



**DEFORME EDİLMİŞ LOGARİTMALAR İLE
GENELLEŞTİRİLMİŞ RİDGE REGRESYON**

Meryem GÖKTÜRK ŞAP

Yüksek Lisans Tezi

İstatistik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Esra Pamukçu

İkinci Danışman: Dr. Öğr. Üyesi M. Niyazi ÇANKAYA

Eylül-2019

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DEFORME EDİLMİŞ LOGARİTMALAR İLE GENELLEŞTİRİLMİŞ RIDGE REGRESYON

YÜKSEK LİSANS TEZİ

Meryem GÖKTÜRK ŞAP

171133104

Tezin Enstitüye Verildiği Tarih : 20 Ağustos 2019

Tezin Savunulduğu Tarih : 05 Eylül 2019

Tez Danışmanı : Dr. Öğr. Üyesi Esra PAMUKÇU (F.Ü)

İkinci Danışman: Dr. Öğr. Üyesi Mehmet Niyazi ÇANKAYA(U.Ü)

Diğer Jüri Üyeleri : Prof. Dr. Sinan ÇALIK (F.Ü)

Doç. Dr. Ayşe Turan BUĞATEKİN (F.Ü)

Dr. Öğr. Üyesi Arzu EKİNCİ DEMİRELLİ

Eylül-2019

ÖNSÖZ

Bu çalışmanın oluşmasında ilgi ve yardımlarını esirgemeyen, özel olarak vakitlerini ayırıp sorularımı cevapsız bırakmayan başta danışman hocalarım Fırat Üniversitesi'nden sayın Dr. Öğr. Üyesi Esra PAMUKÇU ve Uşak Üniversitesi'nden sayın Dr. Öğr. Üyesi Mehmet Niyazi ÇANKAYA olmak üzere, bölüm başkanımız sayın Prof. Dr. Sinan ÇALIK'a ve diğer tüm bölüm hocalarıma teşekkürü bir borç bilirim.

Meryem GÖKTÜRKŞAP
ELAZIĞ-2019

İÇİNDEKİLER

Sayfa No:

ÖNSÖZ	II
İÇİNDEKİLER	III
ÖZET	V
SUMMARY	VI
ŞEKİLLER LİSTESİ	VII
TABLOLAR LİSTESİ	VIII
SEMBOLLER LİSTESİ.....	IX
KISALTMALAR	X
1.GİRİŞ.....	1
2.MATERYAL VE YÖNTEM	4
2.1. En Çok Olabilirlik (EÇO) Tahmin Yöntemi	4
2.2 En Çok Olabilirlik Tahmin Yöntemi ile En Çok Olabilirlik logq ve logk Tahmin Yöntemi Arasındaki Anlamsal Geçiş.....	5
2.3 En Çok Olabilirlik logq ve logk Tahmin Yöntemi ile M-Tahmin Yöntemi Arasındaki Anlamsal Geçiş.....	5
2.4. Basit Doğrusal Regresyon Modeli, Bazı Eğrisel Formlar ve Matris Gösterimi .	6
2.5 En küçük Kareler Yöntemi	7
2.6 En Küçük Kareler Yöntemi hesaplama adımları	8
2.7 Dağılım Bilgisi ve Dağılım Bilgisi Üzerinden Regresyon Denkleminin Yazımı ve Özellikleri.....	11
2.8 Student t-dağılımının Konvekslik ve Konkavlık Özelliğinin İncelenmesi	13
2.8.1 Student t Dağılımı Kullanılarak Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri	13
2.9 Deforme Edilmiş Logaritmalar ve Özellikleri	14
2.10 M-tahmin Yöntemi.....	15
2.11 Regresyon Yapısında M-tahmin Edicileri.....	16
2.12 M-tahmin Yönteminde Kullanılan Fonksiyonun Özelliği	18

2.13 Deforme Edilmiş Logaritmalar Aracılığı ile En Çok Olabilirlik Tahmin Edicileri	19
2.14 Kaniakadis Logaritması logk Özelliğinin İncelenmesi	20
2.15 Dağılım Bilgisi Altında Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri	20
2.16 M-Tahmin Edicileri: En Çok logq Tahmin Edicileri	22
2.17 M-Tahmin Edicileri: En Çok logk Tahmin Edicileri	24
2.18 Genelleştirilmiş Ridge Regresyon.....	25
2.19 Ridge Regresyon Yapısında M-tahmin Edicileri	26
2.20 Dağılım Bilgisi Altında Ridge Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri.....	28
2.21 MR-Tahmin Edicileri: Ridge En Çok logq Tahmin Edicileri	29
2.22 MR-Tahmin Edicileri: Ridge En Çok logk Tahmin Edicileri	30
2.24 M-Tahmin Edicisinin Sayısal Hesaplaması	32
2.24.1 Ridge M-Tahmin Edicisinin Sayısal Hesaplaması	33
2.25 Çoklu İç İlişki Durumu	34
2.25.1 Çoklu İç İlişkinin Ortadan Kaldırılmasına İlişkin Çözüm Yöntemleri	34
2.26 Bilgi Kriterinin Kullanımı.....	35
3. UYGULAMA ve BULGULAR.....	37
3.1 Gerçek Veri 1	37
3.2 Gerçek Veri 2	44
4. SONUÇLAR VE TARTIŞMA.....	53
5. ÖNERİLER.....	54
KAYNAKLAR	55
EKLER	60
ÖZGEÇMİŞ	69

ÖZET

Deforme Edilmiş Logaritmalar ile Genelleştirilmiş Ridge Regresyon

Ridge regresyon, regresyon analizinde bağımsız değişkenler arasında çoklu bağlantı problemi olduğu zaman önerilen alternatif bir yaklaşımdır. Bu araştırmanın amacı, en çok olabilirlik tahmin edicilerinin bir genelleştirmesi olan M-tahmin edicilerini deforme edilmiş logaritmalar ile yeniden tanımlamak ve bu tahmin ediciler yardımıyla ridge regresyon analizi yapabilmektir. Böylelikle hem çoklu bağlantı probleminin üstesinden gelinecek, hem de deforme edilmiş logaritmalar temelinde oluşturulan M-tahmin edicileri ile dayanıklı tahminler elde edilecektir. Önerilen yöntemin etkinliğini ve uygulanabilirliğini göstermek amacı ile gerçek veri setleri üzerinde uygulama yapılmıştır. Deforme edilmiş logaritmalar temelli M-tahmin yöntemiyle uygulanan ridge regresyonda, ridge parametresinin önemi tespit edilmiş ve Huber M-tahmin ve Student t dağılımının regresyon parametrelerinin en çok olabilirlik tahmin edicileri ile karşılaştırmaları yapılmıştır.

Anahtar Kelimeler: Deforme Edilmiş Logaritmalar, En Çok Olabilirlik, Çoklu Bağlantı Problemi, Ridge Regresyon

SUMMARY

Generalized Ridge Regression By Using Deformed Logarithms

Ridge regression is an alternative approach in regression analysis when there is multicollinearity problem between independent variables. The aim of this study is to define M-estimators which is a generalization of the likelihood estimators by using deformed logarithms and to perform ridge regression analysis with M-estimators based on deformed logarithms. In this way, both multicollinearity problem will be overcome and robust estimations will be obtained.. In order to demonstrate the effectiveness and applicability of the proposed method, two applications was performed on the real data sets. The importance of ridge parameter was determined in ridge regression using M-estimation method based on deformed logarithms and comparisons of ridge Huber M-estimation and Student t distribution regression parameters with maximum likelihood estimators were made.

Key words: Deformed Logarithms, Maximum Likelihood, Multicollinearity Problem, Ridge Regression

ŞEKİLLER LİSTESİ

	<u>Sayfa No</u>
Şekil 3.1. Veri setinin normal dağılıma uyum sağlayıp sağlamadığının uygunluğun P-P grafiği ile incelenmesi	38
Şekil 3.2. x_1, x_2 ve x_3 açıklayıcı değişkenleri arasında doğrusal ilişkinin varlığının gösterimi ve her bir değişkenin histogramı	39
Şekil 3.3. x_4 ve x_5 açıklayıcı değişkenleri arasında doğrusal ilişkinin varlığının gösterimi ve her bir değişkenin histogramı	39
Şekil 3.4. Veri setinin normal dağılıma uyum sağlayıp sağlamadığının uygunluğun P-P grafiği ile incelenmesi	45
Şekil 3.5. Baş, boy ve karın çevresi değişkenleri arasındaki korelasyon ve histogramları	47
Şekil 3.6. Glukoz, HDL ve LDL değişkenleri arasındaki korelasyon ve histogramları.	47
Şekil 3.7. Üre, Kreatinin ve Potasyum değişkenleri arasındaki korelasyon ve histogramları	48
Şekil 3.8. HGB, HCT, Diastolik ve Sistolik değişkenleri arasındaki korelasyon ve histogramları	48

TABLÖLER LİSTESİ

Sayfa No

Tablo 3.1. Veri setini modellemede kullanılan amaç fonksiyonları.....	37
Tablo 3. 2. $X_1, X_2, \dots, X_5$ ve Y değişkenleri arasındaki korelasyon matrisi	40
Tablo 3.3. Dayanıksız M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa=10^{-20}$ t: $\nu=100$, Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=15$	40
Tablo 3.4. Dayanıksız M-tahmin değerleri ve $\lambda = 0.05$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k = 4$, $n = 15$	41
Tablo 3.5. Dayanıksız M-tahmin değerleri ve $\lambda = 0.1$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k = 4$, $n = 15$	41
Tablo 3.6. Dayanıksız M-tahmin değerleri ve $\lambda = 0$ iken Kaniadakis: $\kappa = 0.1$, t: $\nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$	42
Tablo 3.7. Dayanıksız M-tahmin değerleri ve $\lambda = 0.05$ iken Kaniadakis: $\kappa = 0.1$, t: $\nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$	43
Tablo 3.8. Dayanıksız M-tahmin değerleri ve $\lambda = 0.1$ iken Kaniadakis: $\kappa = 0.1$, t: $\nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$	43
Tablo 3. 9. Regresyon modelinde kullanılan değişkenler ve her birinin ilgili birimleri ..	44
Tablo 3.10. Yaş, Boy, Karın çevresi, Glukoz, HDL, LDL, Üre, Kreatinin, Potasyum, HGB, HCT, Diastolik ve Sistolik değişkenleri arasındaki korelasyon matrisi 46	
Tablo 3.11. Dayanıksız M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=128$	49
Tablo 3.12. Dayanıksız M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa = 10^{-4}$ t: $\nu = 1.5$, Tsallis: $q=1-10^{-4}$, Huber: $k=1.5$, $n=128$	50
Tablo 3.13. Dayanıksız M-tahmin değerleri ve $\lambda=0.1$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=128$	51
Tablo 3.14. Dayanıksız M-tahmin değerleri ve $\lambda=0.1$ iken Kaniadakis: $\kappa = 10^{-4}$ t: $\nu = 1.5$, Tsallis: $q=1-10^{-4}$, Huber: $k=1.5$, $n=128$	52

SEMBOLLER LİSTESİ

$E()$	: Beklenen değer
e_i	: IRA'da kabul edilen tolerans hata değeri
IRA	: İteratif tekrardan ağırlıklandırılmış algoritması
p	: Parametre sayısı
k	:Huber tuning (ayarlama) katsayısı parametresi
Kov / Cov	: Kovaryans
Kor / Cor	: Korelasyon
n	: Gözlem sayısı
N()	: Normal dağılım
t()	: Student t dağılımı
Var	: Varyans
$\hat{\beta}$	: En küçük kareler tahmin edicisi
$\hat{\beta}_M$	:M-tahmin edicisi
$\hat{\beta}_{MR}$	:Ridge M-tahmin edicisi
${}_q\hat{\beta}_{MR}$	: q-tabanında Ridge M-tahmin edicisi
${}_{\kappa}\hat{\beta}_{MR}$	: κ -tabanında Ridge M-tahmin edicisi
ϵ_i	:Artık (anakütle)
λ	: Ridge parametresi

KISALTMALAR

EKK	: En Küçük Kareler
MATLAB	: Matrix Laboratory (Matrix Labaruatoru)
HKO	:Hata Kareler Ortalaması
HDL	:High Density Lipoprotein(Yüksek Yoğunluklu Lipoprotein)
LDL	:Low Density Lipoprotein (Düşük Yoğunluklu Lipoprotein)
HGB	:Hemoglobin
HCT	:Hematokrit
EÇO	:En çok olabilirlik

1. GİRİŞ

Bir veri setini modelleme de bir parametrik model öne sürülmektedir. Diğer bir ifadeyle, verilerin de bu dağılımdan geldiği varsayılmaktadır. Bu verilerin her biri özünde bir rastgele değişkenin üyesi olup, aynı dağılımlılık ilkesine göre çalışmaktadır. Yani her bir rastgele değişkenler benzer dağılımlı bir olasılık yoğunluk $f(x; \theta)$ veya olasılık $P(x; \theta)$ fonksiyonundan gelmektedir (Öztürk vd., 2015).

Benzer dağılımlılık rastgele gözlemleri $x_1, x_2, \dots, x_n$ 'nin tek bir olasılık yoğunluk fonksiyonundan gelmesidir. Benzer dağılımlılık varsayımı sağlanmadığında, dayanıklı tahmin ediciler elde edilemez. Benzer dağılımlı olmaması durumuna ilişkin örnek olarak, sapan (outlier) gözlem verilebilir. Sapan gözlem yapısı, bir gözlem ya da gözlem topluluğunun $(x_i, x_j \mid i, j = 1, 2, \dots, n, i \neq j)$ dağılımın büyük bir çoğunluğundan (underlying) uzak olması sonucunda oluşmaktadır.

Regresyon parametrelerinin dayanıklı ve yansız tahmini, çıkarım süreci için önemlidir. Literatürde iyi bilinen en çok olabilirlik tahmin edicileri asimptotik etkin, yansız ve tutarlı tahmin edicilerdir. Dayanıklı tahmin yöntemi ise son zamanlarda büyük ilerlemeler kat eden ve sık kullanılan bir yöntemdir. Bu tahmin yöntemi, özellikle konum modeli ve aynı zamanda regresyon modeli için önerilmiştir. En çok olabilirlik tahmin yönteminin bir genelleştirilmesi olan M-tahmin edicileri ve M-tahmin edicileri sınıfında yer alan en küçük medyan kareler ve en küçük kırılmış kareler tahmincileri ile sıralı istatistiklerin doğrusal kombinasyonundan oluşan L-tahmin edicileri ve dağılım parametrelerinin tahmini için kullanılan klasik moment tahmincisiinden esinlenerek elde edilen kırılmış moment tahmincisi farklı dayanıklı tahmin etme yöntemleri olarak karşımıza çıkmaktadır (Huber, 1964; Huber, 1973; Rousseeuw, 1984; Bajorunaite ve Brazauskas, 2008).

Tahmin edicilerin dayanıklı olmasının yanı sıra etkin olması da önemlidir. Eğer gerçek veri seti benzer dağılımlı ise, en çok olabilirlik tahmin yöntemi veya yukarıda da söz edilen diğer farklı tahmin yöntemleri kullanılarak ilgili olasılık yoğunluk fonksiyonu $f(x; \theta)$ 'in parametreleri tahmin edilebilir. Eğer benzer dağılımlılık durumu yok ve veride sapan gözlem var ise; en çok olabilirliğin bir genelleştirmesi olan

esasinda en çok olabilirlik gibi davranan M-tahmin edicileri (M-fonksiyonları) önerilebilir (Huber, 1964;Hampel vd., 1986). Literatürde kullanılan M-fonksiyonları genellikle, regresyon veya konum parametrelerinin tahmini için düşünülmüş veya uygulana gelmiştir. Herhangi bir ilgili $f(x; \theta)$ olasılık yoğunluk fonksiyonunun θ parametrelerini dayanıklı tahmin etmek için son zamanlarda da çalışmalar yer almaktadır (Çankaya ve Korbel, 2018).

Bir olgu, ekonomik, iktisadi veya tıbbi bir olay olabilir, birden fazla sebebin ya da açıklayıcının yer aldığı ortak bir sonuç olarak ortaya çıkabilir. İlgilenilen olgunun değişkenleri arasında bir ilişkinin varlığı düşünülebilir ve böylesi bir ilişki var ise, bir matematiksel model aracılığı ile regresyon modeli şeklinde yazılabilir. Böylesi bir regresyon modelinde, açıklayıcı değişkenler ya da bağımsız değişkenler X ile bu değişkenlerin neyi açıkladığını ya da bağımlı değişkenin ne olduğunu gösteren Y değişkeni vardır. Böylece, biri bağımlı Y ; diğeri bağımsız X olmak üzere iki tip değişken bulunur. Bu iki tip değişken arasında verilerin varsayılan bir matematiksel modelden geldiği düşünülerek

$$Y = X\beta + \varepsilon \quad (1.1)$$

şeklinde bir form ele alınabilir. Bu formda, rastgele değişken $\varepsilon \in R$ ilgili matematiksel modelin etrafındaki gözlemleri temsil etmeyi sağlar.

Doğrusal regresyon modeli

$$y_i = X\beta + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.2)$$

şeklinde yazılabilir. Denklem 1.2'deki ε_i , $E[\varepsilon|x] = 0$ ve $V[\varepsilon|x] = \sigma^2 I$ olan birbirinden bağımsız rastgele hataların vektörüdür. Denklem 1.1 veya 1.2'deki doğrusal regresyon modelinin katsayılarını tahmin etmek için en çok kullanılan yöntem en küçük kareler (EKK) yöntemidir. Ancak bu yöntemin uygulanabilmesi için çoklu regresyon varsayımlarının sağlanması gerekir. Bu varsayımlardan biri, çoklu regresyon modelinde açıklayıcı değişkenlerin ilişkisiz olduğu varsayımdır. Ancak bu her zaman doğru olmayabilir. Açıklayıcı değişkenler ilişkili olduğunda parametre tahminleri yapmak için EKK tahmin edicisine alternatif birçok tahmin edici geliştirilmiştir (Montgomery, 1991).

Deforme edilmiş logaritmlar fizikçiler tarafından önerilmiştir (Tsallis, 1988;Kaniadakis, 2005). İstatistikte de son zamanlarda önemli bir yere sahiptir (Çankaya, ve Korbel, 2018;Ferrari ve Paterlini, 2009). Bu çalışmada, deforme edilmiş logaritmların regresyon analizinde uygulamaları üzerinde durulacaktır. Bu amaçla, deforme edilmiş logaritmlara bağlı yeni regresyon katsayılarının (β) tahmin edicileri elde edilecektir. Böylelikle, hem çoklu iç ilişkiye hem de sapmalara (outlier) karşı dayanıklı regresyon katsayısı tahmin edicilerinin elde edilmesi hedeflenmektedir.

Bu çalışmada, öncelikle dayanıklı (robust) tahmin yöntemi olan M-tahmin yöntemi üzerinde durulacaktır. M-tahmin yöntemlerinden Huber M-tahmin edicisi kullanılacaktır. En çok olabilirlik tahmin yönteminin bir genelleştirmesi olan M-tahmin yöntemi aynı zamanda literatürde incelenmiş olan en çok loq_q olabilirlik fonksiyonuna değinilecektir. loq_q fonksiyonu deforme edilmiş logaritmlar içinde yer almaktadır ve Tsallis tarafından önerilen konkav bir fonksiyon olup tahmin yönteminde kullanılabilmeyi sağlamaktadır. Benzer şekilde, literatürde farklı deforme edilmiş logaritmlar önerilmiştir. Monotonik özelliğini sağlayan Kaniakadis loq_k fonksiyonu da kullanılmıştır. Bu fonksiyonlar, en çok olabilirlik tahmin yönteminde yer alan logaritma fonksiyonu yerine kullanılmıştır. Ayrıca, çalışmada bu M-tahmin yöntemi kullanılarak ridge parametresi λ ile gerekli olan kısıtlar altında ridge M-tahmin edicileri ($\widehat{\beta}_{MR}$) elde edilmiştir.

Sonuç olarak, tez iki ana bölümden oluşmaktadır. İlk olarak dayanıklı M-tahmin edicileri ve bunların deforme edilmiş logaritmlar üzerinden gösterilmesidir. İkinci olarak, en çok olabilirlik tahmin yöntemine kısıt konularak elde edilen ridge dayanıklı M-tahmin edicileri, literatürde ilk kez deforme edilmiş logaritmlar olan loq_q ve loq_k fonksiyonları üzerinden önerilecektir. M-tahmin edicisi olan Huber M-tahmin edicisi ile Student t dağılımının regresyon (konum) parametrelerinin EÇO tahmin edicileri ile karşılaştırılması yapılacaktır. Çalışmanın uygulama bölümünde, $\widehat{\beta}_M$ M-tahmin edicileri Huber M, Student t, loq_q ve loq_k fonksiyonları kullanılarak, ilgili fonksiyonların sayısal örnekler üzerinde verdiği değerleri karşılaştırılacaktır. Aynı zamanda ridge parametresinin de yer aldığı durumdaki $\widehat{\beta}_{MR}$ MR-tahmin edicileri sayısal örnekler üzerinden hesaplanıp karşılaştırılacaktır.

2. MATERYAL VE YÖNTEM

2.1. En Çok Olabilirlik (EÇO) Tahmin Yöntemi

EÇO tahmin yöntemi, asimptotik özellikleri gereği parametre tahmininde kullanılan en popüler yöntemlerden biridir. En çok olabilirlik tahmin edicileri, olabilirlik fonksiyonunu maksimum yapan değerlerdir.

$X = (x_1, x_2, \dots, x_n)$ rastgele gözlem vektörü, olasılık veya olasılık fonksiyonu $f(x; \theta)$ olan bir dağılımdan çekilen bir örneklem olsun. Bu durumda, $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ parametre vektörünün olabilirlik fonksiyonu,

$$L(\theta; x) = \prod_{i=1}^n f(x_i; \theta) = f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta), \theta \in R^p \quad (2.1)$$

şeklinde tanımlanır.

L olabilirlik fonksiyonu ve bağımsız aynı dağılımlı $x_1, x_2, \dots, x_n$ rastgele değişkenlerinin sahip olduğu $f(x; \theta)$ olasılık veya olasılık yoğunluk fonksiyonlarının çarpımı şeklindedir. L fonksiyonu, θ parametresine göre maksimum yapılır ve böylelikle en çok olabilirlik tahmin yöntemi adını alır. Bazen θ parametresini x_i 'lerin bir fonksiyonu olarak elde etmek yani analitik adımların işlenebilmesini sağlamak için, L olabilirlik fonksiyonunun logaritması alınır ve $\log(L)$ şeklinde gösterilir. Bu durumda, denklem (2.1)

$$\begin{aligned} \log(L(\theta; x)) &= \sum_{i=1}^n \log(f(x_i; \theta)) \\ &= \log f(x_1; \theta) + \log f(x_2; \theta) + \dots + \log f(x_n; \theta) \end{aligned} \quad (2.2)$$

şeklinde yazılır. $\log$ yerine $\log_q$ alınabileceği önerilmiştir (Ferrari ve Yang, 2010). $\log$ fonksiyonu konkav bir fonksiyon olup, en çok olabilirlikte de kullanabilmeyi sağlamaktadır ve monoton olma özelliğini de sağlamaktadır (Öztürk vd. 2015) $\log_q$ fonksiyonu da konkav bir fonksiyondur (Çankaya ve Korbel, 2018) Bilinmektedir ki; en

çok olabilirlik tahmincileri, regülerlik koşulları altında, asimptotik yansız ve tutarlı tahmincilerdir (Öztürk vd. 2015).

M-tahmin edicileri en çok olabilirlik (EÇO) tahmin yönteminin bir genelleştirilmesi olan ve en çok olabilirlik yöntemini taklit eden bir yöntemdir (Huber, 1964). M-tahmin yöntemi dayanıklı tahmin ediciler elde edilmesini sağlamaktadır. Eğer veri setinde bir sapan gözlem yani belli bir yığının dışında yer alan bir gözlem veya gözlemler var ise, bu durumda yığının modellenmesi ve/veya yığın parametrelerinin tahmini sürecinde bu sapan gözlemlerden etkilenmemesini sağlanmak için M-tahmin edicilerindeki M-fonksiyonları olarak ta isimlendirilebilen amaç fonksiyonları kullanılmaktadır.

2.2 En Çok Olabilirlik Tahmin Yöntemi ile En Çok Olabilirlik $\log_q$ ve $\log_\kappa$ Tahmin Yöntemi Arasındaki Anlamsal Geçiş

Eşitlik (2.2)'deki $\log(L)$ yerine $\log_q L$ yazılabilir ve böylelikle de

$$\log_q(L(\theta; x)) = \sum_{i=1}^n \log_q(f(x_i; \theta)) \quad (2.3)$$

şeklindeki fonksiyon elde edilir. Buna ise en çok olabilirlik $\log_q$ tahmin yöntemi denir (Ferrari ve Yang, 2010).

2.3 En Çok Olabilirlik $\log_q$ ve $\log_\kappa$ Tahmin Yöntemi ile M-Tahmin Yöntemi Arasındaki Anlamsal Geçiş

Eşitlikler (2.2) ve (2.3) karşılaştırıldığında; $\log$ ve $\log_q$ farklılıklarının olduğu gözlenmektedir. Böylelikle, $\log$ fonksiyonu yerine, EÇO' nun genelleştirilmiş bir formu kullanılmaktadır. Bu durumda, (2.2) eşitliği

$$-\log_q(L(\theta; x)) = -\sum_{i=1}^n \log_q(f(x; \theta)) \quad (2.4)$$

şeklinde de yazılabilir. Belirtmek gerekir ki, (2.3) eşitliğinin θ parametresine göre maksimum yapılması ile (2.4) eşitliğinin ilgili parametreye göre minimum yapılması denk ifadelerdir. (2.4) eşitliğini ele almış olmak, M-tahmin edicilerini ele almaya

karşılık gelecektir (Huber, 1964). Dolayısıyla, (2.4) eşitliği, θ parametresine göre minimum yapılarak $\hat{\theta}$ M-tahmin edicileri elde edilecektir (Ferrari ve Yang, 2010).

Benzer şekilde, κ –tabanında M-tahmin yöntemi

$$-\log_{\kappa}(L(\theta; x)) = -\sum_{i=1}^n \log_{\kappa}(f(x; \theta)) \quad (2.5)$$

şeklinde ifade edilir ve $\hat{\theta}$ M-tahmin edicileri elde edilir (Kaniadakis vd., 2005).

2.4. Basit Doğrusal Regresyon Modeli, Bazı Eğrisel Formlar ve Matris Gösterimi

$X = (x_1, x_2, \dots, x_n)$ ve $Y = (y_1, y_2, \dots, y_n)$ serileri verilmiş olsun. Bunlar arasındaki ilişki doğrusal bir yapı gösterdiğinde, aşağıda verilen basit regresyon modeli ortaya çıkar:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.6)$$

Denklem 2.6'da β_0 ve β_1 regresyon modelinin parametreleri olup, β_0 , y – eksenini kestiği noktadır, β_1 ise doğru denkleminin eğimidir. Eğer katsayı değeri pozitif ise, artan bir doğrudur. Negatif ise de azalan bir doğrudur. i her bir gözlemi gösteren indis ve n örnek sayısı veya gözlem sayısıdır.

$(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$ olmak üzere, n tane gözlem çifti gerçek dünyadaki herhangi bir olgudan geldikten sonra eşitlik 2.6'daki bir regresyon modeli ile oluşturulabileceği gibi ilişkinin şekline göre uygun seçilen bir eğrisel model ile de oluşturulabilir. Aşağıda eğrisel ilişkiye ait örnekler verilmektedir:

$$y_i = \beta_0 + \beta_1 x_i^2 + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.7)$$

ya da

$$y = \beta_0 + \beta_1 x_{1i}^2 + \beta_2 x_{2i}^2 + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.8)$$

şekillerinde olabileceği gibi üstel, trigonometrik ve başka formlar da olabilir. Serpilme diyagramındaki davranışa göre en uygun matematiksel model seçilmektedir.

Eğer saçılım doğrusal ise en genel ifade ile,

$$Y = X\beta + \varepsilon \quad (2.9)$$

regresyon modeli yazılabilir. Burada X vektörü açıklayıcı değişkenlerin yer aldığı $x_1, x_2, \dots, x_p$ açıklayıcı değişkenlerini, $\beta = \beta_1, \beta_2, \dots, \beta_p$ ise ilgili değişkenlerin önündeki katsayıları gösterir. Belirtmek gerekir ki, β katsayıları ile X vektörünün eşitlik 2.9'da verildiği şekilde yazılabilmesi için ilgili matris yazımına izin verilmesi gereklidir (Monthomery ve Peck, 1991; Marquardt, 1970).

2.5 En Küçük Kareler Yöntemi

Regresyon denkleminde parametrelerine göre lineerlik var ise, regresyon katsayılarının $\beta_1, \beta_2, \dots, \beta_p$ tahmini için en popüler yöntem En Küçük Kareler (EKK) yöntemidir. Regresyon doğrusunun ve çoklu regresyon için regresyon düzleminin gözlem değerlerini iyi bir şekilde temsil edebilmesi için, bu gözlem noktalarını tam olarak ortalaması gerekir. Bu şekilde, ε_i artıkları minimize edilmiş olacaktır. Bunun için EKK yönteminde gerçek ilişkiye bir terim olarak eklenen ε değişkeni hakkında şu varsayımlar geçerlidir:

1. ε bir rastgele değişkendir ve $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ ilgili rastgele değişkenden gelen gözlemlerdir.
2. ε rastgele değişkeninin beklenen değeri sıfırdır. $E(\varepsilon_i) = 0$
3. ε rastgele değişkeni her X seviyesinde varyansı sabittir. $Var(\varepsilon_i/X = x_i) = \sigma^2$
4. ε rastgele değişkeni normal dağılıma sahiptir. $\varepsilon_i \sim N(0, \sigma^2)$
5. ε rastgele değişkeninin farklı terimleri olan $\varepsilon_i, \varepsilon_j$ $i, j = 1, 2, \dots, n$ ve $i \neq j$ olmak üzere arasındaki doğrusal ilişkiyi gösteren korelasyon sıfırdır. $Kor(\varepsilon_i, \varepsilon_j) = 0$
6. Eğer $Kor(\varepsilon_i, \varepsilon_j) = 0$ ise, ε rastgele değişkeni ve açıklayıcı değişkenleri $X = (x_1, x_2, \dots, x_p)$ arasında doğrusal ilişki mevcut değildir, kovaryans $Kov(\varepsilon, X) = 0$ ya da korelasyon $Kor(\varepsilon, X) = 0$.

Bu şartlar altında, artık kareleri toplamını minimize eden $\hat{\beta}_0$ ve $\hat{\beta}_1$ değerleri tespit edilerek regresyon katsayılarının bulunması EKK yöntemi olarak bilinmektedir (Monthomery ve Peck, 1991; Marquardt, 1970)

2.6 En Küçük Kareler Yöntemi Hesaplama Adımları

Regresyon modelinin belirlenmesi için serpilme diyagramı incelendiğinde; doğrusal bir eğilim görülüyorsa, X 'in Y 'ye göre matematik modelinin doğrusal olduğuna kesin olmasa da karar verilebilir. Ancak, gözlem noktaları arasından çok sayıda doğru geçebilir. Bu doğrulardan en uygunu, tüm doğrusal fonksiyonlar arasından Y gözlem değerine en yakın tahmini $\hat{Y}$ değerini minimum hata ile veren doğrusal fonksiyon olur. Yani;

$$\varepsilon = Y - \hat{Y} = Y - \beta_0 - \beta_1 X \quad (2.10)$$

minimum olacak şekilde bir fonksiyon araştırmacı tarafından seçilmelidir. Tüm gözlem noktaları için bu durum geçerli olacağına göre,

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \arg \min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.11)$$

olmalıdır. $\beta = (\beta_1, \beta_2)$. Bu ifadeyi minimum yapacak şekilde, çözüm yapılarak parametre kestirimlerinin elde edilmesine EKK regresyonu denmektedir. Bu ifadenin minimum olabilmesi için β_0 ve β_1 parametrelerine göre birinci mertebeden kısmi türevleri sıfıra eşit olması gerekir.

$$\frac{\partial}{\partial \beta_0} \sum_{i=1}^n \varepsilon_i^2 = 2 \sum_{i=1}^n (-1)(y_i - \beta_0 - \beta_1 x_i) = - \sum_{i=1}^n y_i + n\beta_0 + \beta_1 \sum_{i=1}^n x_i = 0 \quad (2.12)$$

$$\frac{\partial}{\partial \beta_1} \sum_{i=1}^n \varepsilon_i^2 = 2 \sum_{i=1}^n (-x_i)(y_i - \beta_0 - \beta_1 x_i) = - \sum_{i=1}^n x_i y_i + \sum_{i=1}^n x_i \beta_0 - \quad (2.13)$$

$$\sum_{i=1}^n x_i^2 \beta_1 = 0$$

$\sum x_i$; X değişkenindeki x_i gözlemlerinin, $\sum x_i^2$; X değişkeninin kareler toplamı, $\sum y_i$; Y değişkenindeki y_i gözlemlerinin toplamı, $\sum x_i y_i$; iki değişkenin çarpımlarının toplamı yerine konularak yapılabilir. Bilindiği gibi aritmetik ortalamadan farkların cebirsel toplamı sıfırdır. O halde yukarıdaki denklemlerde X ve Y değerleri yerine bunların aritmetik ortalamadan farklarını yazarak bazı kısaltmalar yapabiliriz. (2.12) ve (2.13) denklemlerinde,

$$\sum_{i=1}^n (y_i - \bar{y}) = 0 \text{ ve } \sum_{i=1}^n (x_i - \bar{x}) = 0 \quad (2.14)$$

ifadeleri yerine yazılabilir. Böylece,

$$\sum_{i=1}^n (y_i - \bar{y}) = n\beta_0 + \beta_1 \sum_{i=1}^n (x_i - \bar{x}) \quad (2.15)$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \beta_1 \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.16)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i \quad (2.17)$$

denklemleri elde edilir.

(2.16) ve (2.17) denklemlerinin çözümleri;

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}}{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}} \quad (2.18)$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (2.19)$$

şeklindedir.

(2.12) ve (2.13) eşitlikleri düzenlendikten sonra, matris gösterimi ile

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

olmak üzere

$$Y = X\beta + \varepsilon \quad (2.20)$$

şeklinde ifade edilir. Bu ifadeler yardımıyla normal denklemler $X'X\hat{\beta} = X'Y$ olarak elde edilir ve basit matris işlemleri ile katsayıların en küçük kareler tahmincisi,

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.21)$$

olarak elde edilir.

Herhangi iki değişken arasında bir ilişki olabilir ve bu ilişkinin varlığı matematiksel bir model ile oluşturulabilir. Doğrusal veya eğrisel bir ilişki önerilen doğrusal veya eğrisel bir denklem ile modellenebilir. Tez kapsamında, basit doğrusal regresyon ve bağımsız değişkenlerin ikiden fazla olduğu durum ile ilgilenilecektir.

Regresyonda bir bağımlı değişken ve bir ya da daha çok bağımsız değişken vardır. Bağımsız değişkenin birden fazla olma durumu ilerleyen bölümlerde çoklu regresyon olarak incelenecektir. Bu bölümde bir bağımlı değişken ve bir bağımsız değişken varken, bu iki değişken arasındaki ilişkiyi belirleyecek modelin nasıl olduğunun ortaya çıkartılması açıklanmaya çalışılacaktır. n tane birimin her birinden bağımlı değişken (Y) ve bağımsız değişken (X) değerleri saptanmış olsun. Bu durumda $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$ olmak üzere n tane gözlem ikilisi yer alır. Y ve X değişkenleri arasındaki ilişki nasıldır? Bu ilişkiyi matematiksel olarak bir regresyon denklemi aracılığı ile modellenebilir mi? Benzer soruların araştırılıp yanıtlarının verilebilmesi için $(y_i, x_i) \quad i = 1, 2, \dots, n$ gözlem ikililerini koordinat eksenlerine işaretleme süreci gerçekleştirilir.

Bu yürütülen sürece, regresyonda **serpme/saçılım diyagramının** oluşturulması denir. n tane gözlem çiftinin her biri için serpme diyagramında kesişim noktaları bulunduğu n tane nokta oluşacaktır. Bu noktaların kartezyen bölgesindeki konumuna bakılarak modelin nasıl olduğuna ilişkin karar verilir. Eğer noktalar bir doğru etrafında toplanıyorsa doğrusal bir model kullanılmalıdır (Montgomery ve Peck, 1991; Orhunbilge, 2002).

O zaman X ve Y arasındaki,

$$\text{Gerçek ilişki} \quad Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.22)$$

$$\text{Gerçek regresyon doğrusu} \quad E(Y_i) = \beta_0 + \beta_1 X_i \quad i = 1, 2, \dots, n \quad (2.23)$$

olur.

Bu şekilde, Y de meydana gelen deęişmeler, deterministik ve rassal deęişmelerin toplam etkisiyle ortaya çıkmaktadır. Bu gerçek ilişki ve gerçek regresyon doğrusu, Y ve X e ait tüm anakütle deęerleri elde edilmedięi sürece bilinemezler. Ancak aynı model, X ve Y anakütlelerinden alınmış bir grup gözlem deęeri (örneklem) üzerinden řu şekilde tanımlanabilir.

$$\text{Tahmin edilen ilişki} \quad \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.24)$$

$$\text{Tahmin edilen regresyon} \quad \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad i = 1, 2, \dots, n \quad (2.25)$$

Burada Y bağımlı deęişkeni ile X bağımsız deęişkeni arasındaki ilişki ortaya konmuştur. Regresyonda $\hat{\beta}_0$ ve $\hat{\beta}_1$, sırasıyla gerçek (ilgilenilen olgunun doğası gereęi sergiledięi durum) ilişkideki β_0 ve β_1 regresyon parametrelerinin tahmin deęerleridir. Öte yandan, regresyondaki ε_i ' ler gerçek ilişkideki ε_i ' lerin tahminidir. ε_i deęerleri regresyon artıkları adını alır. Grafik üzerinde (X_i, Y_i) gözlem deęerlerinin $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ ile gösterilen regresyon doğrusundan sapmalarının nedeni işte bu ε_i artıklarıdır. Yani, $Y_i = \hat{Y}_i + e_i$ ilişkisi vardır. Regresyon doğrusunun üzerinde yer alan gözlem deęerleri için ε_i ' ler pozitif, altında kalan gözlem deęerleri için ε_i ' ler negatif olmakla beraber, ε_i ' lerin cebirsel toplamı sıfırdır. Burada verilen, $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ regresyon denklemi sadece bir bağımsız deęişken içerdięinden, **basit doğrusal regresyon modeli** adını alır. Denklemden görüldüęü gibi, basit regresyonda sayısal deęeri bulunması gereken iki katsayı vardır; $\hat{\beta}_0$ ve $\hat{\beta}_1$. X deęişkeni 0 deęerini aldığıında regresyon doğrusu Y eksenini $\hat{\beta}_0$ noktasında keser. Bu yüzden $\hat{\beta}_0$ katsayısına **kesme terimi** denirken, $\hat{\beta}_1$ katsayısına regresyon doğrusunun eğimini göstermek üzere **eğim katsayısı** adı verilir (Neter vd., 1990).

2.7 Dağılım Bilgisi ve Dağılım Bilgisi Üzerinden Regresyon Denkleminin Yazımı ve Özellikleri

Çoklu doğrusal regresyon modeli için, $i = 1, 2, \dots, n$ için ε_i hata terimlerinin birbirinden bağımsız, sıfır ortalamalı ve σ^2 sabit varyanslı f dağılımının sahip oldukları, başka bir ifadeyle, $\varepsilon \sim f(\mathbf{0}, \sigma^2 \mathbf{I})$ olduęu varsayılın. Rastgele deęişkenler $X_1, X_2, \dots, X_n$, bağımsız ve aynı dağılımlı aşağıda verilen olasılık yoğunluk fonksiyonuna sahip olsun.

Regresyon formatında olasılık yoğunluk fonksiyonları aşağıdaki şekilde yazılır. Normal dağılımın olasılık yoğunluk fonksiyonu

$$f_N(x; \beta, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{(y - X\beta)}{\sigma}\right)^2\right), \quad x \in R, \beta \in R, \sigma > 0 \quad (2.26)$$

şeklinde tanımlıdır.

Uygulamalı sahalarda gerçek veri setlerinin modellenmesinde kalın kuyruklu (x ekseninden uzak) dağılımların, normal dağılıma kıyasla daha elverişli olduğu gösterilmektedir. Normal dağılımın verilerin büyük bir çoğunluğundan uzak olan veri setinin modellenmesinde kullanımı doğru değildir. Normal dağılıma alternatif olabilen benzeri durumlarda normal dağılım kullanılarak hata terimlerini modellemek yerine Student t dağılımı kullanılabilir.

Student t dağılımı ilk olarak Gosset (1908) tarafından ortalamanın hatası üzerine yaptığı çalışmada önerilmiştir. Daha sonra Fisher (1925) ve diğerlerinin katkılarıyla dağılımla ilgili gelişmeler devam etmiştir (Arslan ve Genç 2009).

X rastgele değişkeninin Student t dağılımına sahip olması durumunda,

$$f_t(x; \beta, \sigma) = \frac{c_v}{\sigma} \left(1 + \frac{(y - X\beta)^2}{v\sigma^2}\right)^{-\frac{v+1}{2}}, \quad x \in R, \beta \in R, \sigma > 0, \quad (2.27)$$

şeklinde olasılık yoğunluk fonksiyonu verilmektedir. $c_v = \frac{\Gamma(\frac{v+1}{2})}{(v\pi)^{\frac{1}{2}}\Gamma(\frac{v}{2})}$, $v > 0$ olmak üzere

kuyruk şeklini ayarlayan parametredir. $v \rightarrow 0$ ise t-dağılımının kuyrukları kalın olup, dayanıklı tahminler üretmeyi sağlar. $v \rightarrow \infty$ ise, t-dağılımının kuyrukları ince (x-eksenine yakın) olur ve normal dağılıma yakınsar (Arslan ve Genç, 2009).

Student t dağılımının bazı özellikleri aşağıda verilmiştir:

(i) $v > 1$ için $E(X)=\mu$ ve $v = 1$ için birinci moment $E(X)$ analitik olarak yoktur.

(ii) $v > 2$ için $Var(X) = \frac{v\sigma^2}{v-2}$ ve $v \leq 2$ için $Var(X)$ analitik olarak yoktur.

(iii) $v = 1$ için özel hali Cauchy veya çift Laplace dağılımıdır.

(iv) $v \rightarrow \infty$ iken Student t dağılımı normal dağılıma yakınsar çünkü v parametresi t

dağılımının kuyruk şeklini belirler ve $\lim_{v \rightarrow \infty} \left(1 + \frac{\varepsilon^2}{v\sigma^2}\right)^{-\frac{v-1}{2}} = \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right)$

.(Ahsanullah vd. 2014)

Student t dağılımı kullanılarak elde edilen M-tahmin edicilerinin dayanıklılık özelliği Arslan ve Genç tarafından incelenmiştir. Regresyon durumunda, β parametresinin M-tahmin edicileri önerilmiştir (Arslan ve Genç, 2009 ve Güney, 2018).

2.8 Student t-dağılımının Konvekslik ve Konkavlık Özelliğinin İncelenmesi

Student t dağılımının tek modlu olup olmadığı incelenecektir. Student t dağılımında, x β ifadesi $y - x \beta$ için konuma karşılık gelmektedir. σ ise ölçek parametresi olup, ilgili f_t fonksiyonunda sadece bir ölçeklendirme yapar ve fonksiyonun ana özelliklerini korur. Buradan hareketle,

$$\frac{\partial}{\partial \beta} f_t(y; \beta, \sigma) = -\frac{\left(\frac{y^2+v}{v}\right)^{-\frac{1}{2}v-\frac{1}{2}}(v+1)y}{y^2+v} \quad (2.28)$$

$$\frac{\partial^2}{\partial \beta^T \beta} f_t(y; \beta, \sigma) = \frac{(vy^2+2y^2-v)(v+1)\left(\frac{y^2+v}{v}\right)^{-\frac{1}{2}v-\frac{1}{2}}}{(y^2+v)^2} \quad (2.29)$$

$\frac{\partial^2}{\partial \beta^T \beta} f_t(y; \beta, \sigma)$ ifadesinde, $(vy^2 + 2y^2 - v)$ parçasının negatif ve pozitif olması durumu incelenmelidir.

$$(vy^2 + 2y^2 - v) = \begin{cases} (vy^2 + 2y^2 - v) < 0 & 0 < y < 1 \text{ için} \\ (vy^2 + 2y^2 - v) \geq 0 & y \geq 1 \text{ için} \end{cases}$$

olarak elde edilir. Dolayısıyla, f_t fonksiyonu $0 < y < 1$ aralığı için konkav ve $y \geq 1$ için konvektir. Bu ise, f_t fonksiyonunun tek modlu olduğunu gösterir (Çankaya, 2018).

2.8.1 Student t Dağılımı Kullanılarak Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri

$x_1, x_2, \dots, x_n$ rastgele gözlemleri Student t dağılımına sahip olsun. Regresyon parametrelerinin olabilirlik ve log-olabilirlik fonksiyonları sırasıyla

$$L(\boldsymbol{\beta}|y, x) = (c_v)^n \prod_{i=1}^n \left(1 + \frac{(y_i - x_i \boldsymbol{\beta})^2}{v} \right)^{-\frac{v+1}{2}} \quad (2.30)$$

$$\log L(\boldsymbol{\beta}, y, x) = n \log(c_v) - \frac{v+1}{2} \sum_{i=1}^n \log(v + (y_i - x_i \boldsymbol{\beta})^2)$$

biçimindedir.

Log-olabilirlik fonksiyonunun $\boldsymbol{\beta}$ parametre vektörüne göre türevlerinin alınıp sıfıra eşitlenmesiyle

$$\frac{\partial \log L}{\partial \boldsymbol{\beta}} = v + 1 \sum_{i=1}^n \frac{(y_i - x_i \boldsymbol{\beta}) x_i}{[v + (y_i - x_i^T \boldsymbol{\beta})^2]} = \mathbf{0} \quad (2.31)$$

denklem sistemi elde edilir. Bu eşitliklerin yeniden düzenlenmesiyle $\boldsymbol{\beta}$ parametrelerinin EÇO tahmin edicileri

$$\hat{\boldsymbol{\beta}} = \left[\sum_{i=1}^n x_i^T w_i x_i \right]^{-1} \left[\sum_{i=1}^n x_i^T w_i y_i \right] \quad (2.32)$$

şeklinde elde edilir. Burada w_i ağırlıkları

$$w_i = \frac{v + 1}{[v + (y_i - x_i^T \hat{\boldsymbol{\beta}})^2]} \quad (2.33)$$

biçiminde tanımlıdır (Güney, 2018).

Aykırı değer yani y değerlerinin büyük olması durumunda, Student t-dağılımının ağırlık fonksiyonu gözlemin sayısal büyüklüğüne göre sıfıra yaklaşan bir ağırlık vermektedir. Bu ise, regresyon parametrelerinin tahminlerinin dayanıklı olmasını sağlamaktadır.

2.9 Deforme Edilmiş Logaritmalar ve Özellikleri

Deforme edilmiş bir logaritma olan $\log_q$ Tsallis tarafından öne sürülmüştür ve

$$\log_q(f) = \frac{f^{1-q} - 1}{1 - q}, \quad 0 < q < 1 \quad (2.34)$$

şeklinde tanımlanır ve q parametresi $\log$ fonksiyonunun q –deformasyonu olarak ifade edilir (Tsallis, 1988).

Kaniakadis κ –deformasyonlu bilinen logaritma fonksiyonu ise

$$\log_{\kappa}(f) = \frac{f^{\kappa} - f^{-\kappa}}{2\kappa}, \kappa > 0 \quad (2.35)$$

şeklinde tanımlanır ve κ parametresi $\log$ fonksiyonunun κ -deformasyonu olarak ifade edilir (Kaniadakis, 2005).

2.10 M-tahmin Yöntemi

M-tahmin yöntemi önerilen amaç fonksiyonun minimizasyonu

$$\hat{\theta} := \arg \min_{\theta} \sum_{i=1}^n \rho(x_i; \theta) \quad (2.36)$$

şeklinde veya

$$\sum_{i=1}^n \psi(x_i; \theta) = 0 \quad (2.37)$$

denklemin θ parametresine göre çözümü olarak tanımlanmaktadır. Araştırmacı amacına göre farklı ρ fonksiyonları seçerek dayanıklı tahmin ediciler $\hat{\theta}$ elde edebilir (Huber, 1964).

2.10 Huber M-Tahmin Edicisinde Kullanılan Fonksiyonlar

$$\rho_k(u) = \begin{cases} u^2 & , \quad |u| \leq k \\ 2k|u| - k^2 & , \quad |u| > k \end{cases} \quad (2.38)$$

$\psi_k(u) = \frac{d}{du} \rho_k(u)$ olduğundan $2\psi_k(u)$ elde edilir ve burada

$$\psi_k(u) = \begin{cases} u, & |u| \leq k \\ \text{sgn}(u)k, & |u| > k \end{cases} \quad (2.39)$$

şeklindedir. Dayanıklılıkta ağırlık fonksiyonu $w = \psi(u)/u$ şeklinde tanımlanır ve Huber M-tahmin edicisinin ağırlık fonksiyonu

$$w_k(u) = \begin{cases} 1, & |u| \leq k \\ \frac{\text{sgn}(u)}{u} k, & |u| > k \end{cases} \quad (2.40)$$

şeklinde verilir (Huber, 1964).

2.11 Regresyon Yapısında M-tahmin Edicileri

M-tahmin edicileri EÇO tahmin yönteminin bir genelleştirmesidir. Eşitlik (2.2)'de yer alan EÇO tahmin yönteminde, $\log$ fonksiyonu yerine $\log_q$ ve $\log_\kappa$ alınarak q ve κ deforme edilmiş logaritmalar aracılığı ile; ${}_q\hat{\beta}_M$ ve ${}_\kappa\hat{\beta}_M$ tahmin edicileri elde edilecektir.

EÇO tahmin yönteminde kullanılan fonksiyon $\log(f(x; \theta))$ 'dir. Böylesi bir durumda ise, θ parametresi için elde edilen tahmin edici $\hat{\theta}$, EÇO tahmin edicisidir. Eğer $\log(f(x; \theta))$ yerine keyfi bir $\rho(x; \theta)$ fonksiyonu alınırsa, bu durumda θ parametresi için elde edilen tahmin edici $\hat{\theta}$ M-tahmin edicisidir (Hampel vd., 1986; Çankaya ve Korbek, 2018).

ρ fonksiyonu yapısı kullanılarak

$$Y = X\beta + \varepsilon \quad (2.41)$$

regresyon modelindeki ε rastgele değişkeninin herhangi bir dağılım bilgisi olmayıp sapan gözlem içermesi durumunda M-tahmin yöntemindeki ρ fonksiyonu aracılığı ile β parametrelerinin M-tahmin edicilerini nasıl elde edildiği ele alınacaktır. Eşitlik (2.41)'deki verilen ifade ile gerçekte regresyon modelinin ilgili bu şekilde olduğu düşünülmektedir. Böylesi bir gerçeklikten rastgele gözlemler olan y_i 'ler gelmektedir. İlgili regresyondaki X 'lerin de sabit olduğu varsayılır ve her sabit X için rastgele gelen y_i gözlemlerinin rastgele ε_i hata terimlerini içerdiği düşünülmektedir. Ayrıca sabit ve bağımsız değişken X ile parametre olan β 'lar aracılığıyla bağımlı değişken Y 'lerin elde edimi regresyon tanımlaması içinde yer alır, y_i gözlemlerinde sapan gözlem olması durumunda $\sum_{i=1}^n \rho(y_i - x_i\beta)$ ifadesinin β parametrelerine göre minimizasyonu sonucu, $\hat{\beta}_M$ M-tahmin edicileri elde edilir.

ρ fonksiyonunun ilgili parametreye göre türevi $\psi = \frac{\partial}{\partial \beta} \rho$ fonksiyonu aracılığı ile gösterilir (Huber, 1981; Hampel vd., 1986). Bu gösterimden sonra, $\hat{\beta}_M$ M-tahmin edicilerini elde etme yöntemi aşağıda verildiği gibidir. Bunun için,

$$\beta = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \rho(y_i - x_i \beta) \quad (2.42)$$

olmak üzere β parametre vektörüne göre minimizasyon gerçekleştirilecektir. Bu optimizasyon süreci türev aracı ile gerçekleştirebilir.

$$\frac{\partial}{\partial \beta} \sum_{i=1}^n \rho(y_i - x_i \beta) = \sum_{i=1}^n \psi(y_i - x_i \beta) (-x_i^T) = \mathbf{0}. \quad (2.43)$$

Tahmin denkleminin β parametresine göre çözümü $\hat{\beta}_M$ M-tahmin edicilerini verecektir. Aynı zamanda, ağırlık fonksiyonu aracılığı ile de gösterilebilir (Huber, 1981).

$$\sum_{i=1}^n \psi(y_i - x_i \beta) \frac{(y_i - x_i \beta)}{(y_i - x_i \beta)} x_i^T = 0 \quad (2.44)$$

olmak üzere $\frac{\psi(y_i - x_i \beta)}{(y_i - x_i \beta)} = w(y_i - x_i \beta)$ şeklinde gösterilir. $w(y_i - x_i \beta)$ yerine kısaca w_i gösterimi tercih edilirse bu durumda, (2.44) eşitliği $\sum_{i=1}^n w_i (x_i^T y_i - x_i^T x_i \beta) = \mathbf{0}$ şeklini alır. Gerekli işlemler yapıldığında,

$$\sum_{i=1}^n (x_i^T w_i y_i - x_i^T w_i x_i \beta) = \mathbf{0} \quad (2.45)$$

olmak üzere

$$\hat{\beta}_M = \left[\sum_{i=1}^n x_i^T w_i x_i \right]^{-1} \sum_{i=1}^n x_i^T w_i y_i \quad (2.46)$$

şeklinde elde edilir.

(2.46) eşitliği matris gösterimi ile

$$\hat{\beta}_M = (X^T W X)^{-1} (X^T W Y) \quad (2.47)$$

şeklindedir ve burada

$$W_{n \times n} = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{bmatrix}_{n \times n} \quad (2.48)$$

olmak üzere ağırlık matrisidir (Güney, 2018).

2.12 M-tahmin Yönteminde Kullanılan Fonksiyonun Özelliği

Amaç fonksiyonu olarak adlandırılan $\rho(\cdot)$ fonksiyonunun özellikleri aşağıda verilmiştir:

(i) Eğer $|x_i| > |j_i|$ ise $\rho(x_i) \geq \rho(x_j)$ ise monotoniktir. Eğer $|x_i| \geq |x_j|$ ise $\rho(x_i) \geq \rho(x_j)$ yada kesinlikle monotoniktir.

(ii) $\rho(0) = 0$

(iii) $x > 0$ için $\rho(\cdot)$ fonksiyonu $\rho(x) < \rho(\infty)$ olacak şekilde artan bir fonksiyondur.

(iv) $\rho(\cdot)$ fonksiyonu sınırlı ise $\rho(\infty) = 1$ olduğu varsayılır (Maronna vd. 2006).

Burada $\psi(\cdot)$ skor fonksiyonu olarak adlandırılır ve aşağıdaki özellikleri sağlar:

(i) $\psi(-r) = -\psi(r)$ olmak üzere skor fonksiyonunun simetrik olduğunu gösterir.

(ii) $r > 0$ için $\psi(r) > 0$ dır.

(iii) $\psi(\cdot)$ sürekli ve parçalı türevlenebilir bir fonksiyondur.

Huber M-tahmin edicisi kesinlikle (strictly) monotonik olması özelliğini sağlamaktadır (Huber, 1981).

2.13 Deforme Edilmiş Logaritmalar Aracılığı ile En Çok Olabilirlik Tahmin Edicileri

Deforme edilmiş logaritmalar ile en çok olabilirlik tahmin yöntemini herhangi bir Λ fonksiyonu aracılığı ile aşağıdaki şekilde ifade edilebilir:

$$\arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n \Lambda(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n; \boldsymbol{\beta}) \quad (2.49)$$

Λ fonksiyonunun özelliğinin incelenmesinin yanı sıra $f(x; \boldsymbol{\theta})$ olasılık yoğunluk fonksiyonunun özelliği de incelenebilir. Normal dağılımın genelleştirmesi olan $f(x; \boldsymbol{\theta})$ olasılık yoğunluk fonksiyonunun mod özelliği incelenmiştir. Bu fonksiyon normal dağılımın genelleştirmesi olduğundan, bu fonksiyon ilgili parametre değerlerinde tek modlu olmaktadır ve normal dağılıma karşılık gelip tek modlu olma özelliğini sağlamaktadır (Çankaya, 2018). Sonuç olarak, $f(x; \boldsymbol{\theta})$ olasılık yoğunluk fonksiyonu Λ fonksiyonu aracılığı ile, başka yere taşınmaktadır ve konkav veya konveks olma özelliğini sağlamak zorundadır. Eğer bu özellik sağlanıyorsa, fonksiyon monotonik olur, yani $g(x_1) < g(x_2)$, $x_1 > x_2$ veya $x_1 < x_2$ için $g, x_1, x_2 \in R$.

İlgili amaç fonksiyonun $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$ parametre vektörüne göre minimum edilmesi ile bulunur. Belirtmek gerekir ki ilgili ifadenin $\boldsymbol{\beta}$ göre minimizasyonu ile

$$\arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n -\Lambda(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n; \boldsymbol{\beta}) \quad (2.50)$$

ifadesi optimizasyon sürecinde denktir. Burada, minimizasyon yapılıyor ise, ilgili Λ fonksiyonunun konveks; maksimizasyon yapılıyor ise, ilgili fonksiyonun konkav olması gereklidir. Çalışmada; Λ fonksiyonu, $\log_q[f(x; \boldsymbol{\beta})]$ ve $\log_\kappa[f(x; \boldsymbol{\beta})]$ fonksiyonları şeklinde seçilecektir (Ferrari ve Yang, 2010; Kaniadakis vd., 2005). Burada, $f(x; \boldsymbol{\beta})$ olasılık yoğunluk fonksiyonudur. Olasılık yoğunluk fonksiyonu $[0,1]$ aralığında değer alan bir fonksiyondur. Bu fonksiyonun normal dağılım veya student t dağılımı gibi tek modlu olduğu durum ile ilgilenilecektir.

M-tahmin edicilerinin skor fonksiyonunun sınırlı olması, M-tahmin edicinin dayanıklı olduğu anlamına gelmektedir (Hampel vd., 1986). Aşağıda verilen skor fonksiyonu $\boldsymbol{\Psi} = (\psi_1, \psi_2, \dots, \psi_p)$ ilgili parametreye göre Λ fonksiyonunun türevi alınması ile elde edilir:

$$\sum_{i=1}^n \Psi(\varepsilon_i; \beta) = \sum_{i=1}^n \Lambda(\varepsilon_i; \beta) = 0 \quad (2.51)$$

ve bu ifadeye tahmin denklemlerinin oluşturduğu denklem sistemi denir (Huber 1964; Huber, 1981. Çalışma kapsamında, Λ fonksiyonu yerine $\log_q(f(x; \theta))$ ve $\log_\kappa(f(x; \theta))$ fonksiyonları kullanılacaktır.

2.14 Kaniakadis Logaritması $\log_\kappa$ Özelliğinin İncelenmesi

Kaniakadis logaritması $\log_\kappa(f)$ fonksiyonunun argümanı f 'e göre bir kez türevi alındığında,

$$\frac{d}{df} \log_\kappa(f) = \frac{1}{2} f^{\kappa-1} + \frac{1}{2} f^{-\kappa-1} \quad (2.52)$$

şeklindeki fonksiyon elde edilir. Bu fonksiyonun kökü $e^{-\frac{1}{2} \ln \pi / \kappa}$ olup bir tek kompleks kökü vardır.

$\log_\kappa(f)$ fonksiyonunun argümanı f 'e göre iki kez türevi alındığında,

$$\frac{d^2}{df^2} \log_\kappa(f) = -\frac{1}{2} f^{-\kappa-2} \kappa + \frac{1}{2} f^{\kappa-2} \kappa - \frac{1}{2} f^{-\kappa-2} - \frac{1}{2} f^{\kappa-2} \quad (2.53)$$

şeklindeki fonksiyon elde edilir. $\kappa > 0$ olmak üzere, $\frac{d^2}{df^2} \log_\kappa(f)$ fonksiyonu negatif veya pozitif olabilir. Dolayısıyla, fonksiyonun konkav veya konveks olduğu gösterilemez. Buradan da, kesinlikle monotonik olduğu görülemez. Ancak, $\log_\kappa(f)$ fonksiyonu her f ve $\kappa > 0$ için farklı değerlere sahip olduğu açıktır. Bu ise, ilgili fonksiyonun kesinlikle monotonik olduğunu gösterir. Benzer şekilde, $\log_q(f)$ fonksiyonu düşünülebilir. Bu fonksiyonun kesinlikle monotonik olduğu aynı zamanda ikinci türev testi ile gösterilebilir (Çankaya ve Korbel, 2018).

2.15 Dağılım Bilgisi Altında Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri

Eşitlik (2.26) verilen normal dağılımın olasılık yoğunluk fonksiyonu kullanılarak, regresyon parametreleri $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ için EÇÖ tahmin edicileri aşağıda verilen log-olabilirlik fonksiyonu ile elde edilebilir.

$$\log(L(\boldsymbol{\beta}, \sigma; x)) = \sum_{i=1}^n \log \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y-x\boldsymbol{\beta})^2}{2\sigma^2}} \right] \quad (2.54)$$

Bunun için $\boldsymbol{\beta}$ parametre vektörüne göre $\log L$ fonksiyonunun türevi alınırsa,

$$\frac{\partial}{\partial \boldsymbol{\beta}} \log(L(\boldsymbol{\beta}, \sigma; x)) = \sum_{i=1}^n \left[\frac{-1}{2\sigma^2} 2(y_i - x_i \boldsymbol{\beta})(-x_i^T) \right] \quad (2.55)$$

şeklinde elde edilir. $\boldsymbol{\beta}$ parametre vektörünün EÇO tahmin edicisi $\hat{\boldsymbol{\beta}}$, (2.54) eşitliğindeki ifadenin $\boldsymbol{\beta}'$ ya göre türev alınması sonucu elde edilir.

$$\begin{aligned} \frac{-2}{2\sigma^2} \sum_{i=1}^n (x_i^T y_i - x_i^T x_i \boldsymbol{\beta}) &= 0 \\ \sum_{i=1}^n x_i^T y_i - \sum_{i=1}^n x_i^T x_i \boldsymbol{\beta} &= 0 \end{aligned} \quad (2.56)$$

$$\boldsymbol{\beta} \sum_{i=1}^n x_i^T x_i = \sum_{i=1}^n x_i^T y$$

$$\hat{\boldsymbol{\beta}} = [\sum_{i=1}^n x_i^T x_i]^{-1} [\sum_{i=1}^n x_i^T y] \quad (2.57)$$

şeklindedir. İlgili EÇO tahmin edici $\hat{\boldsymbol{\beta}}$ aynı zamanda ε hata terimleri için bir dağılım varsayımı olmadığında en küçük kareler (EKK) tahmin edicileridir. Eşitlik (2.57), matris gösterimi ile

$$\hat{\boldsymbol{\beta}} = (X^T X)^{-1} (X^T Y) \quad (2.58)$$

şeklinde yazılır ve burada X matrisi

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ & \vdots & & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}_{n \times p} \quad (2.59)$$

ve y kolon vektörü aynı zamanda bağımlı değişken y için gözlem vektörü

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{nx1} \quad (2.60)$$

şeklindedir. $\hat{\beta}$ EKK tahmin edicisinin boyutu hesaplanırsa,

$$\hat{\beta}_{px1} = (X_{pxn}^T X_{npx})^{-1} (X_{pxn}^T Y_{nx1}) \quad (2.61)$$

olmak üzere $px1$ bir kolon vektörüdür. Eğer modelde kesen parametresi β_0 varsa, açıklayıcı değişken sayısından bir fazlası kadar regresyon modelinde parametre sayısı vardır (Montgomery ve Peck, 1992).

2.16 M-Tahmin Edicileri: En Çok $\log_q$ Tahmin Edicileri

Şimdi eşitlik (2.2)' da yer alan $\log$ fonksiyonu yerine $\log_q$ yazılması ile maksimum $\log_q$ olabilirlik tahmin edicileri elde edilebilir (Ferrari ve Yang, 2010):

$$\log_q(L(\beta, \sigma; x)) = - \sum_{i=1}^n \log_q \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(y_i - x_i\beta)^2}{\sigma^2}} \right] \quad (2.62)$$

burada $\log_q(f) = \frac{f^{1-q}-1}{1-q}$ olmak üzere, (2.62) eşitliği

$$\log_q(L(\beta, \sigma; x)) = - \sum_{i=1}^n \frac{\left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(y_i - x_i\beta)^2}{\sigma^2}} \right]^{1-q} - 1}{1-q} \quad (2.63)$$

şeklinde elde edilir. Şimdi (2.63) eşitliği β' ya göre minimize edilmelidir. Bunun içinde β' ya göre türev alınarak optimizasyon (minimizasyon ve maksimizasyon) süreci türev aracı ile gerçekleştirilirse,

$$\begin{aligned} & - \sum_{i=1}^n \frac{(\sigma\sqrt{2\pi})^{q-1}}{1-q} \left[-\frac{1}{2} \frac{(1-q)}{\sigma^2} 2(y_i - x_i\beta)(-x_i^T) \right] e^{-\frac{1-q}{2\sigma^2} (y_i - x_i\beta)^2} = 0 \\ & - \sum_{i=1}^n [x_i^T y_i - x_i^T x_i \beta] e^{-\frac{1-q}{2\sigma^2} (y_i - x_i\beta)^2} = 0 \end{aligned} \quad (2.64)$$

burada $e^{-\frac{1-q}{2\sigma^2}(y_i-x_i\beta)^2} = w_i$ olmak üzere

$$-\sum_{i=1}^n x_i^T w_i y_i + \sum_{i=1}^n x_i^T w_i x_i \beta = 0 \quad (2.65)$$

$${}_q\hat{\beta}_M = \left[\sum_{i=1}^n x_i^T w_i x_i \right]^{-1} \sum_{i=1}^n x_i^T w_i y_i \quad (2.66)$$

şeklinde elde edilir. (2.66) eşitliği matris formatında yazılırsa

$${}_q\hat{\beta}_M = (X^T W X)^{-1} X^T W Y \quad (2.67)$$

şeklindedir. Burada (2.61) eşitliğinden farklı olarak ağırlık matrisi W gelmiştir ve

$$W_{n \times n} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix} \quad (2.68)$$

şeklinde ifade edilir.

(2.68) eşitliğinde verilen W matrisi $n \times n$ boyutundadır ve

$$\hat{\beta}_{px1} = (X_{pxn}^T W_{n \times n} X_{n \times p})^{-1} (X_{pxn}^T W_{n \times n} Y_{n \times 1}) \quad (2.69)$$

$(px1)$ boyutundaki kolon vektörü olmak üzere M-tahmin edicileri $\hat{\beta}_M$ elde edilmiş olur.

$\log_q(f)$ kullanıldığında, (2.67) eşitliğinde elde edilen ${}_q\hat{\beta}_M$ tahmin edicisi ile ρ fonksiyonu kullanıldığında, (2.47) eşitliğinde elde edilen $\hat{\beta}_M$ tahmin edicileri analitik olarak aynı forma sahiptirler. Burada değişen ifadeler seçilen ağırlık fonksiyonları W' dir.

2.17 M-Tahmin Edicileri: En Çok $\log_\kappa$ Tahmin Edicileri

Ferrari ve Yang 2009 maksimum $\log_q$ olabilirlik tahmin edicilerini önermiştir ve eşitlik (2.63)'de q tabanında M-tahmin edicilerini elde etmek için bir amaç fonksiyonu kullanılmıştır. Şimdi ise benzer süreç $\log_q$ yerine $\log_\kappa$ alındığında yürütülecek ve κ tabanında M-tahmin edicileri elde edilecektir.

$$\log_\kappa(L(\boldsymbol{\beta}, \sigma; \mathbf{x})) = \sum_{i=1}^n \log_\kappa \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(y_i - x_i \boldsymbol{\beta})^2}{\sigma^2}} \right] \quad (2.70)$$

şeklindeki amaç fonksiyonunu ele alalım. $\log_\kappa$ fonksiyonu (2.35) eşitliğinde verilmektedir. Buna göre

$$\begin{aligned} \log_\kappa(L(\boldsymbol{\beta}, \sigma; \mathbf{x})) &= \sum_{i=1}^n \frac{1}{2\kappa} \left[(\sigma\sqrt{2\pi})^{-\kappa} e^{-\frac{1}{2\sigma^2}(y_i - x_i \boldsymbol{\beta})^\kappa} \right. \\ &\quad \left. - (\sigma\sqrt{2\pi})^\kappa e^{-\frac{1}{2\sigma^2}(y_i - x_i \boldsymbol{\beta})(-\kappa)} \right] \end{aligned} \quad (2.71)$$

şeklindeki amaç fonksiyonu kullanılacaktır. Bu fonksiyonun $\boldsymbol{\beta}$ parametresine göre optimizasyonu (minimizasyonu- ya da maksimizasyon) κ - tabanında $\hat{\boldsymbol{\beta}}_M$ M-tahmin edicilerinin elde edimini sağlayacaktır. Optimum noktayı bulmak için türev aracı kullanılırsa ve (2.71) eşitliğinin $\boldsymbol{\beta}$ 'ya göre türevi ile:

$$\frac{\partial}{\partial \boldsymbol{\beta}} \log_\kappa(L(\boldsymbol{\beta}, \sigma; \mathbf{x})) = \sum_{i=1}^n \left[\frac{(\sigma\sqrt{2\pi})^{-\kappa}}{2\kappa} \frac{-\kappa}{2\sigma^2} 2(y_i - x_i \boldsymbol{\beta})(-x_i^T) e^{-\frac{1}{2\sigma^2}(y_i - x_i \boldsymbol{\beta})^\kappa} \right] \quad (2.72)$$

$$-\sum_{i=1}^n \left[\frac{(\sigma\sqrt{2\pi})^{-\kappa}}{2\kappa} \frac{-\kappa}{-2\sigma^2} 2(y_i - x_i \boldsymbol{\beta})(-x_i^T) e^{-\frac{1}{2\sigma^2}(y_i - x_i \boldsymbol{\beta})(-\kappa)} \right] = 0 \quad (2.73)$$

şeklindeki tahmin denklemi elde edilir. Gerekli analitik düzenlemeler yapıldığında, (2.73) eşitliği

$$\begin{aligned}
& \sum_{i=1}^n \left[\frac{(\sigma\sqrt{2\pi})^{-\kappa}}{2\kappa\sigma^2} \kappa (x_i^T y_i - x_i^T x_i \beta) w_{1i} - \frac{(\sigma\sqrt{2\pi})^{\kappa}}{2\kappa \cdot \sigma^2} (-\kappa) (x_i^T y_i - x_i^T x_i \beta) w_{2i} \right] \\
& = 0 \\
& (\sigma\sqrt{2\pi})^{-2\kappa} \sum_{i=1}^n [x_i^T w_{1i} y_i - x_i^T w_{1i} x_i \beta] = \sum_{i=1}^n [-x_i^T w_{2i} y_i + x_i^T w_{2i} x_i \beta] \\
& \sum_{i=1}^n (\sigma\sqrt{2\pi})^{-2\kappa} x_i^T w_{1i} y_i - \sum_{i=1}^n (\sigma\sqrt{2\pi})^{-2\kappa} x_i^T w_{1i} x_i \beta = \\
& - \sum_{i=1}^n x_i^T w_{2i} y_i + \sum_{i=1}^n x_i^T w_{2i} x_i \beta
\end{aligned} \tag{2.74}$$

olmak üzere

$$\begin{aligned}
& \sum_{i=1}^n (\sigma\sqrt{2\pi})^{-2\kappa} x_i^T w_{1i} y_i + \sum_{i=1}^n x_i^T w_{2i} y_i = \sum_{i=1}^n (\sigma\sqrt{2\pi})^{-2\kappa} x_i^T w_{1i} x_i \beta + \\
& \sum_{i=1}^n x_i^T w_{2i} x_i \beta
\end{aligned} \tag{2.75}$$

şeklinde elde edilir. Buradan da, her iki taraf $(\sigma\sqrt{2\pi})^{2\kappa}$ ile çarpılırsa

$$\begin{aligned}
\hat{\beta}_M = & \left[\sum_{i=1}^n x_i^T w_{1i} x_i + x_i^T w_{2i} x_i (\sigma\sqrt{2\pi})^{2\kappa} \right]^{-1} \left[\sum_{i=1}^n x_i^T w_{1i} y_i \right. \\
& \left. + x_i^T w_{2i} y_i (\sigma\sqrt{2\pi})^{2\kappa} \right]
\end{aligned} \tag{2.76}$$

olarak elde edilir. Burada $w_{1i} = e^{-\frac{1}{2\sigma^2}(y_i - x_i \beta)^2 \cdot \kappa}$ ve $w_{2i} = e^{-\frac{1}{2\sigma^2}(y_i - x_i \beta)^2 (-\kappa)}$ şeklindedir. (2.76) eşitliği matris gösterimiyle,

$${}_{\kappa} \hat{\beta}_M = [X^T W_1 X + X^T W_2 X (\sigma\sqrt{2\pi})^{2\kappa}]^{-1} [X^T W_1 Y + X^T W_2 Y (\sigma\sqrt{2\pi})^{2\kappa}] \tag{2.77}$$

şeklindedir.

2.18 Genelleştirilmiş Ridge Regresyon

Ridge regresyon çoklu bağlantı iç ilişki durumunda regresyon parametrelerinin tahmini için önerilmiştir (Hoerl ve Kennard, 1970). Bu regresyon, rastgele olmayan bir parametre λ ile regresyon katsayısının β bir fonksiyonu aracılığı ile oluşturulmuştur. Genel olarak ifade edilirse,

$$\lambda r(\beta^T \beta)$$

şeklindeki bir kısıtın yer alması ile oluşturulmaktadır. r konveks ya da konkav bir fonksiyondur. λ sabit bir katsayıdır. Seçilen r fonksiyonu genelleştirmeyi belirlemektedir. Çalışmada, özel olarak $r(\beta^T \beta) = \beta^T \beta$ alınmıştır. $\beta^T \beta$ ifadesinin konveks ya da konvaxlığını işlemek için ikinci türev testi kullanılsın (Arslan, 2016; Hua ve Gunst, 1983). Buradan hareketle, $\frac{\partial}{\partial \beta} \beta^T \beta = 2\beta$ ve $\frac{\partial^2}{\partial \beta^T \partial \beta} \beta^T \beta = 2$ olacak pozitif sayı olduğu için konvekslik sağlanacaktır.

Ridge regresyon aşağıdaki ifadenin

$$\underset{\beta}{\text{argüman optimizasyon}} \sum_{i=1}^n [\Lambda(y - x\beta) + \lambda r(\beta^T \beta)] \quad (2.78)$$

β parametresine göre optimizasyon edilmesidir. Eğer $\Lambda = u^2$ olursa, diğer bir ifadeyle $\Lambda(Y - X\beta) = (Y - X\beta)^2$, bu durumda en küçük kareler (EKK) tahmin edicileri elde edilir. $\lambda r(\beta^T \beta)$ Lagrangian kısıtı ile birlikte ridge EKK tahmin edicileri elde edilir. (Duran,2011). Eğer $\Lambda = \log_q[f(x; \beta)]$ olursa, en çok $\log_q$ olabilirlik tahmin edicileri elde edilir. Benzer şekilde, $\log_q$ yerine $\log_\kappa$ alınırsa, en çok $\log_\kappa$ olabilirlik tahmin edicileri olacaktır. Çalışmada seçilen $\lambda r(\beta^T \beta)$ kısıtı amaç fonksiyonun analitik şekline göre düzenlenmektedir ve Λ fonksiyonu için farklı fonksiyonlar seçilmektedir. Eğer $\Lambda = \log$ ise, klasik bilinen EÇO tahmin edicileri elde edilecektir. Λ fonksiyonunu seçmiş olmak ile EÇO'nun bir genelleştirilmesi yapılmış olur ve M-tahmin edicileri sınıfına girilir.

2.19 Ridge Regresyon Yapısında M-tahmin Edicileri

Ridge tahmin edicilerini elde etmek için bir kısıt fonksiyonu seçilmektedir. ${}_q\hat{\beta}_{MR}$ ve ${}_\kappa\hat{\beta}_{MR}$ M-tahmin edicilerini elde etmek için ridge parametresi λ ile gerekli kısıtlar konulmuştu. Şimdi ise, $\hat{\beta}_M$ tahmin edicilerinin ridge M-tahmin edicisinin elde edilmesi için aşağıdaki fonksiyon ele alınır.

$$\hat{\beta}_{MR} = \underset{\beta}{\text{argmin}} \sum_{i=1}^n [\rho(y_i - x_i\beta) + \lambda \beta^T \beta] \quad (2.79)$$

İlgili fonksiyonun β parametresine göre minimum edilmesi sonucu $\hat{\beta}_{MR}$ ridge M-tahmin edicileri elde edilir. Bu optimizasyon süreci türev aracılığı ile gerçekleştirilebilir ve

$$\frac{\partial}{\partial \beta} \sum_{i=1}^n \left[\rho(y_i - x_i \beta) + \frac{1}{2} \lambda \beta^2 \right] = \sum_{i=1}^n [\psi(y_i - x_i \beta)(-x_i^T) + \lambda \beta] = 0 \quad (2.80)$$

şeklindeki tahmin denkleminin β parametresine göre çözümü sonucu $\hat{\beta}_{MR}$ M-tahmin edicileri elde edilir. (2.80) eşitliği düzenlenirse,

$$\sum_{i=1}^n [w_i(-x_i^T y_i + x_i^T x_i \beta) + \lambda \beta] = 0 \quad (2.81)$$

olmak üzere

$$\begin{aligned} \sum_{i=1}^n -x_i^T w_i y_i + \sum_{i=1}^n [x_i^T w_i x_i \beta + \lambda \beta] &= 0 \\ \beta \sum_{i=1}^n x_i^T w_i x_i + \lambda &= \sum_{i=1}^n x_i^T w_i y_i \end{aligned} \quad (2.82)$$

ifadeleri elde edilir ve

$$\hat{\beta}_{MR} = \left[\sum_{i=1}^n x_i^T w_i x_i + \lambda \right]^{-1} \sum_{i=1}^n x_i^T w_i y_i \quad (2.83)$$

şeklinde ridge M-tahmin edicisi $\hat{\beta}_{MR}$ elde edilir. (2.83) eşitliği matris gösterimi ile

$$\hat{\beta}_{MR} = [X^T W X + \lambda I]^{-1} X^T W Y \quad (2.84)$$

şeklinde yazılır. λ ridge parametresidir.

Belirtmek gerekir ki; (2.67) eşitliğindeki ${}_q \hat{\beta}_{MR}$ ile (2.84) eşitliğindeki $\hat{\beta}_{MR}$ karşılaştırıldığında aynı analitik formda olduğu görülmektedir ve bunların ağırlık fonksiyonları farklıdır. Ayrıca, (2.84) eşitliğindeki $\lambda = 0$ ise (2.67) eşitliğindeki ${}_q \hat{\beta}_M$ tahmin edicisi elde edilmektedir.

2.20 Dağılım Bilgisi Altında Ridge Regresyon Parametrelerinin En Çok Olabilirlik Tahmin Edicileri

Yukarıdaki paragraflarda ε rastgele değişkenlerinin normal dağıldığı varsayımı altında EÇO tahmin edicileri bulunmuştur. Şimdi yine aynı varsayım altında (benzer şekilde $\hat{\beta}$ EKK tahmin edicileridir) ridge regresyon aracılığı ile $\hat{\beta}_R$ tahmin edicilerini elde edilecektir. (2.54) eşitliği ele alınsın,

$$\log(L(\beta, \sigma; x)) = \sum_{i=1}^n \log \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y-x\beta)^2}{2\sigma^2}} \right] + \sum_{i=1}^n \frac{1}{2} \lambda \beta^T \beta \quad (2.85)$$

fonksiyonu β parametresine göre minimum yapılacaktır. (2.85) eşitliğinin β ' ya göre türevi alınırsa,

$$\frac{\partial}{\partial \beta} \log L = \sum_{i=1}^n \left[\frac{1}{2\sigma^2} 2 (y_i - x_i \beta) (-x_i^T) + \lambda \beta \right] = 0 \quad (2.86)$$

olmak üzere $\sigma = 1$ için

$$\begin{aligned} &= \sum_{i=1}^n [-x_i^T y_i + x_i^T x_i \beta + \lambda \beta] = 0 \\ &\sum_{i=1}^n (x_i^T x_i + \lambda) \beta = \sum_{i=1}^n x_i^T y_i \end{aligned} \quad (2.87)$$

$$\hat{\beta} = \left[\sum_{i=1}^n x_i^T x_i + \lambda \right]^{-1} \sum_{i=1}^n x_i^T y_i \quad (2.88)$$

şeklinde elde edilir. Matris formatında yazarsa

$$\hat{\beta} = (X^T X + \lambda I)^{-1} X^T Y \quad (2.89)$$

olarak elde edilir. Burada λ ridge parametresidir. I birim matrisi göstermektedir. $\hat{\beta}_R$ ridge parametrelerinin EKK veya dağılım bilgisi altında normal dağılımın konumuna karşılık gelen regresyon parametrelerinin ridge tahmin edicisidir.

(2.58) ve (2.88) eşitlikleri karşılaştırıldığında, (2.88) eşitliğinde λ parametresinin geldiği görülmektedir. Bu parametre, (2.85) eşitliğinde koyulan $\sum_{i=1}^n \frac{1}{2} \lambda \boldsymbol{\beta}^T \boldsymbol{\beta}$ kısıtından dolayı gelmektedir. Aynı zamanda, (2.88) eşitliğinde yer alan $\sum_{i=1}^n \frac{1}{2} \lambda \boldsymbol{\beta}^T \boldsymbol{\beta}$ kısıtından dolayı gelmektedir. (2.58) eşitliğine bakıldığında, $X^T X$ matrisinin tersi çoklu bağlantı probleminde sıkıntılı olacağından dolayı, $X^T X$ matrisine λI matrisinin eklenmesi $X^T X$ matrisinde değişime neden olacaktır ve bu matrisin tersinin sıfıra yakın gelmesi durumunu ortadan kaldırmaya yarayacaktır.

2.21 MR-Tahmin Edicileri: Ridge En Çok $\log_q$ Tahmin Edicileri

ε rastgele değişkeninin normal dağıldığı varsayımı altında maksimum $\log_q$ olabilirlik tahmin edicileri (M-tahmin edicileri) eşitlik (2.67)'da bulunmuştur. Şimdi yine aynı varsayım altında ridge regresyon aracılığı ile $\hat{\boldsymbol{\beta}}$ tahmin edicileri elde edilecektir. (2.62) eşitliği ridge kısıtı ile yeniden ele alınırsa

$$\begin{aligned} \log_q(L(\boldsymbol{\beta}, \sigma; x)) &= - \sum_{i=1}^n \log_q \left[\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \frac{(y_i - x_i \boldsymbol{\beta})^2}{\sigma^2}} \right] \\ &\quad + \sum_{i=1}^n \frac{1}{2} \lambda \boldsymbol{\beta}^T \boldsymbol{\beta} (\sigma \sqrt{2\pi})^{q-1} \frac{1}{\sigma^2} \end{aligned} \quad (2.90)$$

$$= (\sigma \sqrt{2\pi})^{q-1} \left[\frac{1}{\sigma^2} (x_i^T y_i - x_i^T x_i \boldsymbol{\beta}) \right] e^{-\frac{1-q}{\sigma^2} (y_i - x_i \boldsymbol{\beta})^2} + \lambda \boldsymbol{\beta} (\sigma \sqrt{2\pi})^{q-1} \frac{1}{\sigma^2} = 0$$

yazılabilir. $e^{-\frac{1-q}{\sigma^2} (y_i - x_i \boldsymbol{\beta})^2} = w_i$ olmak üzere

$$-(x_i^T w_i y_i - x_i^T w_i x_i \boldsymbol{\beta}) + \lambda \boldsymbol{\beta} = 0 \quad (2.91)$$

$${}_q \hat{\boldsymbol{\beta}}_{MR} = \left[\sum_{i=1}^n x_i^T w_i x_i + \lambda \right]^{-1} \sum_{i=1}^n x_i^T w_i y_i \quad (2.92)$$

şeklinde elde edilir. Matris formatında yazılırsa,

$${}_q\hat{\boldsymbol{\beta}}_{MR} = [X^T W X + \lambda I]^{-1} X^T W Y \quad (2.93)$$

olarak elde edilir. Burada, λ ridge parametresidir. I birim matrisidir. ${}_q\hat{\boldsymbol{\beta}}_{MR}$ eşitlik (2.93)'deki ridge regresyon ifadesinden elde edilen ridge maksimum $\log_q$ olabilirlik (M-tahmin edicisi) olarak ifade edilmektedir.

(2.67) ve (2.93) eşitlikleri karşılaştırıldığında, (2.93) eşitliğinde λ parametresinin geldiği gözlenmiştir. Bu parametre, (2.90) eşitliğinde yer alan $\sum_{i=1}^n \frac{1}{2} \lambda \boldsymbol{\beta}^T \boldsymbol{\beta} (\sigma \sqrt{2\pi})^{q-1} \frac{1}{\sigma^2}$ kısıtından dolayı gelmektedir. Aynı zamanda, (2.93) eşitliğine bakıldığında, $X^T W X$ matrisinin tersi çoklu bağlantı probleminde sıkıntılı olacağından dolayı, $X^T W X$ matrisine yapılan λI eklemesi $X^T W X$ matrisinde değişime neden olacaktır ve tersin sifıra yakın gelmesi durumunu ortadan kaldırmaya çalışacaktır.

2.22 MR-Tahmin Edicileri: Ridge En Çok $\log_\kappa$ Tahmin Edicileri

ϵ rastgele değişkeninin normal dağıldığı varsayımı altında maksimum $\log_\kappa$ olabilirlik tahmin edicileri (M-tahmin edicileri) eşitlik (2.96)'de bulunmuştu. Şimdi yine aynı varsayım altında ridge regresyon aracılığı ile $\hat{\boldsymbol{\beta}}_{MR}$ tahmin edicilerini elde edilecektir. (2.70) eşitliğinin negatif ile çarpılmış halini alır, $\sum_{i=1}^n \frac{1}{2} \lambda \boldsymbol{\beta}^T \boldsymbol{\beta} \kappa \left[\frac{\sigma (\sqrt{2\pi})^{-\kappa}}{\sigma^2 2\kappa} \right]$ kısıtı eklenir ve bu kısıt altında (2.70) eşitliğinin $\boldsymbol{\beta}$ 'ya göre türevi alınırsa gerekli düzenlemelerden sonra

$$\sum_{i=1}^n [-x_i^T w_{1i} y_i + x_i^T w_{1i} x_i \boldsymbol{\beta} - A x_i^T w_{2i} y_i + A x_i^T w_{2i} x_i \boldsymbol{\beta} + \lambda \boldsymbol{\beta}] = 0 \quad (2.94)$$

ifadesi elde edilir. $A = (\sigma \sqrt{2\pi})^{2\kappa}$ için

$${}_k\hat{\boldsymbol{\beta}}_{MR} = \left[\sum_{i=1}^n x_i^T w_{1i} x_i + x_i^T w_{2i} x_i A + \lambda \right]^{-1} \left[\sum_{i=1}^n x_i^T w_{1i} y_i + x_i^T w_{2i} y_i A \right] \quad (2.95)$$

olarak elde edilir.

(2.95) eşitliği matris gösterimiyle,

$${}_{\kappa}\widehat{\boldsymbol{\beta}}_{MR} = (X^T W_1 X + X^T W_2 X A + \lambda I)^{-1} (X^T W_1 Y + X^T W_2 Y A) \quad (2.96)$$

şeklindedir.

(2.77) ve (2.96) eşitlikleri karşılaştırıldığında, ridge regresyondan gelen λ parametresi ile $X^T W_1 X + X^T W_2 X A$ matrisindeki olabilecek çoklu bağlantı problemi ve çoklu bağlantının derecesi kadar sıfır yakın gelme durumu söz konusudur. q -tabanında elde edilen ${}_q\widehat{\boldsymbol{\beta}}_{MR}$ ve κ -tabanında elde edilen ${}_{\kappa}\widehat{\boldsymbol{\beta}}_{MR}$ M-tahmin edicileri sırasıyla eşitlikler (2.93) ve (2.96)'da yer almaktadır. (2.93) ve (2.96) eşitlikleri karşılaştırıldığında, (2.96) eşitliğinde $X^T W_2 X A$ ve $X^T W_2 Y$ ifadelerinin geldiği görülmektedir. Bu ise, $\log_{\kappa}(f) = \frac{f^{\kappa} - f^{-\kappa}}{2\kappa}$ ifadesindeki $f^{-\kappa}$ ifadesi w_2 matrisinin yer aldığı ifadelere denk gelmektedir. Ancak, $\log_q(f) = \frac{f^{1-q} - 1}{1-q}$ şeklinde bir analitik ifade mevcut olup, f^{1-q} yapısı kendi içinde sadece bir tane ağırlık fonksiyonu matris W getirmektedir. W matrisi

$$W_{1_{n \times n}} = \begin{bmatrix} w_{11} & 0 & \dots & 0 \\ 0 & w_{12} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{1n} \end{bmatrix}_{n \times n} \quad (2.97)$$

$$W_{2_{n \times n}} = \begin{bmatrix} w_{21} & 0 & \dots & 0 \\ 0 & w_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{2n} \end{bmatrix}_{n \times n} \quad (2.98)$$

şeklindedir.

2.23 Korelasyon: Doğrusal İlişki

Korelasyon değeri, sadece iki tane değişken arasındaki doğrusal ilişkiyi göstermek için kullanılır. $i \neq j$ olmak üzere X_i ve X_j , $i, j = 1, 2, \dots, p$ değişkenleri arasında doğrusal ilişki mevcut ise, korelasyon formülü

$$\rho_{X,Y} = \text{corr}(X,Y) = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (2.99)$$

$$r_{xy} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.100)$$

aracılığı ile ilişkinin derecesi belirlenir. $0 \leq \rho \leq 1$ aralığında değerler alır. Eğer $\rho \rightarrow 0$ ise, doğrusal ilişkinin derecesi düşük olduğunu gösterir. Eğer $\rho \rightarrow 1$ ise, doğrusal ilişkinin derecesinin yüksek olduğunu gösterir. İlişki derecesine göre, çoklu bağlantının da yüksek olacağını gösterir (Şenesen, 2001).

2.24 M-Tahmin Edicisinin Sayısal Hesaplaması

M-tahmin yönteminde elde edilen M-tahmin edicileri, kullanılan amaç fonksiyonları ρ ve bunların türevinden elde edilen Ψ fonksiyonları β parametrelerinin tahminlerinin analitik olarak yalnız kalmasını sağlayamayabilir. Bu durumda, $\hat{\beta}$ elde etmek için nümerik yöntemler kullanılmaktadır. M-tahmin edicilerinin çözümünde, ilgili fonksiyondan dolayı analitik ifadeler oluşmadığında sayısal yöntemler kullanılır. Bunlardan birisi Newton – Raphson algoritmasıdır. Burada Ψ fonksiyonunun türevi gerekli olduğundan, ilgili Ψ fonksiyonunda kök bulma problemi tanımsız veya güvensiz olur. Bu tür durumda, iteratif olarak tekrarlanan ağırlıklandırma (IRA) olarak isimlendirilen sayısal yöntem kullanılabilir.

$$\hat{\beta}^{(c)} = [\sum_{i=1}^n x_i^T w_i^{(c-1)} x_i]^{-1} \cdot \sum_{i=1}^n x_i^T w_i^{(c-1)} y_i \quad (2.101)$$

Burada c hesaplamadaki adım sayısını göstermektedir. Ağırlık fonksiyonu w_i , $\hat{\beta}$ parametrelerinin bir fonksiyonunu içerdiğinden dolayı, yani $w_i = w(x_i, \hat{\beta})$, burada $\hat{\beta}$ 'ya göre çözülmüş analitik bir ifade yoktur. Dolayısıyla, hesaplama işlemi gerçekleştirilmesi için $\hat{\beta}$ için başlangıç değeri $\hat{\beta}^0$ verilmelidir. Çalışmada, p tane başlangıç değeri $\hat{\beta}^0 = (0,0, \dots, 0)_{p \times p}$ alınmıştır.

IRA algoritmasının adımları aşağıda verilmektedir:

1. $\hat{\beta}^{(0)}$ başlangıç değeri verilir. Gerçek veri setinde başlangıç değerleri sıfır olarak verilmektedir.

2. Her bir adım sayısı (c) sırasıyla $\varepsilon_i^{(c)} = y_i - x_i \hat{\beta}^{(c-1)}$ artıklar hesaplanır.

3. $w_i^{(c)}$ ağırlıkları aşağıdaki eşitlik kullanılarak hesaplanır.

$$w_i^{(c)} = \frac{\psi(\varepsilon_i^{(c)})}{\varepsilon_i^{(c)}} \quad (2.102)$$

4. $W^{(c)} = \text{diagonal} (\{w_i^{(c)} : i = 1, 2, \dots, n\})$ ağırlık matrisi bulunur.

5. $\hat{\beta}^{(c)} = (X^T W^{(c)} X)^{-1} X^T W^{(c)} y$ eşitliğinden $\hat{\beta}^{(c)}$ hesaplanır.

7. Yakınsama yani kabul edilen hata e sağlanana kadar 2-5 adımları tekrar edilir (Arslan, 2004, Arslan ve Genç 2009).

2.24.1 Ridge M-Tahmin Edicisinin Sayısal Hesaplaması

Ridge durumunda sadece λ parametresi gelmektedir. Bu durumda,

$$\hat{\beta}^{(c)} = [\sum_{i=1}^n x_i^T w_i^{(c-1)} x_i + \lambda]^{-1} \cdot \sum_{i=1}^n x_i^T w_i^{(c-1)} y_i \quad (2.103)$$

olmak üzere tahmin değerleri elde edilir. λ parametresi kullanıcı tarafından verildiği için hesaplama işleminde sürekli olarak sabit şekilde yer alır.

Herhangi bir tahmin edicinin hata kareler (HKO) ortalaması aşağıda verilmektedir.

$$HKO(\hat{\beta}) = E(\hat{\beta} - \beta)^2 = \text{Var}(\hat{\beta}) + [E(\hat{\beta}) - \beta]^2 \quad (2.104)$$

ve

$$HKO(\hat{\beta}) = \text{Var}(\hat{\beta}) + (\text{yanlılık}(\hat{\beta}))^2 \quad (2.105)$$

Eğer ridge tahmin edicisi kullanılırsa, $\hat{\beta}$ tahmin edicilerinin yukarı doğru anlamsız yanlılığı ortadan kaldırılmış olur (Şenesen, 2001).

2.25 Çoklu İç İlişki Durumu

$X^T X$ veya $X^T I X$ ($X^T W X$) matrisinin özdeğerlerinin analizi: çoklu iç ilişki problemini ciddi anlamda çalışan ilk araştırmacı Ragnar Frisch'dır. Çoklu iç ilişkiyi özdeğerlerle ilişkilendirmiştir (Frisch, 1934). Fakat bilgisayarların yaygın olmaması nedeni ile $X^T X$ veya $X^T I X$ ($X^T W X$) öz değerlerinin nümerik analizi desteklenememiştir (Vinod ve Ullah,1981). Eğer veride bir veya daha fazla lineer bağımlılık varsa, bir veya daha fazla özdeğer küçük olacaktır. Küçük özdeğer, X in kolonları arasındaki yaklaşık lineer bağımlılığı gösterir. λ_{min} ve λ_{maks} , $X^T X$ veya $X^T I X$ ($X^T W X$) matrisinin minimum ve maksimumun özdeğerleri göstermek üzere $K = \frac{\lambda_{maks}}{\lambda_{min}}$ olarak tanımlanan koşul sayısı çoklu iç ilişkiyi belirlemek için kullanılır. K , 100 den büyükse şiddetli bir çoklu iç ilişki vardır.

Çoklu iç ilişki pek çok teknikle belirlenebilir. Bunlardan bir tanesi. $X^T X$ veya $X^T I X$ ($X^T W X$) determinantıdır. Standartlaştırılmış formda $0 \leq |X^T X$ veya $X^T W X$ ($X^T W X$) ≤ 1 'dir. $|X^T X$ veya $X^T I X$ ($X^T W X$)| = 1| ise açıklayıcı değişkenler ortogonaldır. Aksi durumda, $|X^T X$ veya $X^T I X$ ($X^T W X$)| = 0 iken tam çoklu iç ilişki vardır. $|X^T X$ veya $X^T I X$ ($X^T W X$)| sıfıra yaklaştıkça çoklu iç ilişkinin şiddeti artacaktır (Farrar ve Glauber,1967). Regresyon katsayılarının işaretlerinin veya büyüklüklerinin beklenenden farklı olması, açıklayıcı değişkenlerin eklenmesi veya çıkarılması ile regresyon katsayılarının tahminlerinde büyük değişikliklere neden olması çoklu iç ilişkinin göstergesi olabilir.

2.25.1 Çoklu İç İlişkinin Ortadan Kaldırılmasına İlişkin Çözüm Yöntemleri

1. Modelin yeniden belirlenmesi: Çoklu iç ilişkinin nedeni model seçiminden kaynaklanabilir, örneğin ilişkili açıklayıcı değişkenlerin kullanılması gibi. Bu durumda ya açıklayıcı değişkenler yeniden tanımlanır ya da ilişkili açıklayıcı değişkenlerin biri çıkartılır. Fakat açıklayıcı değişkenlerin çıkartılması modelin etkinliğini azaltabilir. Çünkü Lipovetsky ve Conklin, 2001 söz ettiği üzere, çoklu iç ilişki olsa bile değişkenler birbirini tam olarak temsil etmeyebilirler. Açıklayıcı değişkenlerin her biri veriyi modelleme de (uyumda) ve bağımlı değişkenin yapısını açıklamada özel bir role sahip olabilir.

2. Alternatif regresyon analizinin kullanılması: Regresyon analizinde asıl olarak regresyon katsayıları ile ilgileniyorsak EKK en çok kullanılan yöntemdir ve dayanıklı formları ise son zamanlarda çalışılmıştır (Arslan, 2016; Xu ve Ying, 2010).

2.26 Bilgi Kriterinin Kullanımı

1970'lerin başından beri, geçtiğimiz birkaç on yılda, model seçim algoritması ve kriterleri ile ilgili birçok çalışmaya rastlamak mümkündür. Bunlar arasında klasik model seçim yöntemleri ve bilgi kriterlerine dayanan model seçim metotları yer almaktadır.

Model seçimi, temelde açıklayıcı değişkenlerden hangisi veya hangilerinin cevap değişkeni üzerinde etkili olduğunu ortaya koyarak mevcut modeller altkütmesi içinden en iyi modelin bulunması ile ilgili bir süreçtir. Açıklayıcı değişken sayısının çok olduğu ve model hakkında önsel bilginin bulunmadığı veya yetersiz olduğu durumlarda en iyi modelin belirlenmesi araştırmacıların önündeki en büyük problemlerden biridir.

AIC (Akaike information criterion) ve BIC (Bayesian information criterion) bilgi kriterleri sırasıyla aşağıdaki gibi tanımlanmaktadır:

$$AIC = -2 \log(L(\boldsymbol{\beta}; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)) + 2p \quad (2.106)$$

$$BIC = -2 \log(L(\boldsymbol{\beta}; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)) + p \log(n) \quad (2.107)$$

burada p tahmin edilen parametre sayısıdır. Özellikle, BIC uygulamalı bilimlerde yaygın bir şekilde kullanılmaktadır (Weakliem, 1999). $-2\log(L)$ ilgili bilgi kriterlerinde fonksiyonun veriyi iyi modellememesi (lack of fit) terimi olarak, $2p$ ve $p\log(n)$ parçaları ceza terimi olarak tanımlanmaktadır. Eşitlikler, (2.62) ve (2.70) ile bunların ridge kısıtlı formları, veri setini modellemek için kullanılan amaç fonksiyonlarıdır. Ridge formlarında, veri setini modellemek için aynı zamanda ceza terimi eklenmiştir. Artık kareler toplamı minimumum değerinden fazla uzaklaşmamalıdır. Bilgi kriteri ile en iyi hangi amaç fonksiyonunun modellediği belirlenebilir (McDonald ve Galarneau, 1975). Burada, $\log_\kappa$ ve $\log_q$ fonksiyonlarının karşılaştırılabilir olması için buradan gelen tahmin değerleri $\log$ fonksiyonu temelinde kurulu olan BIC formülüne konularak ilgili bilgi kriterinin değeri elde edilmiştir. Aynı zamanda, karşılaştırılabilirliği homojen yapmak için ridge regresyon elde ediminden dolayı kullanılan $\lambda \boldsymbol{\beta}^T \boldsymbol{\beta}$ kısıtı $-2\log L$ parçasına eklenmiş ve

$$-2\log L - 2\lambda\boldsymbol{\beta}^T\boldsymbol{\beta} \quad (2.108)$$

şeklinde önerilen formül ile bilgi kriterinin uyum eksikliği parçası ridge parametresi ile oluşturulmuştur.



3. UYGULAMA ve BULGULAR

Gerçek veri setinde uygulama şu şekilde yapılmıştır. Ridge parametresinin üç farklı değeri alınmıştır. Bunlar, sırasıyla, $\lambda = 0$, $\lambda = 0.05$ ve son olarak ta $\lambda = 0.1$ 'dir. İlgili λ değerlerine ilişkin olarak hem dayanıklı hem de muhtemel dayanıksız olabilecek durumda tahminler yapılmıştır. Muhtemel dayanıksız olabilecek durumlar tuning parametrelerinin $\kappa \rightarrow 0$ ve $q \rightarrow 1$ klasik bilinen logaritmaya ($\log$) yakın olduğu durumlardır. Dayanıklı durum ise, $\kappa > 0$ ve $0 < q < 1$ dır. Sonuç olarak, toplam 6 sonuç vardır. Bunlar, dayanıklı ve dayanıksız olmak üzere λ parametresinin aldığı 0,0.05 ve 0.1 değerlerine karşılık gelmektedir. Huber M-tahmin edicisi de tuning parametresi $-3 < k < 3$ değerlerini aldığı anda dayanıklı M-tahmin edicileri verir. Aksi halde, dayanıklı M-tahmin edicisi verilmez. k 'nin aldığı bu farklı değerlere göre de M-tahmin edicisinin etkinliği ayarlanabilir. t dağılımda ise $\nu \rightarrow \infty$ iken dayanıksız EÇO tahmin edicileri elde edilirken $\nu \rightarrow 0$ iken dayanıklı EÇO tahmin edicileri elde edilir. Veri setini modellemede kullanılan amaç fonksiyonları Tablo 3.1'de verilmiştir.

Tablo 3.1. Veri setini modellemede kullanılan amaç fonksiyonları

	Kanikadis	t	Tsallis	Huber
Dayanıksız	$\kappa \rightarrow 0$	$\nu \rightarrow \infty$	$q \rightarrow 1$	$-\infty < k < -3$ ve $3 < k < \infty$
Dayanıklı	$\kappa > 0$	$\nu \rightarrow 0$	$0 < q < 1$	$-3 < k < 3$

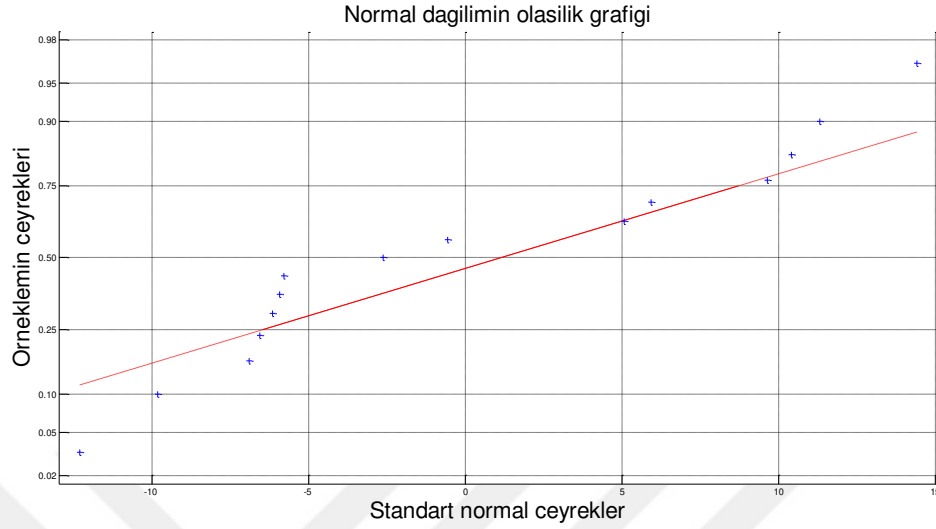
Gerçek veri setlerinde, doğrusal bir regresyon modeli düşünülmüştür. Gerekli hesaplamalar ve kodlamalar MATLAB R2013a ortamında yapılmıştır.

3.1 Gerçek Veri 1

Bu bölümde Hoerl ve Kennard'ın (1981, Ek Tablo 1) ürettiği veri analiz edilecektir. Açıklayıcı değişkenler ile bağımlı değişken arasında

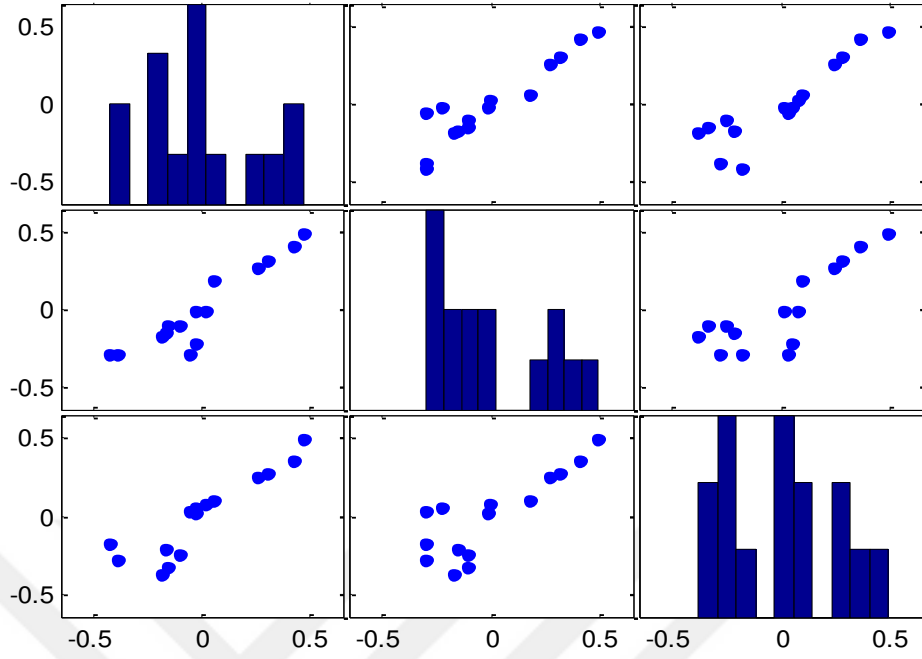
$$y_i = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_5 x_{5i} + \varepsilon_i, i = 1, 2, \dots, 15 \quad (3.1)$$

olmak üzere birden çok bağımsız değişkenin yer aldığı regresyon modeli tercih edilmiştir. Veri setleri Ek Tablo 1'de yer almaktadır.

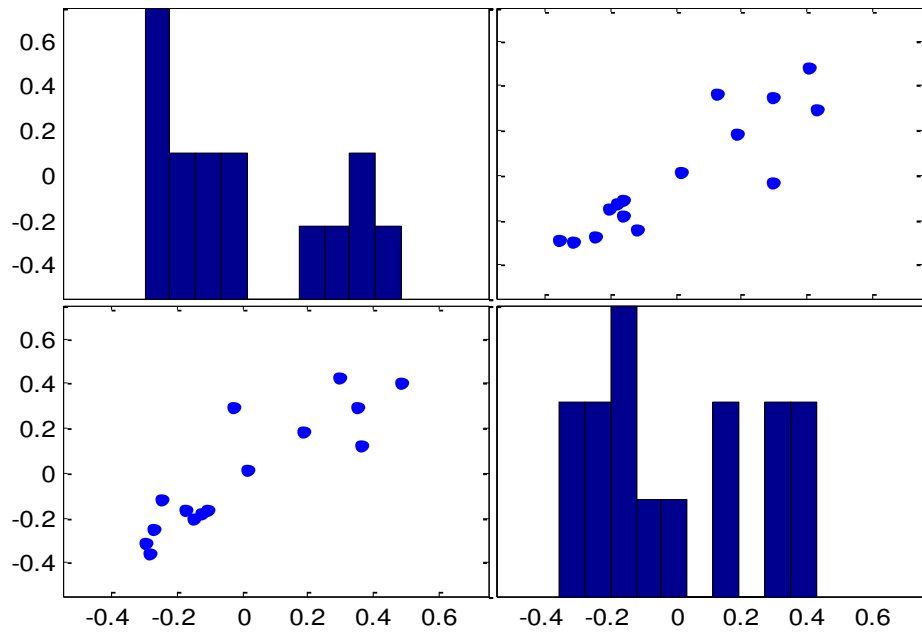


Şekil 3.1. Veri setinin normal dağılıma uyum sağlayıp sağlamadığının uygunluğun P-P grafiği ile incelenmesi

Veri sayısının $n=15$ olması sağlıklı karar vermede yeterli değildir. Düz çizgi normal dağılımın olasılık grafiğini; + şeklindeki imgeler gözlemleri temsil etmektedirler. + işaretlerin çizgiye yakınlığı yeterince gözlenmemektedir. Ancak, veri setinin normal dağılımdan gelip gelmediği uyum iyiliği test istatistiği ile test edilebilir. MATLAB R2013a ortamında; Kolmogorov test istatistiği ile uyum iyiliği yapıldığında, $p = 1.9176e-04$ olmak üzere veri setinin normal dağılımdan geldiğine dair yeterli kanıt yoktur. Dolayısı ile, farklı fonksiyonlar ile modelleme yapılması önerilir ve ilgili fonksiyonların başarısının bilgi kriterleri ile ölçülmesi gereklidir. (Çankayavd.,2019).



Şekil 3.2. x_1, x_2 ve x_3 açıklayıcı değişkenleri arasında doğrusal ilişkinin varlığının gösterimi ve her bir değişkenin histogramı



Şekil 3.3. x_4 ve x_5 açıklayıcı değişkenleri arasında doğrusal ilişkinin varlığının gösterimi ve her bir değişkenin histogramı

Şekiller 3.2 ve 3.3 incelendiğinde, değişkenler arasında doğrusal bir ilişki olduğu gözlenmektedir.

Tablo 3. 2. $X_1, X_2, \dots, X_5$ ve Y değişkenleri arasındaki korelasyon matrisi

X_1	X_2	X_3	X_4	X_5	Y
1	0.927163	0.906799	0.952479	0.857727	0.950248
0.927163	1	0.833904	0.972502	0.87046	0.965068
0.906799	0.833904	1	0.909541	0.821283	0.880922
0.952479	0.972502	0.909541	1	0.889113	0.965429
0.857727	0.87046	0.821283	0.889113	1	0.945048
0.950248	0.965068	0.880922	0.965429	0.945048	1

Tablo 3.2’de değişkenler arasındaki korelasyon değerlerinden oluşan korelasyon matrisi verilmiştir. Tabloya bakıldığında değişkenler arasındaki korelasyonların oldukça fazla olduğu görülmektedir.

Tablo 3.3. Dayanıksız M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa=10^{-20}$ t: $v=100$, Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=15$

	Normal	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	7.9567	7.9567	7.8496	7.9567	8.4648
$\widehat{\beta}_2$	16.6563	16.6563	17.1195	16.6563	17.2043
$\widehat{\beta}_3$	2.6446	2.6446	2.6306	2.6446	2.1825
$\widehat{\beta}_4$	-5.9090	-5.9090	-6.2930	-5.9090	-6.7249
$\widehat{\beta}_5$	12.3692	12.3692	12.5395	12.3692	12.6435
logL	58.0973	58.0973	58.1796	58.0973	58.1995
BIC	-20.9245	-20.9245	-20.9656	-20.9245	-20.9756

Tablo 3.3’de ridge regresyon sonucu elde edilen parametre değerleri verilmektedir. BIC değerleri dikkate alındığında Huber M-tahmin edicisi en iyi tahmini verdiği sonucuna varılmıştır.

Tablo 3.4. Dayanıksız M-tahmin değerleri ve $\lambda = 0.05$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k = 4$, $n = 15$

	Normal	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	6.8725	7.2286	6.1825	6.8725	6.1987
$\widehat{\beta}_2$	9.9813	11.2536	9.5898	9.9813	9.4285
$\widehat{\beta}_3$	1.5772	1.3581	1.8377	1.5772	1.7036
$\widehat{\beta}_4$	3.7309	1.9569	4.3613	3.7309	4.4150
$\widehat{\beta}_5$	10.8713	11.4724	10.7016	10.8713	10.6819
logL	-5.8991	-8.2540	-9.4400	-8.2540	-9.7955
BIC	28.0465	32.7563	35.1283	32.7563	35.8393

Tablo 3.4’de ise BIC değerlerine göre normal dağılımın konum parametresinin (regresyon) EÇO tahmin edicisi ya da EKK en iyi tahmini vermiştir. Burada görülmektedir ki, ridge parametresi λ tahminde etkili olmaktadır. Diğer bir açıdan, $X^T W X$ matrisinde çoklu doğrusal bağlantı problemi vardır. Çoklu doğrusal bağlantıdan arındırma işlemi gerçekleştirildiğinde; ridge kısıtının yer aldığı amaç fonksiyonunun başarılı modellediği gözlenmektedir.

Tablo 3.5. Dayanıksız M-tahmin değerleri ve $\lambda = 0.1$ iken Kaniadakis: $\kappa = 10^{-20}$ t: $\nu = 100$, Tsallis: $q=1-10^{-20}$, Huber: $k = 4$, $n = 15$

	Normal	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	6.5349	6.8725	5.8571	6.5349	5.7928
$\widehat{\beta}_2$	8.9347	9.9813	8.5575	8.9347	8.4006
$\widehat{\beta}_3$	2.2130	1.5772	2.6378	2.2130	2.6072
$\widehat{\beta}_4$	4.9518	3.7309	5.2847	4.9518	5.2681
$\widehat{\beta}_5$	10.0071	10.8713	9.6082	10.0071	9.5127
logL	5.8185	1.8300	-1.0502	1.8300	-2.2410
BIC	4.6113	12.5882	18.3486	12.5882	20.7303

Tablo 3.5’de ise EKK en iyi tahmini vermiştir. Burada, $\lambda = 0.1$ alınmış olmakla ridge parametresinin $X^T W X$ matrisi üzerindeki etkisi artırılmıştır. Böylelikle de, $\widehat{\beta}$ M-

tahmin edici deęerleri daha iyileřtirebilmiřtir. Bu ise, aynı zamanda açıklayıcı deęiřkenler $x_1, x_2, \dots, x_5$ arasında çoklu doęrusal baęlantının varlıęını göstermektedir. Dayanıksız bir yaklařım dikkate alındıęında, önerilen ise arařtırmacının $\hat{\beta}$ tahmin deęerlerinin sayısal büyüklüęü ve iřaretine göre uygun λ deęeri ayarlamaya çalışmasıdır. Tabi aynı zamanda, arařtırmacının BIC kriterini kullanılması da önerilir.

Tablo 3.6. Dayanıklı M-tahmin deęerleri ve $\lambda = 0$ iken Kaniadakis: $\kappa = 0.1$, t : $\nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$

	Kaniadakis	t	Tsallis	Huber
$\hat{\beta}_1$	13.9601	13.7269	23.5133	10.2073
$\hat{\beta}_2$	13.8257	17.2983	14.6458	17.5946
$\hat{\beta}_3$	-7.3327	-1.7072	-5.9251	1.0042
$\hat{\beta}_4$	3.2819	-8.4178	-9.1798	-7.5049
$\hat{\beta}_5$	7.5696	12.2434	14.0584	12.1889
logL	-20.9245	-25.6194	-38.1348	-21.3663
BIC	55.3893	64.7790	89.8098	56.2728

Tablo 3.6’da en küçük BIC deęeri Kaniadakis M-tahmin edicisi tarafından elde edilmektedir. Belirtmek gerekir ki, bir sonraki en küçük deęer Huber M-tahmin edicisidir ve bu deęer Kaniadakis’in verdięi BIC deęerine çok yakındır. Dolayısıyla, deforme edilmiř bir logaritma olan $\log_\kappa$ aracılıęı ile daha başarılı bir modelleme süreci gerçekeřtirilebilmiřtir. Ancak, ridge parametre deęeri sıfır iken ($\lambda = 0$), ridge regresyon etkisinin olmadıęı dikkatten kaçmamalıdır.

Tablo 3.7. Dayanıklı M-tahmin değerleri ve $\lambda = 0.05$ iken Kaniadakis: $\kappa = 0.1$, $t: \nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$

	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	14.0644	4.3831	2.9536	5.7748
$\widehat{\beta}_2$	13.2123	5.6840	6.4001	7.5654
$\widehat{\beta}_3$	-7.2667	3.1278	2.2768	3.0407
$\widehat{\beta}_4$	3.7934	4.4103	4.0488	5.4218
$\widehat{\beta}_5$	7.5036	5.8268	7.9572	7.9140
logL	-8.2540	-40.8480	-61.5640	-21.4541
BIC	30.0482	95.2362	136.6683	56.4485

Tablo 3.7’de en küçük BIC değeri Kaniadakis M- tahmin edicisi tarafından elde edilmiştir. Ridge parametresinin ($\lambda = 0.05$) etkili olduğu durumda deforme edilmiş bir logaritma olan $\log_{\kappa}$ aracılığı ile belirgin bir şekilde daha başarılı bir modelleme süreci gerçekleştirilebilmiştir. Benzer sonuçlar ($\lambda = 0.1$) için de elde edilmiş ve Tablo 3.8’de verilmiştir.

Tablo 3.8. Dayanıklı M-tahmin değerleri ve $\lambda = 0.1$ iken Kaniadakis: $\kappa = 0.1$, $t: \nu = 1.5$, Tsallis: $q = 0.9$, Huber: $k = 1$, $n = 15$

	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	14.1600	3.3961	2.4659	5.1696
$\widehat{\beta}_2$	12.7242	4.2376	4.8775	6.3911
$\widehat{\beta}_3$	-7.1855	3.0536	2.7827	3.7042
$\widehat{\beta}_4$	4.1532	3.6588	3.6460	5.2565
$\widehat{\beta}_5$	7.4585	4.5126	6.2039	6.6737
logL	1.8300	-48.0063	-99.9660	-28.7070
BIC	9.8802	109.5528	213.4723	70.9543

3.2 Gerçek Veri 2

Pamukçu tarafından 2009'da derlenen Sistolik kan basıncını etkileyebilecek olan muhtemel değişkenlerin yer aldığı veri seti analiz edilmiştir. Bu açıklayıcı değişkenler ve bağımlı değişken Tablo 3.9 'da verilmiştir. Açıklayıcı değişkenler ile bağımlı değişken arasında

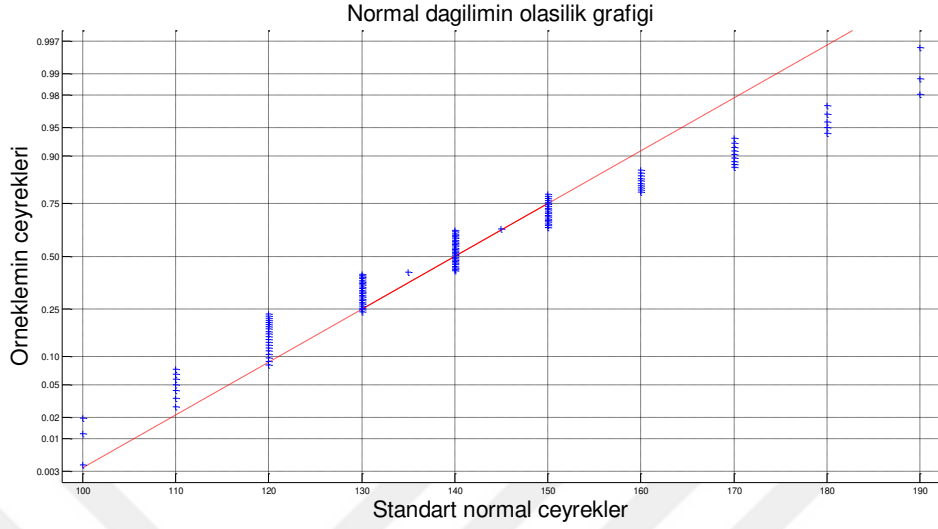
$$y_i = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_{12} x_{12i} + \varepsilon_i, i = 1, 2, \dots, 128 \quad (3.2)$$

olmak üzere birden çok bağımsız değişkenin yer aldığı regresyon modeli tercih edilmiştir. Veri setleri Ek Tablo 2'de yer almaktadır. Birden çok bağımsız değişken ele alındığında, ilgili bağımsız değişkenler arasındaki doğrusal ilişkide eğer varsa $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{12}$ tahmin değerlerinin yukarı yönlü yanlı olmasını tetikler. Ridge parametresi ile bu durumunun üstesinden gelinebilir. Çalışmada sistolik kan basıncını tahmin etmek için kullanılan değişkenler ve ilgili ölçü birimleri aşağıdaki tabloda verilmektedir (Pamukçu 2009).

Tablo 3. 9. Regresyon modelinde kullanılan değişkenler ve her birinin ilgili birimleri

Değişkenin Türü	Değişkenler	Ölçü Birimi
Bağımsız	Yaş	Yıl
	Boy	m
	Karın Çevresi	m
	Glikoz	mg/dl
	HDL	mg/dl
	LDL	mg/dl
	Üre	mg/dl
	Kreatinin	mg/dl
	Potasyum	mg/dl
	HGB	mg/dl
	HCT	mg/dl
	Diastolik kan basıncı	mm Hg
Bağımlı	Sistolik kan basıncı	mm Hg

Tablo 3.9'da verilen birimler üzerinden regresyon modeli oluşturulmuştur.



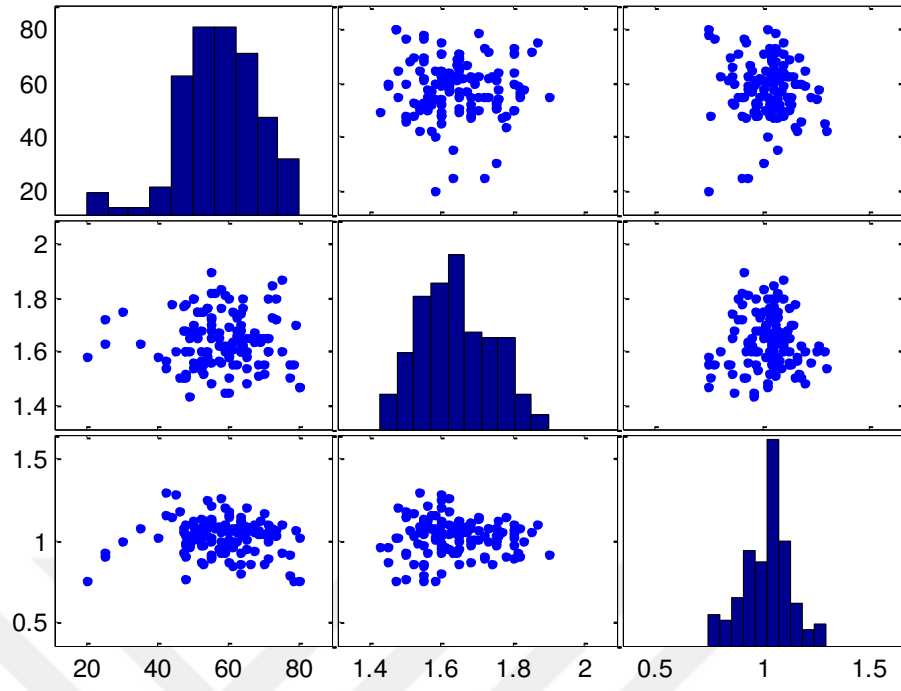
Şekil 3.4. Veri setinin normal dağılıma uyum sağlayıp sağlamadığının uygunluğun P-P grafiği ile incelenmesi

Şekil 3.4 incelendiğinde, sağ taraftaki gözlemlerin normal dağılımın olasılık değeri olan çizgiden önemli derecede uzaklaştığı görülmektedir. Veri sayısının $n=128$ olması sağlıklı karar vermede yeterlidir. Ancak, veri setinin normal dağılımdan gelip gelmediği uyum iyiliği test istatistiği ile test edilebilir. MATLAB R2013a ortamında Kolmogorov test istatistiği ile uyum iyiliği yapıldığında, $p = 7.5755e-114$ olmak üzere veri setinin normal dağılımdan geldiğine dair yeterli kanıt yoktur. Dolayısı ile, farklı fonksiyonlar ile modelleme yapılması önerilir ve ilgili fonksiyonların başarısının bilgi kriterleri ile ölçülmesi gereklidir.

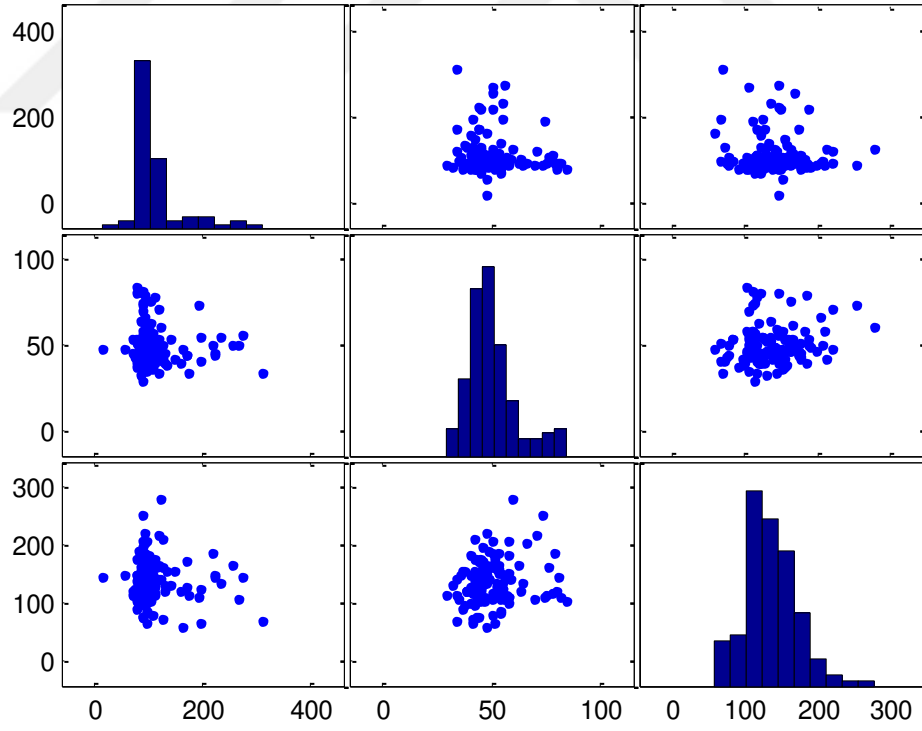
Tablo 3.10. Yaş,Boy, Karın çevresi, Glukoz, HDL, LDL, Üre, Kreatinin, Potasyum, HGB, HCT, Diastolik ve Sistolik değişkenleri arasındaki korelasyon matrisi

Yaş	Boy	Kar çev	Glukoz	HDL	LDL	Üre	Kreatinin	Potasyum	HGB	HCT	Diastolik	Sistolik
1	-0.03082	-0.05663	0.020648	-0.1015	0.036871	-0.00052	-0.1651	0.027122	0.089201	0.111348	-0.05357	0.012794
-0.03082	1	0.022362	0.058706	-0.30192	-0.17625	0.114503	0.119113	0.151822	0.266359	0.234008	-0.12355	-0.09893
-0.05663	0.022362	1	0.104124	0.004312	0.136967	-0.18345	-0.08985	-0.08888	0.160872	0.183033	0.040481	0.196741
0.020648	0.058706	0.104124	1	-0.09519	-0.07827	0.092706	0.077944	0.120084	-0.12058	-0.10208	0.063041	0.080953
-0.1015	-0.30192	0.004312	-0.09519	1	0.155548	-0.12458	-0.11012	-0.01231	-0.06383	-0.01236	-0.04653	-0.08528
0.036871	-0.17625	0.136967	-0.07827	0.155548	1	-0.1472	-0.16111	-0.00823	0.108821	0.114003	0.047561	0.077899
-0.00052	0.114503	-0.18345	0.092706	-0.12458	-0.1472	1	0.856513	0.116248	-0.39219	-0.42521	-0.10343	-0.04058
-0.1651	0.119113	-0.08985	0.077944	-0.11012	-0.16111	0.856513	1	-0.0536	-0.33736	-0.37587	-0.06988	0.027132
0.027122	0.151822	-0.08888	0.120084	-0.01231	-0.00823	0.116248	-0.0536	1	-0.05193	-0.0368	-0.11412	-0.11082
0.089201	0.266359	0.160872	-0.12058	-0.06383	0.108821	-0.39219	-0.33736	-0.05193	1	0.974211	0.038775	0.034962
0.111348	0.234008	0.183033	-0.10208	-0.01236	0.114003	-0.42521	-0.37587	-0.0368	0.974211	1	0.042018	0.050817
-0.05357	-0.12355	0.040481	0.063041	-0.04653	0.047561	-0.10343	-0.06988	-0.11412	0.038775	0.042018	1	0.741741
0.012794	-0.09893	0.196741	0.080953	-0.08528	0.077899	-0.04058	0.027132	-0.11082	0.034962	0.050817	0.741741	1

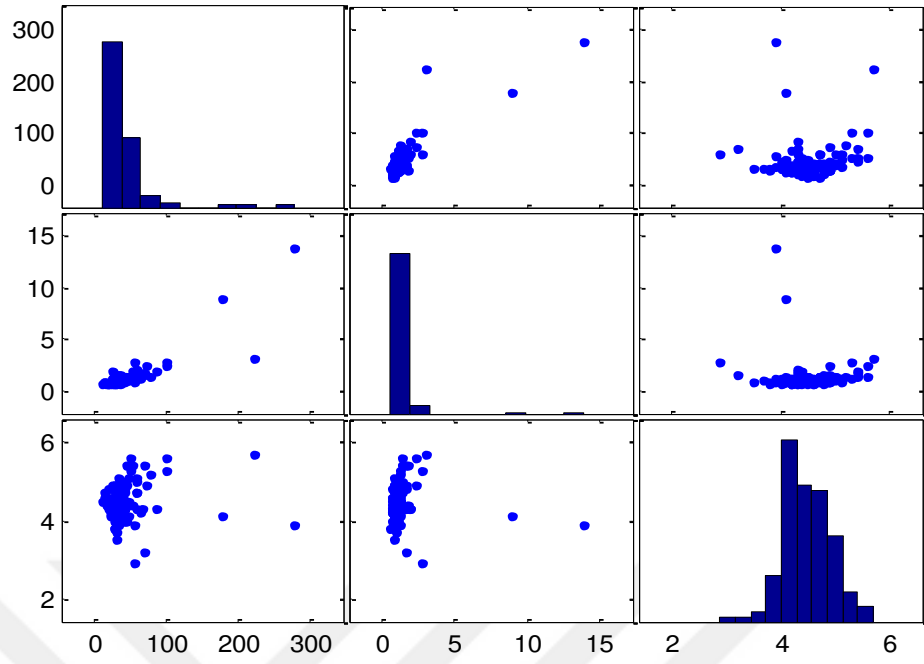
Tablo 3.10 değişkenleri arasındaki doğrusal ilişkiyi görmek için oluşturulmuştur. Böylelikle de, bu ilişkinin sayısal büyüklüğünün ne kadar olduğu gözlenmiştir. Yüksek ilişkiler koyu ile gösterilmiştir.



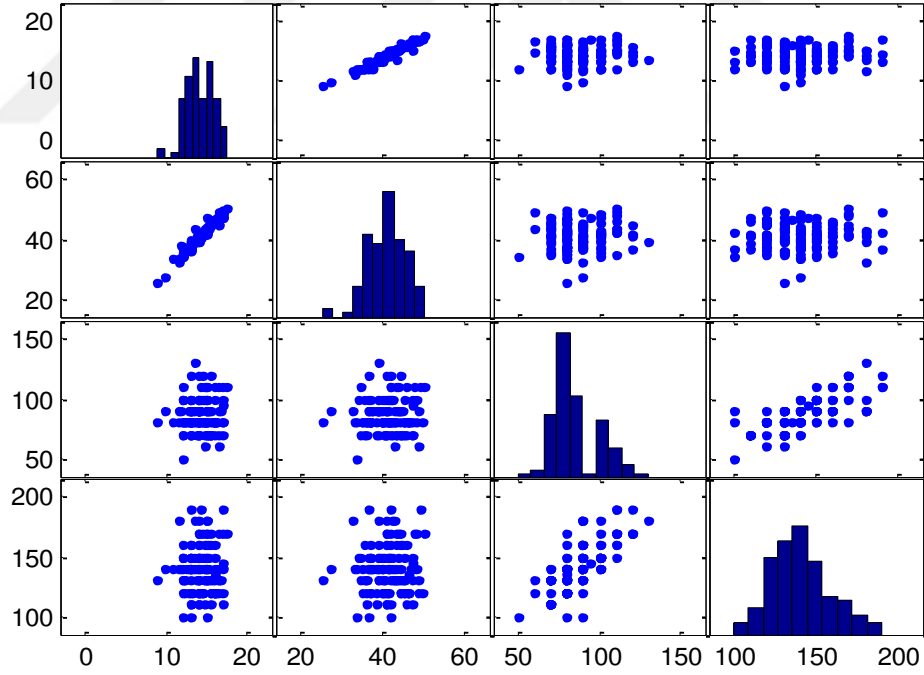
Şekil 3.5. Baş, boy ve karın çevresi değişkenleri arasındaki korelasyon ve histogramları



Şekil 3.6. Glukoz, HDL ve LDL değişkenleri arasındaki korelasyon ve histogramları



Şekil 3.7. Üre, Kreatinin ve Potasyum değişkenleri arasındaki korelasyon ve histogramları



Şekil 3.8. HGB, HCT, Diastolik ve Sistolik değişkenleri arasındaki korelasyon ve histogramları

Tablo 3.11. Dayanıksız M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa=10^{-20}$ t: $\nu = 100$,
Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=128$

	Normal	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	0.1640	0.1640	0.1605	0.1640	0.1623
$\widehat{\beta}_2$	-1.6712	-1.6712	0.6172	-1.6712	-0.4801
$\widehat{\beta}_3$	31.1907	31.1907	33.4141	31.1907	32.1656
$\widehat{\beta}_4$	0.0015	0.0015	0.0050	0.0015	0.0032
$\widehat{\beta}_5$	-0.0870	-0.0870	-0.0556	-0.0870	-0.0729
$\widehat{\beta}_6$	0.0230	0.0230	0.0051	0.0230	0.0147
$\widehat{\beta}_7$	-0.0629	-0.0629	-0.0610	-0.0629	-0.0605
$\widehat{\beta}_8$	3.1941	3.1941	2.6062	3.1941	2.8632
$\widehat{\beta}_9$	0.9283	0.9283	1.3491	0.9283	1.0966
$\widehat{\beta}_{10}$	-3.4969	-3.4969	-5.5546	-3.4969	-4.5397
$\widehat{\beta}_{11}$	1.3322	1.3322	1.8898	1.3322	1.6128
$\widehat{\beta}_{12}$	1.0627	1.0627	1.0351	1.0627	1.0531
logL	-10253.7	-10253.7	-4638.83	-10253.7	-10364.1
BIC	20565.6	20565.6	9335.88	20565.6	20786.4

Tablo 3.11’de t-dağılımının regresyon (konum) parametrelerinin EÇÖ tahmin edicileri en iyi tahmin değerlerini verdiği gösterilmiştir.

Tablo 3.12. Dayanıklı M-tahmin değerleri ve $\lambda=0$ iken Kaniadakis: $\kappa=10^{-4}$ t: $\nu = 1.5$, Tsallis: $q=1-10^{-4}$, Huber: $k=1.5$, $n=128$

	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	0.1704	0.1605	0.1636	0.1623
$\widehat{\beta}_2$	-5.7289	0.6204	2.0236	-0.4801
$\widehat{\beta}_3$	32.5438	33.4395	31.8288	32.1656
$\widehat{\beta}_4$	-0.0027	0.0050	0.0059	0.0032
$\widehat{\beta}_5$	-0.1130	-0.0554	-0.0546	-0.0729
$\widehat{\beta}_6$	0.0275	0.0049	0.0090	0.0147
$\widehat{\beta}_7$	-0.0832	-0.0611	-0.0537	-0.0605
$\widehat{\beta}_8$	4.1322	2.6052	2.3827	2.8632
$\widehat{\beta}_9$	0.9046	1.3529	1.2491	1.0966
$\widehat{\beta}_{10}$	-1.8940	-5.5668	-5.3476	-4.5397
$\widehat{\beta}_{11}$	0.9141	1.8932	1.8178	1.6128
$\widehat{\beta}_{12}$	1.0851	1.0349	1.0228	1.0531
logL	-10253.7	-732.33	-10699.2	-10364.1
BIC	20565.6	1522.88	21456.7	20786.4

Tablo 3.12’de t-dağılımının regresyon (konum) parametrelerinin EÇO tahmin edicileri en iyi tahmin değerlerini verdiği gösterilmiştir.

Tablo 3.13. Dayanaksız M-tahmin değerleri ve $\lambda=0.1$ iken Kaniadakis: $\kappa=10^{-20}$ t: $\nu=100$, Tsallis: $q=1-10^{-20}$, Huber: $k=4$, $n=128$

	Normal	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	0.1649	0.1644	0.1795	0.1649	0.1782
$\widehat{\beta}_2$	-1.0410	-1.3373	0.9838	-1.0410	1.9926
$\widehat{\beta}_3$	29.2453	30.1840	3.0930	29.2453	10.5484
$\widehat{\beta}_4$	0.0022	0.0019	0.0235	0.0022	0.0134
$\widehat{\beta}_5$	-0.0846	-0.0858	0.0157	-0.0846	-0.0392
$\widehat{\beta}_6$	0.0241	0.0236	0.0171	0.0241	0.0249
$\widehat{\beta}_7$	-0.0649	-0.0639	-0.0606	-0.0649	-0.0789
$\widehat{\beta}_8$	3.2357	3.2155	2.6607	3.2357	3.3356
$\widehat{\beta}_9$	0.9444	0.9356	1.7686	0.9444	1.7274
$\widehat{\beta}_{10}$	-3.5202	-3.5099	-2.5938	-3.5202	-3.9887
$\widehat{\beta}_{11}$	1.3487	1.3409	1.2345	1.3487	1.6085
$\widehat{\beta}_{12}$	1.0646	1.0637	1.0897	1.0646	1.0824
logL	-10160.5	-10168.2	-4768.6	-10168.2	-10672.2
BIC	20379.2	20394.7	9595.4	20394.7	21402.7

Tablo 3.13’de t-dağılımının regresyon (konum) parametrelerinin EÇO tahmin edicileri en iyi tahmin değerlerini vermiştir. Çünkü BIC değeri en küçük değer t dağılımı kullanıldığında elde edilmektedir. Normal, Kaniadakis ve Tsallis kullanılarak elde edilen tahmin ediciler birbirine yakın tahmin değerleri vermiştir. Bu ise, beklenen bir sonuçtur. Çünkü, $\kappa=10^{-20}$ ve $q=1-10^{-20}$ alınmış olması normal dağılımın regresyon parametrelerini EÇO tahmin edicilerine karşılık getirir. Ridge parametre değeri $\lambda=0.1$ 'dir

Tablo 3.14. Dayanıklı M-tahmin değerleri ve $\lambda=0.1$ iken Kaniadakis: $\kappa=10^{-4}$ t: $\nu=1.5$, Tsallis: $q=1-10^{-4}$, Huber: $k=1.5$, $n=128$

	Kaniadakis	t	Tsallis	Huber
$\widehat{\beta}_1$	0.1706	0.1669	0.1665	0.1822
$\widehat{\beta}_2$	-5.4289	0.0561	2.8609	1.3561
$\widehat{\beta}_3$	31.8323	0.1043	27.4340	5.4268
$\widehat{\beta}_4$	-0.0024	0.0356	0.0076	0.0174
$\widehat{\beta}_5$	-0.1122	0.0744	-0.0493	-0.0222
$\widehat{\beta}_6$	0.0280	0.0217	0.0110	0.0271
$\widehat{\beta}_7$	-0.0838	0.0364	-0.0590	-0.072
$\widehat{\beta}_8$	4.1455	0.2405	2.5052	3.1981
$\widehat{\beta}_9$	0.8999	0.1341	1.3655	1.7586
$\widehat{\beta}_{10}$	-1.9044	0.0377	-5.3879	-3.0422
$\widehat{\beta}_{11}$	0.9196	0.4632	1.8632	1.3596
$\widehat{\beta}_{12}$	1.0856	1.0999	1.0282	1.0933
logL	-10168.2	-721.7	-10628.2	-10889
BIC	20394.7	1501.63	21314.7	21836.2

Tablo 3.14’de t-dağılımının regresyon (konum) parametrelerinin EÇO tahmin edicileri en iyi tahmin değerlerini vermiştir. Ridge ve sapan gözlem birlikte düşünüldüğünde, t dağılımının regresyon parametrelerinin EÇO tahmin edicileri iyi sonuç vermiştir.

4. SONUÇLAR VE TARTIŞMA

Çalışmada ilk olarak EÇO tahmin yöntemi ve bunun genelleştirmesi olan M-tahmin edicileri ele alınmıştır. M-tahmin yönteminde yer alan genelleştirilmiş logaritmalar Tsallis $\log_q$ ve Kaniakadis $\log_\kappa$ fonksiyonları EÇO tahmin yöntemindeki $\log$ fonksiyonu yerine kullanılmıştır. İlgili fonksiyonların regresyon (konum) modeline göre M-tahmin yönteminde türevleri alındığında, getirdikleri ağırlık fonksiyonları aracılığı ile sapan gözlemlere küçük ya da sıfıra yakın değerler verdikleri için $\widehat{\beta}_M$, ${}_q\widehat{\beta}_M$, ${}_\kappa\widehat{\beta}_M$ M-tahmin edicilerinin sapan gözleminden etkilenmemesini sağlamışlardır. Bu tahmin edicilerin ridge versiyonları da benzer şekilde sapan gözleminden etkilenmemiştir. Çünkü, sapan gözlem durumunda yine ağırlık fonksiyonu işlevini yerine getirmektedir.

İkinci olarak Ridge regresyon yönteminde, birinci bölümde ele alınan yaklaşımın aynısı takip edilmektedir. Değişen yapı ise, amaç fonksiyonuna $\sum_{i=1}^n \Lambda(\varepsilon_i; \beta)$ yapılan ekleme veya konulan kısıttır. Bu kısıt dikkate alındığında, $\sum_{i=1}^n \Lambda(\varepsilon_i; \beta) + \lambda \beta^T \beta$ şeklindeki yeni bir amaç fonksiyonunun β parametre vektörüne göre optimizasyonu söz konusudur.

Ridge regresyon ile deforme edilmiş logaritmalar kullanılarak oluşturulan M-tahmin yöntemi birleştirilmiştir. Böylelikle, ridge deforme edilmiş M-tahmin yöntemi adı verilen yaklaşım geliştirilmiştir. Sonuç olarak; q ve κ deforme parametreleri ve λ ridge parametresi aynı anda kullanılmaktadır. Bu parametrelerin aynı anda kullanılması, dayanıklı ve yansız M-tahmin edicilerinin elde edilmesini sağlamaktadır. İleri ki çalışmalar olarak, bu tahmin edicilerin asimptotik dağılımları ve kovaryansları ile ilgilenilecektir. Aynı zamanda, teorik olarak elde edilen kovaryans yapıları simülasyon ile karşılaştırmalı olarak verilecektir.

5. ÖNERİLER

1. Yansızlığı ayarlamak için farklı ridge λ ile Tsallis q , Kaniakadis κ , Student-t ν ve Huber k dayanıklılığı ayarlama parametre değerleri seçilerek; regresyon parametreleri $\beta_1, \beta_2, \dots, \beta_p$ için farklı değerlerde $\widehat{\beta}_1, \widehat{\beta}_2, \dots, \widehat{\beta}_p$ tahmin değerleri elde edilebilir.
2. Farklı tahmin değerlerine sahip olmak regresyon katsayılarının işaretlerini görmeyi ve araştırmacının amacına uygun olanını seçebilmeyi sağlar. Ancak, her ne kadar en küçük bilgi kriterini veren fonksiyonun seçilmesi önerilse de, uygulamalı bilimlerdeki araştırmacılar regresyon parametre değerlerinin farklı sayısal değerlerine ve/veya işaretlerine ihtiyaç duyabilirler.
3. Gerek Kanikadis ve gerekse de Tsallis deforme logaritmalar kullanmak; ilgili olasılık yoğunluk fonksiyonuna yakın değerleri meydana getireceği için, uygulamalı sahadaki bilimcinin bu deforme logaritmalar ile çalışması önerilebilir.
4. $X^T W X$ matrisinin tersi bu matristeki çoklu bağlantı probleminin derecesine göre, sıfıra yaklaşma göstermektedir. Ridge regresyon tanımından gelen λ parametresinin kullanılması ile sıfıra yakınlık derecesinden uzaklaşılabilir. Araştırmacının amacına göre, λ parametresinin değeri seçilebilir ya da deforme edilmiş logaritma temelli ridge parametresinin eş anlı tahmini süreci gerçekleştirilebilir.

KAYNAKLAR

- [1] **Acıtaş, S., Kasap, P., Şenoğlu, B., and Arslan, O.** 2013. One-step M-estimators: Jones and Faddy's skewed t-distribution. *Journal of Applied Statistics*, Vol. **40(7)**, 1545-1560.
- [2] **Akaike, H.** 1974. A new Look at the Statistical Model Identification. *IEEE Transactions Automatic Control*, Vol. **19(6)**, 716-723.
- [3] **Akdeniz, F. ve Çabuk, A.** Ridge regresyon teorisinde 1970-2001 arası gelişmeler. <http://idari.cu.edu.tr/sempozyum/bil24.htm>.
- [4] **Aktaş, C. ve Yılmaz, V.** 2003. Çoklu bağlantılı modellerde Liu ve Ridge regresyon kestiricilerinin karşılaştırılması. *Anadolu Üniversitesi Bilim Ve Teknoloji Dergisi*, **4, 2**, 189-194
- [5] **Andrews, D.F., Bickel, P.J., Hampel, F.R. and Huber, P.J.** 1972. Robust Estimates of Location: Survey and Advances. *Princeton University Press Princeton*, 373, USA.
- [6] **Arslan, O.** 2004. Convergence behavior of an iterative reweighting algorithm to compute multivariate M-estimates for location and scatter. *Journal of Statistical Planning and Inference*, Vol. **118(1)**, 115 - 128.
- [7] **Arslan, O., and Genc, A. I.** (2009). The skew generalized t distribution as the scale mixture of a skew exponential power distribution and its applications in robust estimation. *Statistics*, **43(5)**, 481-498.
- [8] **Arslan, O.** (2016). Penalized MM regression estimation with L_γ penalty: a robust version of bridge regression. *Statistics*, **50(6)**, 1236-1260.
- [9] **Albayrak, A., S.,** 2005. Çoklu doğrusal bağlantı halinde en küçük kareler tekniğinin alternatifi yanlı tahmin teknikleri ve bir uygulama. *ZKÜ Sosyal Bilimler Dergisi*, **1, 1**.
- [10] **Brazauskas, R. B. V.** (2008). Method of trimmed moments for robust fitting of parametric failure time models. *Metron*, **66(3)**, 341-360

- [11] Çankaya, M. N., Yalçınkaya, A., Altındağ, Ö., And Arslan, O. (2019). On the robustness of an epsilon skew extension for Burr III distribution on the real line. *Computational Statistics*, 1-27.
- [12] Çankaya, M. N., and Korbel, J. (2018). Least informative distributions in maximum q-log-likelihood estimation. *Physica A: Statistical Mechanics and its Applications*, **509**, 140-150.
- [13] Çankaya, M. N., and Korbel, J. (2017). On statistical properties of Jizba–Arimitsu hybrid entropy. *Physica A: Statistical Mechanics and its Applications*, **475**, 1-10.
- [14] Duran, E.A. 2011 Regresyon analizinde çoklu bağlantı: Parametrik ve semiparametrik tahmin, *Doktora Tezi*, Gazi Üniversitesi. Fen Bilimleri Enstitüsü, Ankara.
- [15] Farebrother, R. W., 1984. The Restricted Least Squares Estimator and Ridge Regression. *Comm. Statist. Theory Methods*, **13**, 191-196.
- [16] Farrar, D. E. and Glauber, R. R., 1967. Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics*, **49**, **1**, 92-107.
- [17] Ferrari, D., And Paterlini, S. (2009). The maximum Lq-likelihood method: An application to extreme quantile estimation in finance. *Methodology and Computing in Applied Probability*, **11**(1), 3-19.
- [18] Ferrari, D., And Yang, Y. (2010). Maximum Lq-likelihood estimation. *The Annals of Statistics*, **38**(2), 753-783.
- [19] Fisher, R. A. (1925, July). Theory of statistical estimation. *In Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. **22**, No. **5**, pp. 700-725). Cambridge University Press.
- [20] Frank, L. E., And Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, **35**(2), 109-135.
- [21] FuWJ. Penalized regression: the bridge versus lasso. *J Comput Graph Statist.* **1998**;7:397–416.

- [22] **Freund, J.,E., And Çev: Şenesen,Ü.** 2001. Matematiksel İstatistik. Literatür Yayıncılık, İstanbul.
- [23] **Frisch, R.,** 1934. Statistical Confluence Analysis by Means of Complete Regression Systems. Universitets Okonomiske Institut, Oslo.
- [24] **Gosset, W. S.** (1908). Student. *The Application of the 'Law of Error' to the Work of the Brewery*, 3-6
- [25] **Gujarati,D.,N. , And Çev: Şenesen,Ü, Şenesen,G.,G.** 1999. Temel Ekonometri. Literatür Yayıncılık. İstanbul.
- [26] **Hoerl, A.,E. and Kennard, R.,W.** 1970a. Ridge regression: biased estimation for non-orthogonal problems. *Technometrics*, **12**,1,55-67.
- [27] **Haberman, S.J.** 1989. Concavity and Estimation. *The Annals of Statistics. JSTOR, Vol.17(4)*, 1631-1661.
- [28] **Hampel, F.R.,Ronchetti, E.M., Rousseeuw, P.J. andStahel, W.A.** 1986. Robust Statistics: The Approach Based on Influence Functions. *Wiley Series in Probability and Statistics*, 36-39, New York.
- [29] **Huber, P.J.** 1981. Robust Statistics. *Wiley Series in Probability and Statistics*, 308, New York.
- [30] **Huber, P.J. And Ronchetti, E.M.** 2009. Robust Statistics. *Second Edition, Wiley Series in Probability and Statistics*, 354, USA.
- [32] **Hoerl, A. E., Kannard, R. W., And Baldwin, K. F.** (1975). Ridge regression: some simulations. *Communications in Statistics-Theory and Methods*, **4(2)**, 105-123.
- [33] **Hua, T. A., And Gunst, R. F.** (1983). Generalized ridge regression: a note on negative ridge parameters. *Communications in Statistics-Theory and Methods*, **12(1)**, 37-45.
- [34] **Kaniadakis, G., Lissia, M., and Scarfone, A. M.** (2005). Two-parameter deformations of logarithm, exponential, and entropy: A consistent framework for generalized statistical mechanics. *Physical Review E*, 71(4), 046128.

- [35] **Khalaf, G. And Shukur, G.** (2005). Choosing ridge parameter for regression problems.
- [36] **Li, B. And Yu, Q.** Robust and sparse bridge regression. *Statist Interface*. **2009**;2:481–491.
- [37] **Li, G., Peng, H., And Zhu, L.** (2011). Nonconcave penalized M-estimation with a diverging number of parameters. *Statistica Sinica*, 391-419.
- [38] **Lipovetsky, S. and Conklin, W. M.,** 2001. Multiobjecti ve Regression Modificationsfor Collinearity. *Computers and Operations Research*, **28**,(1333-1345).
- [39] **Çankaya, M.** (2018). Asymmetric bimodal exponential power distribution on the real line. *Entropy*, **20**(1), 23.
- [40] **Marquardt, D.W.** 1970. Generalized inverses ridge regression, biased linear estimation and nonlinear estimation, *Technometrics*, **12**, 3,591-612.
- [41] **Marquardt, D.W. and Snee, R.,D.** 1975. Ridge Regression in Practice. *JASA Journal of The American Statistical Association*, **29**, 1, 3-20.
- [42] **Mcdonald, G.,C. and Galarneau, D.,I.** 1975. A monte carlo evaluation of some ridge type estimators. *JASA Journal of the American Statistical Assocation*, **70**, 350, 407-416.
- [43] **Montgomery,D., And Peck, E.** 1991. *Introduction to Linear Regression Analysis*, New York.
- [44] **Montgomery,D., And Peck,E.**1992. Introduction to Linear Regression Analysis. *John Wiley and Sons*, New York, 526p.
- [45] **Özkale, M.,R.** 2007.Çoklu iç ilişki ile ilgili problemler, *Doktora tezi*, Çukurova Üniversitesi Fen Bilimleri Enstitüsü, Adana.
- [46] **Öztürk, F.** (2015). Parametre tahmini ve hipotez testi. Bıçaklar Kitabevi.

- [47] **Pamukçu, E.** 2009. Sistolik kan basıncını etkileyebilecek faktörlerin ridge regresyon analizi ile incelenmesi ve çoklu bağlantı problemi. *Yüksek Lisans Tezi*. Fırat Üniversitesi Fen Bilimleri Enstitüsü. Elazığ
- [48] **Rao, C. R. and Toutenburg, H.,**1995. Linear Models: Least Squares and Alternatives. *Springer*, New York, 351p.
- [49] **Rousseeuw, P. J.** (1984). Least median of squares regression. *Journal of the American statistical association*, **79(388)**, 871-880.
- [50] **Silvey, S. D.,** 1969. Multicollinearity and Imprecise Estimation. *Journal of the Royal Statistical Society, Series B (Methodological)*, **31, 3**, 539-552.
- [51] **SheY, And OwenAB.** Outlier detection using nonconvex penalized regression. *J Amer. Statist Assoc.* **2011;106**:626–639.
- [52] **Tsallis, C.** (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, **52(1-2)**, 479-487.
- [53] **Vinod, H. D. ve Ullah, A.,** “Recent Advances in Regression Methods”, *Marcel Dekker Inc., New York*, 30-45 (1981).
- [54] **Wetherill, G.,B.** 1986. *Regression Analysis with Applications*, Chapman and Holl, New York.
- [55] **Weakliem, D.L.** 1999. A critique of the Bayesian information criterion for model selection. *Sociological Methods Research*, Vol.**27(3)**, 359-397.
- [56] **Wang, H.,Li, R., &Tsai, C. L.** (2007). Tuning parameter selectors for the smoothly clipped absolute deviation method. *Biometrika*, **94(3)**, 553-568.
- [57] **Xu, J. and Ying, Z.,** (2010). Simultaneous estimation and variable selection in median regression using Lasso-type penalty. *Annals of the Institute of Statistical Mathematics*, **62**, 487- 514
- [58] **Güney, Y.** 2018. Simetrik ve çarpık dağılımlar için bilgi karmaşıklık (ICOMP) kriteri kullanılarak robust model seçimi, *Doktora Tezi* ,Ankara Üniversitesi. Fen Bilimleri Enstitüsü, Ankara.

EKLER

Verilen x veri setinin korelasyon ve kovaryans değerlerini sırasıyla MATLAB2013a da yer alan 'corr' ve 'cov' fonksiyonları aracılığı ile hesaplar.

```
boyut=size(x,2)
for j=1:boyut
for i=1:boyut
    x1=x(:,j);
    x2=x(:,i);
    kore(j,i)=corr(x1,x2);
    kov(:,j,i)=cov(x1,x2);
end
end
kore
kov
```

IRA yürüten adımlar

```
Farkesh=5;iter=0.01;
while norm(Farkesh) > iter
b(:,k)=wreg(b(:,k-1),x,y,lm);
fb(k)=b(k)-b(k-1);
    b(k-1)=b(k);
Farkesh=fb(k);
ktesh=ktesh+1;
end
```

Aşağıda Kaniakadis $\log_k$ fonksiyonu kullanılarak elde edilen regresyon parametrelerinin tahminini hesaplayan program parçası vardır.

```
leb=length(b);n=length(y);b=zeros(leb,1);
function freg=wreg(b,x,y,k,lm)
idex=1;wt1=zeros(1,n);wt2=zeros(1,n);u=zeros(1,n);
sspm1=zeros(leb,1);sspm2=zeros(leb,1);ssdm1=zeros(leb,leb);ssdm2=zeros(leb,leb);
for i=1:n
    u(i)=(abs(y(i)-z(i,:)*b));wt1(idex,i)=exp(-k*u(i)^2/2);wt2(idex,i)=exp(k*u(i)^2/2);
    sspm1(:,idex)=sspm1(:,idex)+wt1(idex,i)*x(i,:)'*y(i);
```

```

sspm2(:,idex)=sspm2(:,idex)+wt2(idex,i)*x(i,:)'*y(i)*((2*pi)^0.5)^(2*k);
ssdm1(:,idex)=ssdm1(:,idex)+wt1(idex,i)*x(i,:)'*x(i,:);
ssdm2(:,idex)=ssdm2(:,idex)+wt2(idex,i)*x(i,:)'*x(i,:) *((2*pi)^0.5)^(2*k);
end
b(:,1)=inv(ssdm1(:,idex)+ssdm2(:,idex)+lm*eye(size(z,2)))*(sspm1(:,idex)+sspm2(:,i
dex));
freg=b;

```



EK Tablo 1. Gerçek veri seti 1

X_1	X_2	X_3	X_4	X_5	Y
-0.4239	-0.2943	-0.1829	-0.2857	-0.3571	-12.3061
-0.3912	-0.2943	-0.2815	-0.2993	-0.3135	-9.8269
-0.1867	-0.1723	-0.38	-0.2452	-0.1176	-6.5725
-0.1704	-0.1501	-0.2158	-0.2722	-0.2482	-6.8947
-0.1581	-0.1057	-0.3307	-0.1235	-0.1829	-6.1596
-0.1009	-0.1057	-0.2486	-0.1505	-0.2047	-2.6332
-0.0559	-0.2943	0.0306	-0.1776	-0.1611	-5.7948
-0.0232	-0.2277	0.0471	-0.11	-0.1611	-5.9351
-0.0232	-0.017	0.0142	0.0117	0.0131	-0.5752
0.0177	-0.0059	0.0799	-0.0288	0.2961	5.0592
0.0586	0.1826	0.0964	0.1875	0.1872	5.9096
0.2549	0.2714	0.2442	0.2957	0.4267	10.4006
0.308	0.3157	0.2771	0.3498	0.2961	9.6367
0.4266	0.4044	0.3592	0.3633	0.1219	11.2918
0.4675	0.4932	0.4906	0.485	0.405	14.4004

EK Tablo 2. Gerçek veri seti 2

SIRA	YAŞ	BOY	KAR, ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
1	52	1,55	1,13	133	38	156	20	0,7	4,3	15	42,7	90	180
2	53	1,52	0,86	86	64	135	28	0,9	4,6	14	39,2	90	140
3	61	1,64	1,08	89	29	114	41	1,4	5	14	39,9	80	130
4	71	1,65	1,1	88	53	186	39	1	4,5	13	38,8	100	160
5	54	1,76	1,12	92	38	96	27	1,2	4,6	17	48,1	100	140
6	58	1,83	1	97	45	106	27	1,3	4,5	15	44,1	90	130
7	80	1,47	0,75	97	35	105	85	1,9	4,3	13	36,4	90	100
8	77	1,5	0,91	100	58	182	43	1,1	4,3	16	46,8	80	140
9	47	1,68	1,07	56	48	150	69	1,6	3,2	15	43,7	110	170
10	61	1,55	0,84	196	55	125	15	0,8	4,6	14	39,1	90	130
11	47	1,5	0,92	90	82	109	29	0,8	4,1	12	35	80	120
12	54	1,6	1,25	120	71	219	44	1,1	4	14	40,1	90	150
13	63	1,74	0,86	89	39	120	57	2,7	2,9	13	37,3	80	160
14	60	1,8	0,97	171	44	126	32	1	4,2	16	45,4	80	120
15	35	1,63	1,07	88	47	144	18	0,7	4,6	13	38,8	90	130
16	52	1,68	0,87	79	84	102	30	1,6	4,9	15	43,2	110	170
17	67	1,64	1,1	91	75	113	37	1,1	4,1	16	46,5	80	120
18	54	1,76	1,03	80	42	139	31	1,1	4,1	14	40,2	100	140
19	64	1,62	1,04	91	40	99	31	0,7	4,2	12	35,4	70	130
20	48	1,6	0,98	124	45	141	27	0,8	4,1	13	36,6	120	190
21	40	1,58	1,02	86	54	153	19	0,7	4,6	13	38,7	70	110

SIRA	YAŞ	BOY	KAR,ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
22	80	1,47	1,02	273	56	146	39	1	5	15	47,1	100	150
23	64	1,76	1,03	93	51	167	31	1	4,9	13	39	70	140
24	59	1,62	1,1	172	44	173	37	0,9	4,9	14	40,9	120	170
25	67	1,64	0,94	94	52	134	60	1,3	5	14	41,6	100	150
26	62	1,68	0,99	94	62	164	30	0,9	4,2	14	41,2	70	110
27	49	1,65	1,03	119	34	143	14	0,8	4,7	12	34,8	110	150
28	63	1,7	1,14	88	45	123	52	1,8	5,4	16	46,5	80	135
29	55	1,9	0,91	88	70	105	73	2,4	4,9	13	36,3	70	110
30	48	1,5	0,76	71	45	122	101	2,8	5,3	9,8	27,5	90	140
31	77	1,55	0,78	93	58	111	101	2,4	5,6	12	33,9	50	100
32	25	1,72	0,9	93	48	156	77	1,3	5,2	14	40,9	80	120
33	65	1,54	1,05	97	48	164	56	0,8	3,9	13	36,2	80	120
34	60	1,69	1,01	95	44	173	31	1,4	4,8	16	46,2	80	120
35	55	1,68	1,11	108	52	129	28	1,2	4,9	15	41,9	80	140
36	60	1,58	1,16	123	60	279	37	1,1	4,7	14	41	90	160
37	52	1,55	0,97	93	46	110	34	0,9	4,9	14	39,9	80	120
38	55	1,56	1,22	88	58	208	29	0,7	4	15	41,5	80	140
39	25	1,63	0,93	90	48	134	12	0,7	4,5	14	41,2	70	110
40	50	1,68	1,01	89	46	198	29	1,1	4,9	17	48,3	80	170
41	52	1,75	1,03	85	32	130	28	1,1	4,9	14	40,4	70	130
42	70	1,52	1,17	112	78	116	29	0,6	3,8	14	41,9	80	130
43	51	1,75	1,07	127	50	163	27	1,1	4,3	15	43,6	70	130
44	53	1,55	1,12	107	57	112	34	0,8	5,1	14	39,7	90	150

SIRA	YAŞ	BOY	KAR,ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
45	56	1,82	1,07	221	45	148	50	1,4	5,6	17	47,4	70	110
46	60	1,5	1,15	95	50	116	28	0,8	3,8	15	42,3	100	180
47	48	1,78	0,9	94	48	221	34	1,1	4,5	15	42	80	100
48	48	1,52	1,03	90	53	113	30	0,7	4,3	16	45,3	100	150
49	47	1,77	1,05	85	47	142	34	1,1	4,6	16	45,2	80	120
50	30	1,75	1	164	48	58	278	13,9	3,9	9	25,4	80	130
51	60	1,45	0,87	90	56	106	23	0,8	4,3	13	37,3	80	140
52	71	1,8	0,89	158	40	122	223	3,1	5,7	12	34,9	80	140
53	62	1,66	1,05	15	48	146	41	1,3	4,4	14	40	80	120
54	48	1,68	0,97	77	43	148	34	1	4,7	17	47,4	95	145
55	71	1,6	1,05	127	42	211	48	1	4,5	12	35,5	100	160
56	60	1,62	1,18	222	44	146	30	0,8	4	13	37,3	80	150
57	73	1,8	1,06	86	37	155	60	1,1	5,1	13	36,5	90	140
58	75	1,87	1,1	82	42	96	28	0,8	4,6	15	42,5	70	130
59	65	1,6	0,95	87	46	163	60	1,3	4,7	17	49,5	80	120
60	49	1,6	0,97	80	40	147	41	1,1	4,8	15	43,3	90	160
61	52	1,53	0,98	82	49	189	32	0,9	4,2	13	39	100	150
62	52	1,6	1,04	268	50	106	28	0,9	4,3	12	34,1	100	150
63	69	1,65	1,12	118	53	164	45	1,4	4,9	17	49,1	110	190
64	63	1,6	0,8	107	47	157	31	1,2	4,5	15	42,6	100	150
65	66	1,65	0,86	84	40	184	34	1,2	4,7	13	37,3	90	150
66	57	1,57	1,07	113	44	141	58	2	4,3	14	40,7	100	150
67	46	1,5	1,18	192	74	109	19	0,7	4,8	12	36,7	100	150

SIRA	YAŞ	BOY	KAR,ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
68	57	1,62	0,95	95	49	140	27	0,8	4,5	13,5	39,2	130	180
69	71	1,65	1	91	53	159	47	1,3	4,1	15,1	42,6	80	140
70	73	1,72	1,03	78	51	175	41	1,4	4,3	16,6	48,7	60	120
71	57	1,67	1,09	91	55	123	46	1,2	4,5	15,2	43,6	70	130
72	50	1,8	1,02	103	63	119	33	1,6	4,8	16,2	47,1	80	140
73	65	1,48	1,2	92	66	203	37	1,2	4,9	13,7	40	80	140
74	59	1,45	0,96	110	49	140	44	1,3	4,5	14,1	40,9	90	160
75	50	1,54	1,09	106	36	150	177	8,9	4,1	11,7	32,5	90	180
76	49	1,43	0,96	94	53	153	36	0,9	4,5	14,8	41,6	80	150
77	69	1,64	1,04	95	50	207	41	1,3	4,3	14,5	41,6	80	120
78	55	1,7	1,13	219	50	186	26	1,8	4,4	15,5	45,1	90	160
79	62	1,73	1,07	106	46	173	35	1,3	4,3	16,6	47,9	110	170
80	45	1,6	1,29	107	53	127	28	0,8	4,3	14,8	43,1	60	130
81	42	1,57	1,16	110	49	140	32	1,1	4,5	12,5	37	80	140
82	50	1,56	1,12	91	45	177	26	0,9	4,6	15,6	46	70	110
83	50	1,7	1,01	110	49	140	44	1,3	4,5	14,1	40,9	100	170
84	63	1,6	1,04	98	45	126	16	0,9	4,4	14,3	41,9	110	190
85	59	1,6	1,2	117	46	163	32	0,9	4,6	14,4	41,8	100	150
86	60	1,58	1,05	96	47	102	38	1	4,3	14,7	41,8	70	110
87	64	1,8	1,1	147	42	154	50	1,2	5,3	12,0	34,5	70	130
88	62	1,55	1,07	105	47	133	28	1	4,3	13,1	36,5	80	140

SIRA	YAŞ	BOY	KAR,ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
89	72	1,73	1,07	196	41	66	28	1	4,3	16,3	46,2	80	120
90	68	1,55	1,04	97	54	84	49	1,1	5,1	11,9	37,7	90	140
91	51	1,75	1,07	90	43	75	26	1,1	4,9	15,5	44,5	70	120
92	63	1,68	1,11	107	50	77	41	1,3	4,7	14	41	90	180
93	78	1,55	0,75	94	39	119	27	0,8	4,4	12,4	36,1	80	130
94	72	1,85	1,05	128	41	73	45	1,4	5,4	14,9	42,9	80	140
95	58	1,56	0,93	98	58	101	41	1	4	13,6	43,4	80	140
96	63	1,63	1,05	102	59	142	43	0,8	4,3	13,6	39,1	90	150
97	48	1,6	1,1	235	55	134	69	1,7	5,4	11,7	33,8	80	140
98	50	1,8	1,04	93	47	111	25	1,1	4,8	15,8	45,5	80	130
99	53	1,75	0,97	256	50	167	31	0,9	4,4	15,3	44,2	70	140
100	57	1,65	0,95	77	37	90	36	1,2	4,2	16,4	47,1	70	130
101	55	1,72	0,89	97	54	83	35	1,0	4,3	14,3	41,3	90	140
102	70	1,55	0,85	89	73	254	39	0,9	4,5	12,9	37,2	80	120
103	68	1,51	1,06	140	54	132	38	0,8	4,6	10,9	33,4	80	140
104	55	1,82	0,9	311	34	68	59	1,4	4,7	12,8	37,5	90	140
105	51	1,65	1,03	129	39	122	30	1	4,4	15,5	44,6	120	170
106	65	1,65	1,09	105	45	103	53	1,2	4,4	14,6	41,5	80	120
107	71	1,6	1,03	76	43	124	29	0,9	4,4	14,6	41,9	100	150
108	63	1,65	1,09	94	43	108	31	1	4,8	17,5	50,1	110	170
109	20	1,58	0,75	79	80	122	40	0,9	4,6	13,4	38,6	90	120
110	44	1,78	1,15	100	41	117	23	0,7	4,2	14,6	41,8	120	170
111	61	1,61	0,96	173	34	115	32	0,9	3,5	14,7	41,5	100	140

SIRA	YAŞ	BOY	KAR,ÇEV	GLUKOZ	HDL	LDL	ÜRE	KREATİNİN	POTASYUM	HGB	HCT	DİASTOLİK	SİSTOLİK
112	58	1,62	1,26	95	58	152	36	0,7	4,4	14,3	40,5	80	120
113	65	1,58	1,01	93	79	185	28	0,9	4,6	13	39,9	80	140
114	61	1,75	1	85	42	175	37	1,6	4,3	14,8	43	80	130
115	79	1,7	1,06	111	40	126	64	1,1	4,2	15,9	46	110	160
116	63	1,6	1,12	77	54	105	28	0,8	4,3	13	37,2	100	150
117	63	1,64	0,93	93	52	160	34	1,2	3,9	15,6	44,5	80	120
118	58	1,76	1,05	89	44	76	32	1	3,7	15,9	45,0	70	130
119	59	1,81	0,93	89	45	122	37	1	4,4	15,3	44,6	90	160
120	75	1,6	0,92	98	51	65	49	1	4,4	12,3	35,9	80	130
121	42	1,54	1,3	130	45	130	24	0,8	4,1	12,3	37,1	80	120
122	67	1,67	1,05	112	43	115	38	1,1	4,3	13,9	40,7	70	130
123	70	1,65	1,04	112	46	130	68	1,7	4,3	15,6	45,3	70	130
124	55	1,48	1,02	105	76	162	28	0,9	4,4	16,7	48,8	90	130
125	52	1,57	1,08	95	58	110	25	1	4,6	13,2	38,7	80	140
126	55	1,73	0,99	70	54	113	31	1	4,4	15,2	44	90	150
127	55	1,65	1,12	90	81	145	33	0,9	4,4	14,7	41,9	110	160
128	54	1,58	1,07	112	43	175	25	0,7	4,2	16,4	44,4	90	130

ÖZGEÇMİŞ

1988 yılında Elazığ'da doğdum. İlk, orta ve lise öğrenimimi Elazığ'da tamamladıktan sonra 2009 yılında Fırat Üniversitesi İstatistik bölümünü kazandım. 2013 yılında bölüm mezun oldum. 2017-2018 güz döneminde Fırat Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstatistiksel Bilgi Sistemler Bilim Dalı'nda yüksek lisansa başladım. Evli ve bir çocuk annesiyim.

