,9092	0,08532	0,1001	0,9480	0,0949
-0,4276	0,0922	0,0992	0,9265	0,0919	0,0949	0,8052	0,07647	0,20245	0,9603	0,19443	0,0999	0,9303	0,0929
-0,6318	0,1061	0,1503	1,2690	0,1908	0,1501	1	0,15019	0,08800	0,8246	0,07257	0,0999	0,9013	0,0901
1,1503	0,1151	0,1027	0,1447	0,0148	0,1147	0,2941	0,03374	0,09496	0,9711	0,09222	0,1072	0,9678	0,1037
1,3580	0,1088	0,1019	0,1733	0,0176	0,1113	0,3274	0,03646	0,11881	0,9761	0,11598	0,1096	0,9865	0,1081
-1,3949	0,0776	0,0925	0,1881	0,0174	0,0833	0,2758	0,02299	0,04724	0,9400	0,04441	0,1020	0,8687	0,0886
0,6149	0,1071	0,1087	0,4274	0,0464	0,1274	0,6446	0,08214	0,12357	0,9961	0,12309	0,1031	0,9947	0,1025
4,9532	0,1021	0,1431	0,1632	0,0233	0,1846	0,2857	0,05278	0,1	0	0	0,1	0	0

Elde edilen p_i ve w_i ağırlıkları kullanılarak elde edilen tahmin değerleri çizelge 5.9’ da verilmiştir.

Çizelge 5.9 Tahmin değerleri

	Gerçek Parametre Değeri	Tahmin Değerleri				
		EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}
β_0	5	4,7588	5,4973	4,9419	4,9291	5,3122
β_1	5	2,6288	2,9136	2,6579	3,8735	4,9322

Çizelge 5.9’a bakıldığında veri setinde kaldıraç noktası olması durumunda robust kısıtların kullanımı ile klasik ampirik olabilirlik yöntemine göre gerçek parametre değerine daha yakın değerler elde edildiği görülmektedir. Özellikle MM tahmin yöntemi yardımıyla elde edilen ağırlıkların kullanılarak elde edilen EL_{MM} tahmininin gerçek parametre değerine yakın olduğu görülmektedir.

Çizelge 5.10 $X \sim (0,95)N(0, 1) + (0,05)N(10, 1)$ ve $\varepsilon_i \sim (0,95)N(0, 1) + (0,05)N(10, 1)$ iken tahmin edicilerin HKO değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}	EKK
2	30	12,5287	9,3522	12,3187	6,8047	3,2960	12,8769
	50	12,0095	6,6944	12,0016	4,3556	1,0904	12,1704
	100	11,0615	8,0749	11,0590	4,4360	0,0228	11,2225
5	30	86,3577	79,6514	73,6284	71,4418	98,7573	100,1528
	50	80,2003	78,6923	73,5550	70,5554	81,8866	89,8267
	100	80,0219	78,0814	72,6232	70,4154	52,7541	84,9259
15	30	431,2791	395,6866	374,8440	414,1845	414,2570	630,5305
	50	393,8606	367,4952	297,280	348,0308	397,7124	475,2467
	100	364,4392	364,4392	302,1951	348,0034	305,3740	389,5919

Çizelge 5.10 ve 5.11 $X \sim (0,95)N(0,1) + (0,05)N(10,1)$ ve $\varepsilon_i \sim (0,95)N(0,1) + (0,05)N(10,1)$ olduğunda ilgili tahmin edicilere ilişkin hata kare ortalamaları ve göreceli etkinlikleri göstermektedir. Veri setinde, tasarım matrisi ve hatalarda sapan gözlem olması durumunda ampirik olabilirlik tahmin edicileri EKK tahmin edicisine göre daha küçük hata kareler ortalaması vermektedir. Ancak diğer durumlar ele alındığında Hem x hem de y yönünde sapan gözlem olması tüm tahmin edicilerin hata kareler ortalamasında artışa neden olmaktadır. Bununla birlikte dayanıklı kısıtlara sahip EL tahmin edicilerinin mevcut EL tahmininden daha küçük hata kare ortalamasına sahip olduğu görülmektedir.

Çizelge 5.11 $X \sim (0,95)N(0,1) + (0,05)N(10,1)$ ve $\varepsilon_i \sim (0,95)N(0,1) + (0,05)N(10,1)$ iken tahmin edicilerin GE değerleri

k	n	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}
2	30	0,9730	0,7263	0,9567	0,5284	0,2560
	50	0,9868	0,5501	0,9861	0,3579	0,0896
	100	0,9857	0,7195	0,9854	0,3953	0,0020
5	30	0,8623	0,7953	0,7352	0,7133	0,9861
	50	0,8928	0,8760	0,8188	0,7854	0,9116
	100	0,9423	0,9194	0,8551	0,8291	0,6212
15	30	0,6840	0,6275	0,5945	0,6569	0,6570
	50	0,8287	0,7733	0,6255	0,7323	0,8368
	100	0,9354	0,8532	0,7757	0,8933	0,7838

Bu çalışmada hatalarda (y yönünde), tasarım matrisinde (x yönünde), hatalarda ve tasarım matrisinde (x ve y yönünde) sapan gözlem olması durumu ele alınmıştır. Yapılan simülasyon çalışmaları sonucu ele alınan veri setlerinde robust tahmin edicilerden faydalanılarak oluşturulan kısıtlarla elde edilen ampirik olabilirlik tahminlerinin, klasik olana göre daha küçük hata kare ortalamalarına sahip olduğu ve daha etkin olduğu gözlemlenmiştir. Sonuç olarak veri setinde sapan gözlem olması durumunda robust yöntemlerden faydalanılarak oluşturulan kısıtlar ile bu gözlemlerden daha az etkilenen ampirik olabilirlik tahminlerine ulaşılabileceği görülmüştür. Bir sonraki bölümde yer alan örneklerinde simülasyon sonuçlarını destekler nitelikte olduğu görülmektedir.

6. GERÇEK DÜNYADAN ÖRNEKLER

Bu bölümde önerilen tahmin edicilerin çalışabilirliğini gösteren gerçek dünyadan örnekler yer almaktadır.

6.1 Stackloss Verisi

Ampirik olabilirlik yönteminin hesaplanmasında kullanılan nümerik yöntem alternatif olarak Bölüm 3'te sunulan algoritmanın işleyişinin örneklemek için Brownlee (1965) tarafından oluşturulan stackloss veri seti kullanılmıştır. Bu veri seti 21 günlük işleme ait amonyağın nitrik aside oksidasyonu için bir bitkinin verilerinden oluşmaktadır. Yanıt değişkeni olan stackloss, bitkinin etkinliğinin bir ölçüsü olarak tanımlanmıştır. Yanıt değişkenini açıklamada kullanılan bağımsız değişkenler ise, soğutucu hava akışı, soğutucu su sıcaklığı ve asit konsantrasyonudur. Gözlemlerin ilgili değişkenler bakımından aldıkları değerler çizelge 6.1'deki gibi derlenmiştir (Brownlee 1965).

Çizelge 6.1 Stackloss verisi

Soğutucu Hava Akışı	Su Sıcaklığı	Asit Kons.	stackloss
80	27	89	42
80	27	88	37
75	25	90	37
62	24	87	28
62	22	87	18
62	23	87	18
62	24	93	19
62	24	93	20
58	23	87	15
58	18	80	14
58	18	89	14
58	17	88	13
58	18	82	11
58	19	93	12
50	18	89	8
50	18	86	7
50	19	72	8
50	19	79	8
50	20	80	9
56	20	82	15
70	20	91	15

Başlangıç noktası olarak kullanılmak üzere parametrelerin EKK ve M tahmin edicileri hesaplanıp bu iki tahmin edicinin ortalaması alınmıştır. Hesaplanan tahmin değerleri çizelge 6.2'de verilmiştir.

Çizelge 6.2 Model parametrelerinin tahminleri

	EL _p	EL _d	EKK	M
$\hat{\beta}_1$	-40,6333	-40,6279	-39,9197	-41,3469
$\hat{\beta}_2$	0,7654	0,6181	0,7156	0,8153
$\hat{\beta}_3$	1,1478	1,3952	1,2953	0,9997
$\hat{\beta}_4$	-0,1417	-0,0992	-0,1521	-0,1315

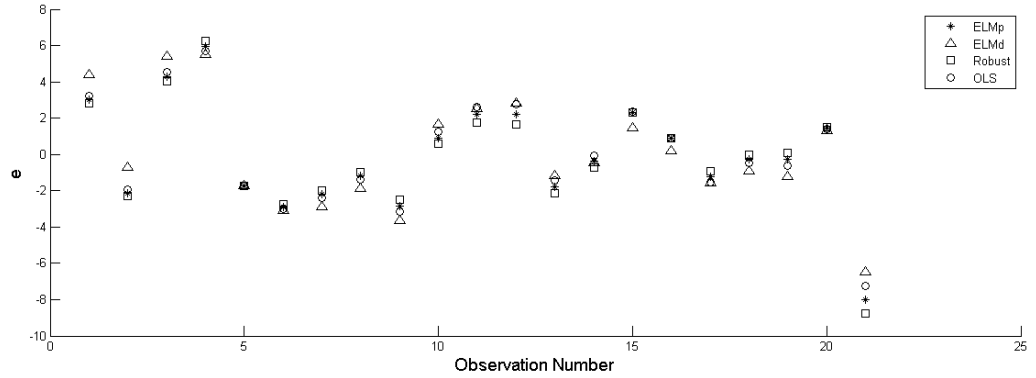
Çizelge 6.2' de verilen tahmin değerleri yardımıyla hesaplanan Hata Kare Toplamları (HKT) çizelge 6.3'de sunulmuştur.

Çizelge 6.3 Tahminlere ilişkin HKT değerleri

	EL _p	EL _d	EKK	M
HKT	180,9758	186,0087	178,8300	187,3410

Veride sapan gözlem olmaması EKK yöntemi ile yapılan tahminlerle diğer tahmin edicilerden daha küçük HKT değerleri hesaplanmıştır. Bu veri ile hesaplanan değerlere göre EL_p ile adlandırılan alternatif yöntemle ilişkin HKT değerlerinin EL_d ile gösterilen klasik ampirik olabilirlik yönteminden daha küçük HKT değerine sahip olduğu anlaşılmaktadır.

Tahmin ediciler kullanılarak elde edilen hatalar şekil 6.1’de verilmiştir. Buna göre EL_p ile hesaplanan hataların EKK ve M tahmin edicisiyle benzerliği EL_d ile olanlardan daha fazladır.



Şekil 6.1 Stackloss verisi için tahmin edicilere ait hatalar

6.2 Belçika’dan uluslararası telefon aramaları örneği

M tahmin edicileri yardımıyla oluşturulan kısıtlarla ampirik olabilirlik tahmin yönteminin işleyişine örnek vermek için çizelge 6.4’te verilen 1950-1973 yılları arasında Belçika’dan yapılan uluslararası aramalar (milyon) kullanılmıştır (Rousseeuw ve Leory 1987).

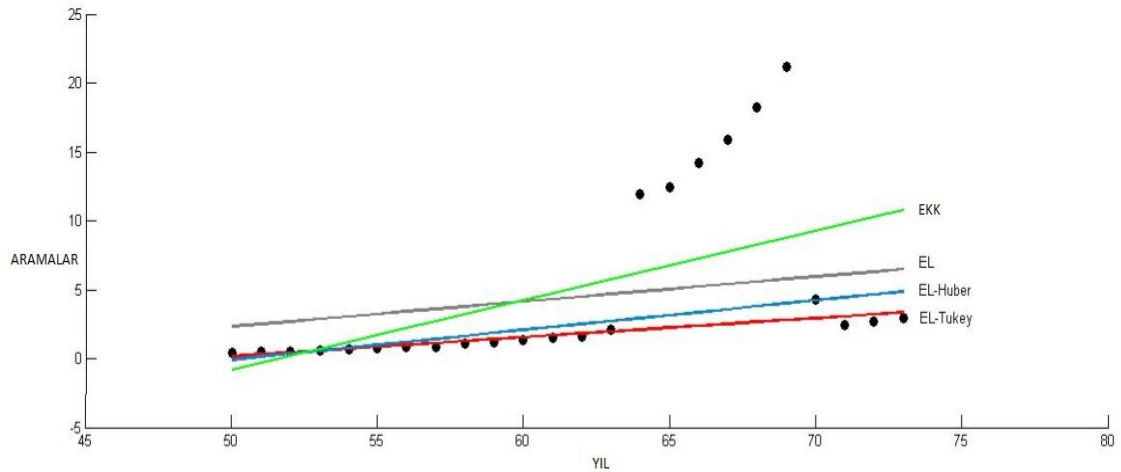
Çizelge 6.4 Belçika’dan yapılan uluslararası telefon aramaları (milyon)

Yıl	Arama sayısı	Yıl	Arama sayısı	Yıl	Arama sayısı
50	0,44	58	1,06	66	14,2
51	0,47	59	1,20	67	15,9
52	0,47	60	1,35	68	18,2
53	0,59	61	1,49	69	21,2
54	0,66	62	1,61	70	4,3
55	0,73	63	2,12	71	2,4
56	0,81	64	11,9	72	2,7
57	0,88	65	12,4	73	2,9

Çizelge 6.4 ve şekil 6.2’de yer alan serpm diyagramında görüldüğü gibi veride y yönünde sapan gözlemler bulunmaktadır. Bu sapan gözlemlerin EKK ve normal dağılımlı hata varsayımı altında elde edilen EÇO tahmin edicileri gibi klasik yöntemleri etkileyeceği açıkça görülmektedir. Dolayısıyla bu tip bir veri seti için robust tahmin yöntemlerinden faydalanmak daha uygun olacaktır. Verinin küçük oluşt dağılım varsayımlarının sağlanabilirliğini güçleştirmektedir. Bu nedenlerle Belçika’dan yapılan uluslararası aramalar ile aramaların yapıldığı yıllar arasında model kurmak için dağılım varsayımı gerektirmeyen ve y yönündeki sapan gözlemlerden etkilenmeyen tahminleri bulunarak çizelge 6.5’te verilmiştir. Bununla birlikte bu tahmin değerleri kullanılarak oluşturulan doğrusal regresyon modelleri şekil 6.2’de görsel olarak sergilenmektedir. Başlangıç değeri olarak robust bir tahmin edici kullanılmış ve hesaplanan parametre tahminleri çizelge 6.5’te verilmiştir.

Çizelge 6.5 Belçika’dan yapılan uluslararası telefon aramaları (milyon) için parametre tahminleri

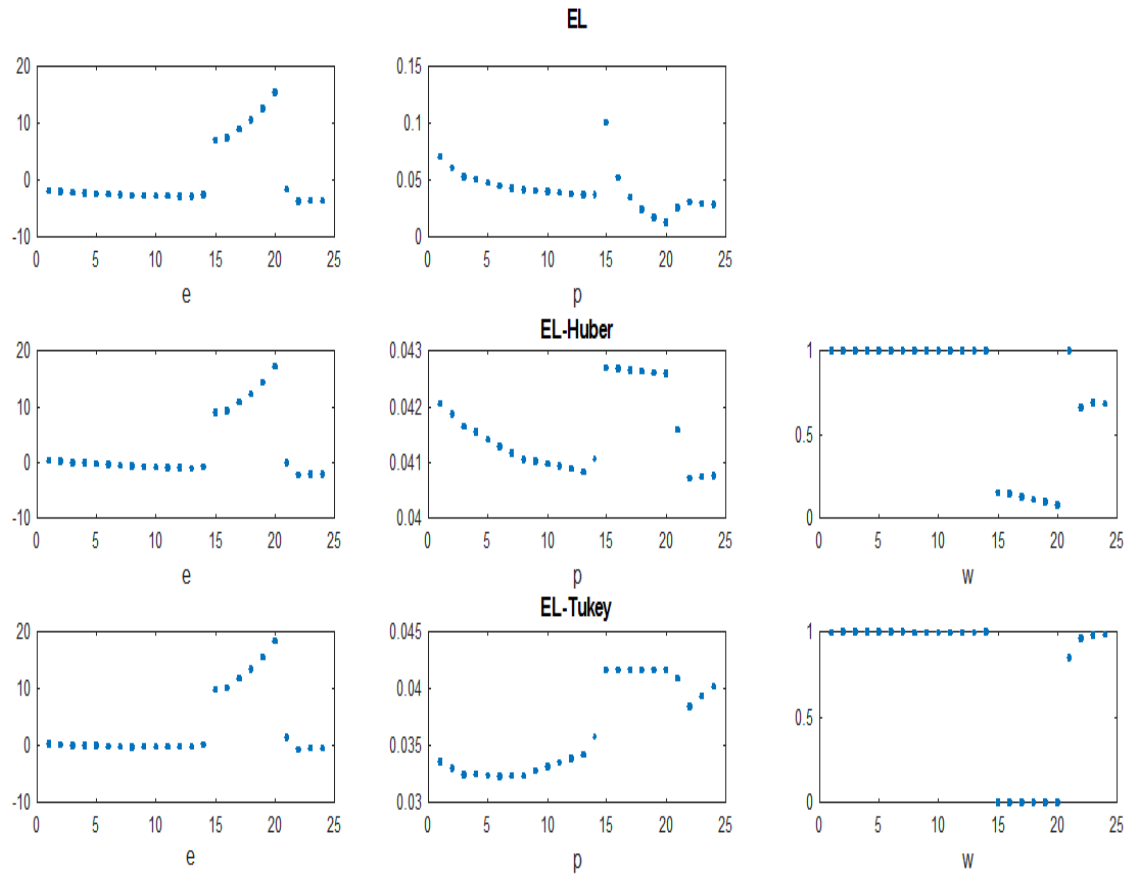
	EL	EL _{TUK}	EL _{HUB}	EKK
$\hat{\beta}_0$	-6,7247	-6,7254	-10,8456	-26,0059
$\hat{\beta}_1$	0,1809	0,1378	0,2152	0,5041



Şekil 6.2 Belçika’dan yapılan uluslararası telefon aramaları (milyon) için serpm diyagramı ve tahmin edicilere ilişkin regresyon doğruları

Şekil 6.2'ye bakıldığında EKK kullanılarak oluşturulan regresyon doğrusunun verideki sapan gözlemler tarafından etkilendiği açıkça görülmektedir. Ampirik olabilirlik tahmin edicisi veriye EKK tahmin edicisinden daha iyi uyum sağlamış ancak sapan gözlemlerin etkisi ile mücadelede tam olarak başarı sağlayamamıştır. Robust kısıtlar yardımıyla elde edilen ampirik olabilirlik tahminlerle, özellikle de Tukey fonksiyonundan faydalanılarak hesaplanan tahminlerle oluşturulan regresyon doğrularının verinin büyük çoğunluğuna oldukça iyi uyum sağladığı görülmektedir. Dolayısıyla y yönünde sapan gözlemler içeren Belçika'dan yapılan uluslararası telefon aramaları ile yıl arasındaki doğrusal bağıntıyı açıklamada M-tahmin edicileri yardımıyla oluşturulan kısıtlara sahip ampirik olabilirlik tahmininin ele alınan diğer tahminlerden daha tercih edilebilir olduğu söylenebilir.

Tahmin edilen regresyon doğrularının yanı sıra her gözleme karşılık gelen p_i ve w_i değerleri de incelenebilir. Şekil 6.3 gözlemlere ilişkin tahmin edilen hata (e), p_i ve w_i değerlerini sunmaktadır.



Şekil 6.3 Belçika'dan yapılan uluslararası aramalar (milyon) için hata, p ve w değerleri

Şekil 6.3'e bakıldığında ampirik olabilirlik yöntemi ile bulunan p_i değerlerinin sapan gözlemlere karşı tutarlı bir tutumu yoktur. Ancak w_i ağırlıklarının doğası gereği sapan gözlemlerde veri setinin genelinden daha küçük değerler aldığı görülmektedir. EL_{HUB} tahmin edicisinin hatalardaki sapan gözlemlere 0,5'ten küçük değerler verdiği, EL_{TUK} tahmin edicisinin ise sıfır değerini verdiği görülmektedir.

6.3 Yıldız Verisi

x yönünde sapan gözlem olması durumunda robust kısıtlı ampirik olabilirlik tahminlerinin nasıl sonuçlar verdiğini görmek için CYB OB1 yıldız kümesinde bulunan 47 yıldızın Hertzsprung-Russell diyagramı kullanılmıştır. Yıldızın yüzeyindeki sıcaklığın (T) logaritması açıklayıcı değişken, ışık yoğunluğunun (L/L_0) logaritması ise yanıt değişkeni olarak ele alınmıştır. Çizelge 6.6'ya bakıldığında açıklayıcı değişkende yani yıldızların yüzeyindeki sıcaklarında veri setinin genelinden daha küçük değerler olduğu göze çarpmaktadır. Bu durum şekil 6.4'te yer alan serpm diyagramında da açıkça görülmektedir.

Çizelge 6.6 Yıldız verisi (Rousseeuw ve Leory 1987)

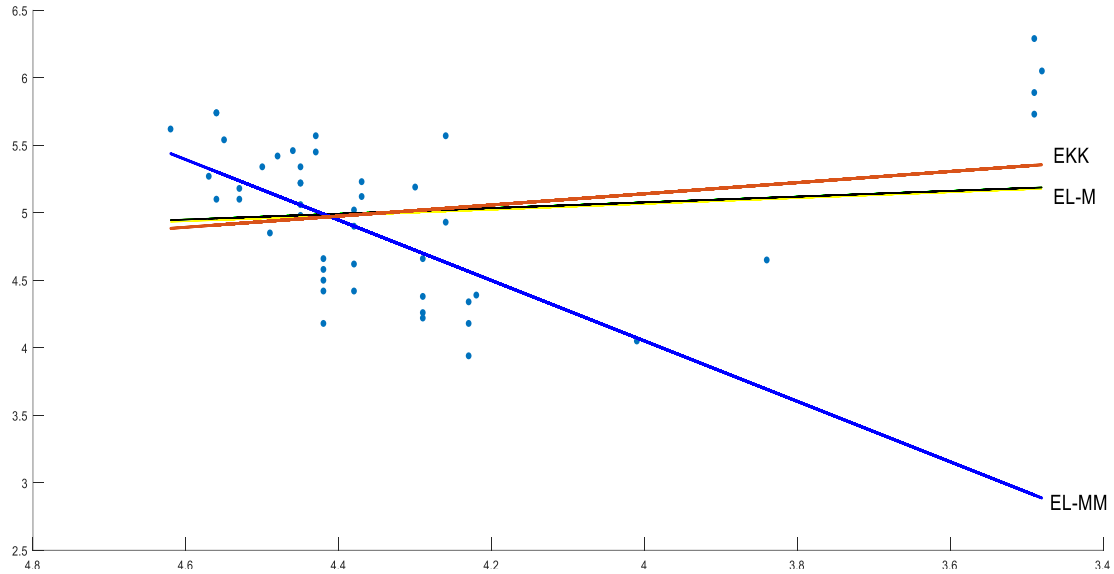
log T	log(L/L ₀)	log T	log(L/L ₀)	log T	log(L/L ₀)
4,37	5,23	4,23	3,94	4,45	5,22
4,56	5,74	4,42	4,18	3,49	6,29
4,26	4,93	4,23	4,18	4,23	4,34
4,56	5,74	3,49	5,89	4,62	5,62
4,3	5,19	4,29	4,38	4,53	5,1
4,46	5,46	4,29	4,22	4,45	5,22
3,84	4,65	4,42	4,42	4,53	5,18
4,57	5,27	4,49	4,85	4,43	5,57
4,26	5,57	4,38	5,02	4,38	4,62
4,37	5,12	4,42	4,66	4,45	5,06
3,49	5,73	4,29	4,66	4,5	5,34
4,43	5,45	4,38	4,9	4,45	5,34
4,48	5,42	4,22	4,39	4,55	5,54
4,01	4,05	3,48	6,05	4,45	4,98
4,29	4,26	4,38	4,42	4,42	4,5
4,42	4,58	4,56	5,1		

Yıldız verisi için kurulan doğrusal regresyon modelinin parametre tahminleri hesaplanarak çizelge 6.7’de verilmiştir. Çizelge 6.7’deki tahmin değerleri kullanılarak oluşturulan doğrusal regresyon modelleri de serpmeye diyagramı ile birlikte şekil 6.4’te verilmiştir.

Çizelge 6.7 Yıldız verisi için parametre tahminleri

	EL	EL _t	EL _{HUB}	EL _{TUK}	EL _{MM}	EKK
$\hat{\beta}_0$	5,9255	5,9247	5,9255	5,9255	-4,9016	6,7934
$\hat{\beta}_1$	-0,2119	-0,2150	-0,2150	-0,2111	2.2380	-0,4133

Şekil 6.4’e göre M-tahmin edicileri yardımıyla oluşturulan ampirik olabilirlik tahminleri de EKK gibi kaldırma noktalarından etkilenmektedir. Ancak kısıtlarda MM tahmin edicileri kullanılarak oluşturulan ampirik olabilirlik tahminlerinin kaldırma noktalarına karşı dayanıklı olduğu ve veri setinin büyük çoğunluğuyla uyumlu olacak şekilde parametre tahmini yapıldığı görülmektedir.



Şekil 6.4 Yıldız verisi için serpmeye diyagramı ve tahmin edicilere ilişkin regresyon doğruları

6.4 Coleman Verisi

Coleman vd. (1966) Mid-Atlantic ve New England eyaletlerindeki 20 okul üzerinde bir çalışma yapmış ve aşağıdaki veri setini toplamıştır. Bu veri, sözel ortalama test notu (altıncı sınıf öğrencilerinin tamamı) yanıt değişkeni, öğrenci başına personel maaşları, memur babaların yüzdesi, sosyo-ekonomik durum kompozit sapma (aile büyüklüğü, aile dokunulmazlığı, baba eğitimi, annenin eğitimi ve ev eşyaları), öğretmenlerin ortalama sözel test puanı, annelerin eğitim düzeyi (bir birim iki okul yılı) açıklayıcı değişkenler olmak üzere Mosteller ve Tukey (1977) tarafından kullanılmıştır. 20 okula ait ilgili değişkenlerin aldığı değerler çizelge 6.8’de verilmiştir. Çizelge 6.8 incelendiğinde X_2 ve X_3 açıklayıcı değişkenlerinin aldığı değerlerin kendi içlerinde gösterdiği değişkenlik fazladır, yani sezgisel olarak bu değişkenlerin basık olduğu görülebilir.

Çizelge 6.8 Coleman verisi (Rousseeuw ve Leory 1987)

X_1	X_2	X_3	X_4	X_5	Y
3,83	28,87	7,2	26,6	6,16	37,01
2,89	20,1	-11,71	24,4	5,17	26,51
2,86	69,05	12,32	25,7	7,04	36,51
2,92	65,4	14,28	25,7	7,1	40,7
3,06	29,59	6,31	25,4	6,15	37,1
2,07	44,82	6,16	21,6	6,41	33,9
2,52	77,37	12,7	24,9	6,86	41,8
2,45	24,67	-0,17	25,01	5,78	33,4
3,13	65,01	9,85	26,6	6,51	41,01
2,44	9,99	-0,05	28,01	5,57	37,2
2,09	12,2	-12,86	23,51	5,62	23,3
2,52	22,55	0,96	23,6	5,34	35,2
2,22	14,3	4,77	24,51	5,8	34,9
2,67	31,79	-0,96	25,8	6,19	33,1
2,71	11,6	-16,04	25,2	5,62	22,7
3,14	68,47	10,62	25,01	6,94	39,7
3,54	42,64	2,66	25,01	6,33	31,8
2,52	16,7	-10,99	24,8	6,01	31,7
2,68	86,27	15,03	25,51	7,51	43,1
2,37	76,73	12,77	24,51	6,96	41,01

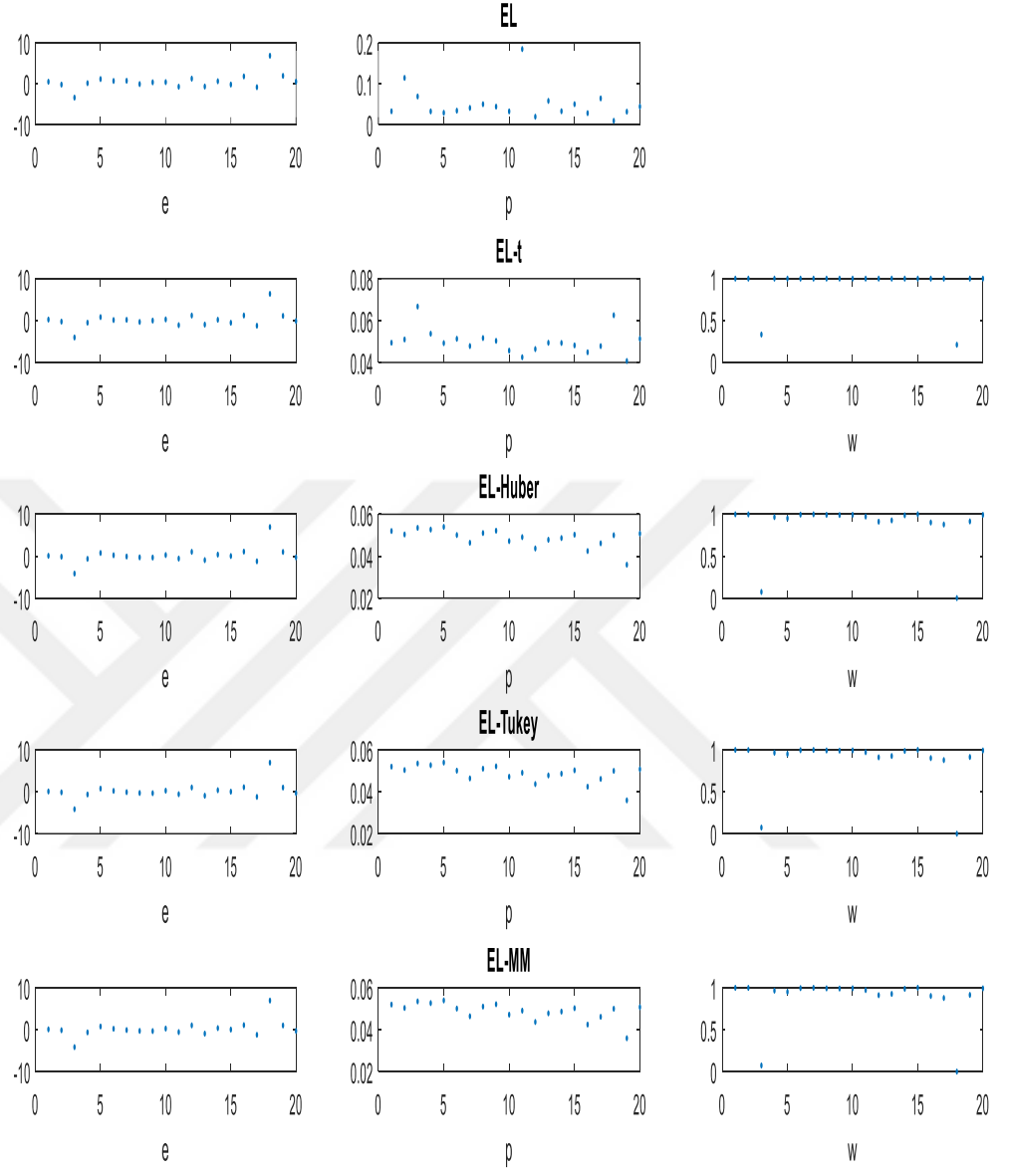
Coleman verisi için hesaplanan parametre tahminleri çizelge 6.9’da verilmiştir. Çizelge 6.9’da verilen parametre tahminlerine bakıldığında EL_{MM} tahmin edicilerinin diğer tahmin edicilerden özellikle EKK tahmin edicisinden farklı değerler aldığı görülmektedir. Açıklayıcı değişken değerlerinde olan farklılıklar EL_{MM} haricindeki tahmin edicileri benzer şekilde etkilemiştir.

Çizelge 6.9 Coleman verisi için parametre tahminleri

	EL	EL_t	EL_{HUB}	EL_{TUK}	EL_{MM}	EKK
$\hat{\beta}_0$	28,7329	28,7326	26,7332	28,7327	30,4029	19,9561
$\hat{\beta}_1$	-1,6857	-1,6853	-1,7517	-1,6851	-1,7068	-1,8106
$\hat{\beta}_2$	0,0720	0,0743	0,0695	0,0780	0,0817	0,0441
$\hat{\beta}_3$	0,6319	0,6525	0,6227	0,6545	0,6713	0,5551
$\hat{\beta}_4$	1,1628	1,1644	1,1442	1,1659	1,1692	1,1123
$\hat{\beta}_5$	-3,7864	-3,7881	-3,2752	-3,7875	-4,0909	-1,81673

Coleman verisi için hesaplanan parametre tahminleri kullanılarak hesaplanan hatalar, p ve w değerleri şekil 6.5’te verilmiştir.

Şekil 6.5’e bakıldığında klasik EL tahmin edicisinin dayanıklı kısıtlara sahip EL tahmin edicisinden farklı ağırlıklar aldığı görülmektedir. Klasik EL tahmin edicisi ile hesaplanan olasılıksal ağırlıkların hatalara göre tutarlı bir davranışı olduğu söylenemezken, dayanıklı kısıtlara sahip EL tahmin edicileri bazı gözlemlere diğerlerinden daha küçük ağırlıklar atamıştır. Bu sayede veri setindeki bozulmalara karşı daha etkin tahmin ediciler elde edilmiştir.



Şekil 6.5 Coleman verisi için tahmin edilen hata, p ve w değerleri

7. TARTIŞMA VE SONUÇ

Doğrusal regresyon modelinin parametre tahmininde sıklıkla kullanılan EKK ve EÇO tahmin edicilerinin kullanılabilmesi için sağlanması gereken model varsayımları vardır. Dağılım varsayımının sağlanmadığı durumlarda, buna ihtiyaç duymayan ampirik olabilirlik yöntemi kullanılabilir. Bununla birlikte, hataların normal dağıldığı varsayımı altında bağımsız değişken sayısı ile örnek çapının birbirine yaklaşması durumunda ampirik olabilirlik yöntemi ile elde edilen tahminlerin EKK tahminlerinden daha etkin olduğu simülasyon çalışmalarından kolaylıkla görülmektedir.

Ampirik olabilirlik yöntemi bütün gözlemlere olasılıksal ağırlıklar vererek, bu ağırlıklardan oluşan ampirik olabilirlik fonksiyonunu en büyükecek şekilde model parametrelerini bazı kısıtlar altında tahmin eder. Ancak söz konusu olasılıksal ağırlıkların verideki sapan gözlem, çarpıklılık ve\veya kalın kuyrukluluk durumları ile mücadele etmede yetersiz kaldığı sezgisel olarak anlaşılabılır. Böyle bozulmalar olması durumunda robust tahmin yöntemlerinin kullanımı oldukça yaygındır. Robust tahmin yöntemlerinden faydalanılarak oluşturulan kısıtlarla elde edilen ampirik olabilirlik tahminlerinin, bu çalışma kapsamında ele alınan sınırlı simülasyon çalışmalarında veri setlerine eklenen sapan gözlemlerden daha az etkilendiği görülmektedir. y yönünde sapan gözlem olması durumunda sıklıkla kullanılan M tahmin edicilerinin ampirik olabilirlik yöntemi ile birlikte yorumlanması ile y yönünde sapmalar olduğunda klasik ampirik olabilirlik yönteminden daha küçük hata kare ortalaması veren tahmin ediciler elde edilebileceği anlaşılmaktadır. Çünkü ampirik olabilirlik yönteminde kullanılan olasılıksal ağırlıklara ek olarak, hataların bir fonksiyonu olduğu için sapan gözlemlerin etkisini azaltan ilave bir ağırlık kullanılmaktadır. Böylece sapan gözlemlerin etkisinin azaltılması ile veri setinin geneliyle uyumlu parametre tahminleri elde edilmektedir.

M-tahmin edicileri kullanılarak elde edilen bu robust kısıtlar y yönündeki sapan gözlemlerin etkisi azaltılabilirken, x yönündeki sapan gözlemlerin etkisini azaltmak için başka bir robust tahmin yönteminden faydalanılabilir. Etkin ve kırılma noktası yüksek olup kaldıraç noktalarına karşı dayanıklı olan MM tahmin edicileri yardımıyla oluşturulan kısıtlarla elde edilen ampirik olabilirlik tahminlerinin x yönündeki sapmalar olması durumunda verinin genelini davranışını yansıtan modeller kurulabileceği yapılan

simülasyon çalışmaları ve gerçek veri seti örneğinden anlaşılmaktadır. Benzer şekilde hem x hem de y yönünde sapan gözlem olması durumunda da, MM tahmin yönteminden faydalanılarak oluşturulan kısıtlar kullanılarak tanımlanan ampirik olabilirlik tahminlerinin ele alınan diğer tahmin yöntemlerinden daha küçük hata kare ortalamasına sahip olduğu yine simülasyon çalışmaları yardımıyla sergilenmektedir.

Bu çalışmada ampirik olabilirlik yöntemi robust kısıtlarla yeniden ele alınmış ve doğrusal regresyon modelinde parametre tahminine nasıl uygulanabileceği gösterilmiştir. Bir sonraki aşamada bu tezde önerilen robust kısıtlarla elde edilen ampirik olabilirlik tahmin edicilerinin asimptotik özellikleri ve dayanıklılık ölçüleri yani etki fonksiyonu ve kırılma noktası bakımından davranışları incelenecektir. Buna ek olarak, ampirik olabilirlik yöntemlerinin kullanılabilmesi için hipotez testleri ve güven aralıkları oluşturmada da problemin robust tahmin edicilerle birlikte ele alınması araştırmacılara daha geniş kullanım alanları tanıyacaktır. Ayrıca bu çalışmada ele alınan ampirik olabilirlik yöntemi yerine, geliştirilmiş ampirik olabilirlik yönteminin kullanılması da gelecek çalışmalara konu oluşturulmaktadır.

Bu çalışmada yer alan veri setleri, tahmin edicilerin bazı özelliklerini vurgulamak için literatürde sıklıkla kullanılmaktadır. İlerleyen çalışmalarda, laboratuvar ortamında yapılan, tekrarlanması güç olan deneylerden veya üzerinde ölçüm yapılması zor olan birimlerden derlenen veri setleri, yani örneklem çapı az olan veri setleri için de kullanışlı olduğunu gösterebilmek için, önerilen tahmin yöntemi güncel veri setlerine uygulanacaktır.

KAYNAKLAR

- Bondell, H. D. and Stefanski, L.A. 2013. Efficient Robust Regression via Two-Stage Generalized Empirical Likelihood, Journal of the American Statistical Association, 108:502, 644-655
- Brownlee, K. A. 1965. Statistical Theory and Methodology in Science and Engineering. New York: Wiley. 2nd ed.
- Chen, S. X. 1993. On the accuracy of empirical likelihood confidence regions for linear regression model. Ann. Inst. Statist. Math., 45, 621–637.
- Chen, S. X. 1994. Empirical likelihood confidence intervals for linear regression coefficients. J. Multiv. Anal., 49, 24–40.
- Chen, S. X. 1996. Empirical likelihood confidence intervals for nonparametric density estimation. Biometrika, 83, 329–341.
- Chen, S. X. and Cui, H. 2003. An extended empirical likelihood for generalized linear models. Statist. Sinica, 13, 69–81.
- Chen, S. X. and Keilegom, I. V. 2009. A review on empirical likelihood methods for regression. TEST, 19,3,415-447.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. L. vd. 1966. Equality of Educational Opportunity, two volumes, Office of Education, U.S. Department of Health, Washington.
- Hall, P. and La Scala, B. 1990. Methodology and algorithms of empirical likelihood. Internat. Statist. Review, 58, 109–127.
- Hardle, W. 1990. Applied Nonparametric Regression. Cambridge University Press.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. 1986. Robust Statistics: The Approach Based on Influence Functions. New York: Wiley.
- Huber, P. J. 1964. Robust estimation of a location parameter. Annals of Mathematical Statistics, 35, 73-101.
- Huber, P. 1981. Robust Statistics. New York: Wiley.
- Huber P. and Ronchetti E. M. 2009. Robust Statistics. New York: Wiley.
- Imbens, G. W. 2002. Generalized Method of Moments and Empirical Likelihood, Journal of Business & Economic Statistics, 20:4, 493-506.

- Kitamura, Y. 2001. Asymptotic optimality of empirical likelihood for testing moment restrictions. *Econometrica*, 69, 1661–1672.
- Kitamura, Y. 2006 Empirical likelihood methods in econometrics: Theory and practice. (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.9883&rep=rep1&type=pdf>)
- Kolaczyk, E. D. 1994. Empirical likelihood for generalized linear model. *Statist. Sinica*, 4, 199–218.
- Khouider, B. 2012. Numerical Methods with Application in Finance and Economics Lecture Notes, (http://www.math.uvic.ca/~khouider/CLASSES/MATH348_S12/math348s12notes.pdf)
- Li, G. R., Tian P. and Xue L.C. 2008. Generalized Empirical Likelihood Inference in Semiparametric Regression Model for Longitudinal Data, *Acta Mathematica Sinica, English Series*, 24, 12, 2029–2040.
- Maronna, R. A., Martin, R. D. and Yohai, V.J. 2006. *Robust Statistics: Theory and Methods*. Wiley, New York.
- Mendes, B. and Tyler, D. E. 1996. Constrained M-Estimation for Regression. *Robust Statistics, Data Analysis, and Computer Intensive Methods Lecture Notes in Statistics Volume 109*, 299–320.
- Montgomery, D. C. and Peck, E.A. 1992. *Introduction to Linear Regression Analysis*. John Wiley and Sons, Inc., New York.
- Mosteller, F. and Tukey, J. W. 1977. *Data Analysis and Regression*, Addison- Wesley, Reading, MA.
- Newey, W. and Smith, R. J. 2004. Higher-order properties of GMM and generalized empirical likelihood estimators. *Econometrica* 72, 219–255.
- Owen, A. B. 1988. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75, 237–249.
- Owen, A. B. 1990. Empirical likelihood confidence regions. *Ann. Statist.*, 18, 90–120.
- Owen, A. B. 1991. Empirical likelihood for linear models. *Ann. Statist.*, 19, 1725–1747.
- Owen, A. 2001. *Empirical Likelihood*. Chapman and Hall, New York.
- Owen, A. 2013. Self-concordance for empirical likelihood. *Canadian Journal of Statistics*, 41(3), 387–397.

- Özdemir, Ş. and Arslan, O. 2018. An Alternative Algorithm of the Empirical Likelihood Estimation for the Parameter of a Linear Regression Model. *Communications in Statistics - Simulation and Computation*. Doi: 10.1080/03610918.2018.1435801.
- Qin, J. and Lawless, J. 1994. Empirical likelihood and general estimating equations. *Ann. Statist.*, 22, 300–325.
- Rao, J. N. K. and Wu, C. 2008. Empirical Likelihood Methods, *Handbook of Statistics, Volume 29, Sample Surveys: Theory, Methods and Inference*.
- Rousseeuw, P. J. 1984. Least Median of Squares Regression. *Journal of the American Statistical Association*, 79, 871-880
- Rousseeuw, P. J. and Leroy. A.M. 1987. Robust regression and outlier detection.. John wiley and sons, Newyork.
- Rousseeuw, P. J. and Yohai, V. J. 1984. Robust Regression by Means of S-Estimators. *Robust and Nonlinear Time Series*, edited by. J. Franke, W. Hardle, ve D. Martin, *Lecture Notes in Statistics*, 26, Berlin: Springer-Verlag, 256–272.
- Yohai, V. J. 1987. High breakdown-point and high efficiency estimates for regression. *The Annals of Statistics*, 15, 642–65.
- Yohai, V. J., and Zamar, R.H. 1988. High Breakdown-Point Estimates of Regression by Means of the Minimization of an Efficient Scale. *Journal of the American Statistical Association*, 83, 406–413

ÖZGEÇMİŞ

Adı Soyadı : Şenay Özdemir
Doğum Yeri : İkizdere
Doğum Tarihi : 01/01/1988
Medeni Hali : Bekar
Yabancı Dili : İngilizce

Eğitim Durumu

Lise : Keçiören Lisesi (2005)
Lisans : Gazi Üniversitesi Fen Edebiyat Fakültesi
İstatistik Bölümü (2009)
Yüksek Lisans : Gazi Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı (2012)
Doktora : Ankara Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı (Eylül 2012-Mart 2018)

Çalıştığı Kurum

Afyon Kocatepe Üniversitesi / Fen-Edebiyat Fakültesi / İstatistik Bölümü /
İstatistiksel Bilgi Sistemleri Anabilim Dalı (2011-Devam Ediyor)

SCI-Exp

1. **Özdemir Şenay** and Arslan Olcay (2018). An Alternative Algorithm of the Empirical Likelihood Estimation for the Parameter of a Linear Regression Model. Communications in Statistics - Simulation and Computation. Doi: 10.1080/03610918.2018.1435801.
2. Ebegil Meral, **Özdemir Şenay** (2016). Two Different Shrinkage Estimator Classes for the Shape Parameter of Classical Pareto Distribution. Hacettepe Journal of Mathematics and Statistics, 45(4), 1231-1244 Doi: 10.15672/HJMS.20158812908

ESCI

1. **Özdemir Şenay**, Güney Yeşim, Tuğ Yetkin, and Arslan Olcay (2019). Maximum LQ-Likelihood Estimation for the Parameters of Marshall-Olkin Extended Burr XII Distribution. Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat., 68. (1), 17-34 DOI: 10.1501/Commua1_0000000887

Hakemli Dergiler

1. Saraçlı Sinan,Kılıç İbrahim,Türkan Ayça Hatice,Dursun Pınar,Doğan İlkay, **Özdemir Şenay**, Gazeloğlu Cengiz (2014). Üniversite Öğrencilerinin Çevresel Bilgi ve Davranışlarının Yapısal Eşitlik Modeli ile İncelenmesi. International Anatolian Academic Online Journal, Scientific Science, 2(1), 1-11.
2. Karakaş Özge, Türkan Ayça Hatice, **Özdemir Şenay** (2013). Fen Edebiyat Ve Eğitim Fakültesi Öğrencilerinin Dil Kirliliğine Duyarlılığı Üzerine Bir Araştırma Afyonkarahisar Örneği. Turkish Studies - International Periodical For The Languages, Literature and History of Turkish or Turkic, 8(4), 927-937.
3. **Özdemir Şenay**, Ebeğil Meral (2012). Klasik Pareto Dağılımının Şekil Parametresi İçin Shrinkage Tahmin Ediciler. Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü Dergisi, 16(1), 116-121.

Uluslararası Kongre Sunum

1. **Özdemir Şenay**, Arslan Olcay (2017). Empirical Likelihood Estimation of Linear Regression Model with Robust Constraint, 18. Uluslararası Ekonometri, Yöneylem Araştırması Ve İstatistik Sempozyumu.(Özet Bildiri / Sözlü sunum)
2. Güney Yeşim,Tuaç Yetkin,**Özdemir Şenay**,Arslan Olcay (2017). Maximum Lq-likelihood estimation for autoregressive error term regression model. 31st EMS (The European Meeting of Statisticians) (Özet Bildiri/Sözlü Sunum)
3. Tuaç Yetkin,Güney Yeşim,**Özdemir Şenay**,Arslan Olcay (2017). Robust variable selection in an autoregressive error term regression model. 31st EMS (The European Meeting of Statisticians) (Özet Bildiri/Sözlü Sunum)
4. **Özdemir Şenay**,Güney Yeşim,Tuaç Yetkin,Arslan Olcay (2017). Empirical likelihood estimation for linear regression models with AR(p) error terms. 31st EMS (The European Meeting of Statisticians) (Özet Bildiri/Sözlü Sunum)
5. Tuaç Yetkin,Güney Yeşim,**Özdemir Şenay**,Arslan Olcay (2017). Parameter estimation for skew normal distribution via the maximum Lq-likelihood estimation. 3RD International Researchers, Statisticians and Young Statisticians Congress (Özet Bildiri/Sözlü Sunum)
6. Güney Yeşim,**Özdemir Şenay**,Tuaç Yetkin,Arslan Olcay (2017). Optimal B-robust estimators for the parameters of the MOEBXII distribution. 3RD International Researchers, Statisticians and Young Statisticians Congress (Özet Bildiri/Sözlü Sunum)
7. **Özdemir Şenay**,Tuaç Yetkin,Güney Yeşim,Arslan Olcay (2017). Maximum Lq-Likelihood Estimation for the Marshall-Olkin Extended Burr XII Distribution. 3RD International Researchers, Statisticians and Young Statisticians Congress (Özet Bildiri/Sözlü Sunum)
8. **Özdemir Şenay**,Güney Yeşim Arslan Olcay (2016). Variable Selection in Empirical Likelihood based Regression with Information Complexity (ICOMP). International Conference on Information Complexity and Statistical Modeling in High Dimensions with Applications (IC-SMHD-2016) (Özet Bildiri/Sözlü Sunum)
9. **Özdemir Şenay**, Ebeğil Meral (2012). A Shrinkage Estimator for the Scale Parameter of Laplace Distribution. 8th International Symposium Of Statistics (Özet Bildiri/Sözlü Sunum)