
GAZİ ÜNİVERSİTESİ
Gazi
1926

ÇOKLU BAĞLANTI DURUMUNDA REGRESYON KATSAYILARININ ANLAMLILIĞI İÇİN İKİ PARAMETRELİ YANLI TAHMİN EDİCİLERE DAYALI HİPOTEZ TESTLERİ

Hilal KAPLAN TABAK

YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

OCAK 2024

ETİK BEYAN

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

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

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

Hilal KAPLAN TABAK

12/01/2024

ÇOKLU BAĞLANTI DURUMUNDA REGRESYON KATSAYILARININ ANLAMLILIĞI İÇİN İKİ PARAMETRELİ YANLI TAHMİN EDİCİLERE DAYALI HİPOTEZ TESTLERİ

(Yüksek Lisans Tezi)

Hilal KAPLAN TABAK

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Ocak 2024

ÖZET

Doğrusal regresyon modelinde, bağımsız değişkenler arasında ilişki olmaması varsayımı sağlanmadığında çoklu bağlantı sorunuyla karşılaşılmaktadır. Bu durumda, En Küçük Kareler tahmin edicisi ile elde edilen parametre tahminleri gerçek değerinden uzaklaşarak yanlış bulguların elde edilmesine sebep olabilmektedir. Çoklu bağlantının etkisini gidermek için etkili olan yöntemlerden biri yanlı tahmin yöntemleridir. Bu tahmin yöntemleri ile artan varyansın azaltılması ve tutarlı parametre tahminlerinin elde edilmesi amaçlanmaktadır. Bu çalışmada, Halawa ve El-Bassiouni'nin (2000) çalışmasında Ridge tahmin edicisi için elde edilen test istatistiğinden yararlanarak regresyon katsayılarının anlamlılığının testi için Liu Tipi tahmin edicisine ilişkin bir test istatistiği elde edilmiştir. Gerçek veri uygulaması ve simülasyon çalışması ile Ridge, Liu ve Liu Tipi tahmin edicilerine ilişkin regresyon modelinin katsayılarının anlamlılık testleri yapılmıştır. Farklı durumlar altında bu tahmin edicilerin 1.tip hataları ve güç değerleri hesaplanarak elde edilen bulgular karşılaştırılmıştır. Elde edilen sonuçlara göre, yüksek derecede çoklu bağlantı bulunduğunda, Liu Tipi tahmin edicisine ait elde edilen LT1 testinin diğer testlere göre daha güçlü olduğu gözlenmiştir.

Bilim Kodu	:	20513
Anahtar Kelimeler	:	Çoklu bağlantı, yanlı tahmin ediciler, hipotez testi, 1.tip hata, testin gücü
Sayfa Adedi	:	45
Danışman	:	Prof. Dr. Meral EBEGİL

HYPOTHESIS TESTS BASED ON TWO-PARAMETER BIASED ESTIMATORS FOR SIGNIFICANCE OF REGRESSION COEFFICIENTS IN CASE OF MULTICOLLINEARITY

(M. Sc. Thesis)

Hilal KAPLAN TABAK

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

January 2024

ABSTRACT

In the linear regression model, multicollinearity problems arise when the assumption of no relationship between independent variables is not met. In this case, the estimations of the model parameters obtained by the Least Squares Estimator method deviate from the true value, leading to erroneous results. The Biased Estimator method is one of the effective methods used to eliminate the negative effects caused by multicollinearity. With this estimation method, it is aimed to reduce the increased variance and obtain consistent parameter estimates. The aim of this study is to obtain a test statistic for the Liu-Type Estimator to test the significance of the regression coefficients, using the test statistic obtained for the Ridge estimator in the study of Halawa and El-Bassiouni (2000). The significance tests of the regression model coefficients for the Ridge, Liu and Liu type biased estimators were conducted with a real data application and simulation study. The type-1 errors and power values of these estimators in different situations are calculated and the results obtained are compared. According to the results, when a high degree of multicollinearity is found, the LT1 test obtained for the Liu Type estimator is stronger than the other tests.

Science Code : 20513

Key Words : Multicollinearity, biased estimators, hypothesis testing, type 1 error, powers of test

Page Number : 45

Supervisor : Prof. Dr. Meral EBEGİL

TEŞEKKÜR

Tez çalışmam boyunca kıymetli tecrübelerinden faydalandığım, bilgi ve katkılarıyla bana yol gösteren, ilgisini ve güler yüzünü esirgemeyen değerli danışmanım Prof. Dr. Meral EBEGİL hocama, tez çalışmalarımın yürütülmesi sırasında yardımlarını esirgemeyen Doç. Dr. Esra YİĞİT GÖKPINAR hocama ve yayınlarından faydalandığım bütün hocalarıma emeklerinden ve katkılarından dolayı teşekkür ederim. Gelir İdaresi Başkanlığında göreve başladığım günden itibaren her konuda kolaylık sağlayan müdürüm Sayın Göksel BAYRAM' a desteğinden dolayı teşekkür ederim. Beni her zaman destekleyen ve yanımda olan canım anneme, babama ve kardeşime, birlikte yürüdüğümüz bu yolda mutluluk ve moral kaynağım olan sevgili eşim Ziya'ya ve bu tezi yazmaya başladığım süreçte henüz karnımda olan bana şans getiren minik kızıma teşekkür ederim.

İÇİNDEKİLER

Sayfa

ÖZET iv

ABSTRACT v

TEŞEKKÜR v

İÇİNDEKİLER vii

ÇİZELGELERİN LİSTESİ ix

ŞEKİLLERİN LİSTESİ xiv

SİMGELER VE KISALTMALAR xv

1. GİRİŞ 1

2. GENEL KAVRAMLAR 5

2.1. Regresyon Modeli ve Varsayımlar 5

2.2. Çoklu Bağlantı Sorunu 7

2.3. Çoklu Bağlantının Belirlenmesi 8

2.3.1. Korelasyon katsayısı 8

2.3.2. Varyans şişirme faktörü (Variance inflation factors -VIF) 8

2.3.3. Korelasyon matrisinin özdeğerler ve özvektörleri 9

2.3.4. Belirtme katsayısı (R^2) 9

2.3.5. $X'X$ matrisinin determinantı 10

2.4. Çoklu Bağlantının Ortadan Kaldırılması 10

3. YANLI TAHMİN EDİCİLER 11

3.1. Ridge Tahmin Edicisi 11

3.2. Liu Tahmin Edicisi 15

3.3. Liu Tipi Tahmin Edicisi 18

4. UYGULAMA 21

4.1. EKK Tahmin Edicisi ile Elde Edilen Sonuçlar 23

Sayfa

4.2. Ridge Tahmin Edicisi ile Elde Edilen Sonuçlar 24
4.3. Liu Tahmin Edicisi ile Elde Edilen Sonuçlar 26
4.4. Liu Tipi Tahmin Edicisi ile Elde Edilen Sonuçlar 27
5. SİMÜLASYON ÇALIŞMASI 31
6. SONUÇ 41
KAYNAKLAR 43
ÖZGEÇMİŞ 45

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

Çizelge 4.1. Portland çimento verisi 21

Çizelge 4.2. Standartlaştırılmış veriler 22

Çizelge 4.3. X′X korelasyon matrisi 22

Çizelge 4.4. X′X korelasyon matrisine ilişkin öz değer matrisi 23

Çizelge 4.5. Regresyon katsayılarının EKK tahmin edicisine dayalı test sonuçları 24

Çizelge 4.6. Regresyon katsayılarının Ridge tahmin edicisine dayalı test sonuçları 26

Çizelge 4.7. Regresyon katsayılarının Liu tahmin edicisine dayalı test sonuçları 27

Çizelge 4.8. Regresyon katsayılarının Liu Tipi tahmin edicisine dayalı test sonuçları . 28

Çizelge 5.1. ρ=0.85 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları 32

Çizelge 5.2. ρ=0.95 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları 32

Çizelge 5.3. ρ=0.99 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları 32

Çizelge 5.4. ρ=0.85 ve σ=1 iken testlerin deneysel 1.tip hata oranları 33

Çizelge 5.5. ρ=0.95 ve σ=1 iken testlerin deneysel 1.tip hata oranları 33

Çizelge 5.6. ρ=0.99 ve σ=1 iken testlerin deneysel 1.tip hata oranları 33

Çizelge 5.7. ρ=0.85 ve σ=2 iken testlerin deneysel 1.tip hata oranları 34

Çizelge 5.8. ρ=0.95 ve σ=2 iken testlerin deneysel 1.tip hata oranları 34

Çizelge 5.9. ρ=0.99 ve σ=2 iken testlerin deneysel 1.tip hata oranları 35

Çizelge 5.10. ρ=0.85 ve σ=0,5 iken testlerin güç değerleri 36

Çizelge 5.11. ρ=0.95 ve σ=0,5 iken testlerin güç değerleri 36

Çizelge 5.12. ρ=0.99 ve σ=0,5 iken testlerin güç değerleri 36

Çizelge 5.13. ρ=0.85 ve σ=1 iken testlerin güç değerleri 37

Çizelge 5.14. ρ=0.95 ve σ=1 iken testlerin güç değerleri 37

Çizelge 5.15. ρ=0.99 ve σ=1 iken testlerin güç değerleri 37

Çizelge 5.16. ρ=0.85 ve σ=2 iken testlerin güç değerleri 38

Çizelge **Sayfa**

Çizelge 5.17. ρ=0.95 ve σ=2 iken testlerin güç değerleri 38

Çizelge 5.18. ρ=0.99 ve σ=2 iken testlerin güç değerleri 39

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

Şekil 3.1. HKO fonksiyonları 13

Şekil 3.2. Ridge ve Liu tahmin edicilerinin yanlılık parametreleri..... 16

SİMGELER VE KISALTMALAR

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

Simgeler	**Açıklamalar**
$\widehat{\boldsymbol{\beta}}_{\mathbf{EKK}}$	En küçük kareler tahmin edicisi
$\widehat{\boldsymbol{\beta}}_{\mathbf{L}}$	Liu tahmin edicisi
$\widehat{\boldsymbol{\beta}}_{\mathbf{LT}}$	Liu Tipi tahmin edicisi
$\widehat{\boldsymbol{\beta}}_{\mathbf{R}}$	Ridge tahmin edicisi
$\mathbf{R^2}$	Belirtme katsayısı
$\boldsymbol{\lambda_i}$	X′X korelasyon matrisinin özdeğerleri
$\widehat{\boldsymbol{\alpha}}$	Kanonik formda parametre tahmin vektörü
d	Liu tahmin edicisinin yan parametresi
k	Ridge tahmin edicisinin yan parametresi
n	Gözlem sayısı
q	Bağımsız değişken sayısı
X	Bağımsız değişken gözlem matrisi
Y	Bağımlı değişken vektörü
β	Bilinmeyen parametreler vektörü
ε	Hata vektörü
X′X	Korelasyon matrisi

Kısaltmalar	**Açıklamalar**
EKK	En küçük kareler
HKO	Hata kareler ortalaması
MF	Most favorable
VIF	Variance inflation factors

1. GİRİŞ

Doğrusal regresyon analizi, bağımlı değişken ve bağımsız değişkenler arasındaki ilişkiyi belirlemek ve bilinmeyen regresyon parametrelerinin değerini tahmin etmek için kullanılan istatistiksel bir yöntemdir. Birçok farklı bilimsel uygulamada regresyon analizinden yararlanılmaktadır. En küçük kareler (EKK) yöntemi, regresyon modelinin parametrelerinin tahmin edilmesinde en sık kullanılan yöntemlerden biridir. EKK tahmin yöntemi ile doğru ve güvenilir tahminlerin elde edilebilmesi için bazı varsayımların sağlanması gerekmektedir. Bu varsayımlar, bağımsız değişkenler arasında doğrusal bir ilişki olmaması, hata terimlerinin normal dağılımlı, ortalaması sıfır ve sabit varyanslı olması, hata terimlerinin birbirinden bağımsız olması şeklinde özetlenebilir. Parametre tahminlerinin yansız ve güvenilir olması bu varsayımların sağlanmasına bağlıdır. Ancak bu varsayımlarda bazı sapmalar söz konusu olabilir.

Bağımsız değişkenler arasında ilişki olmaması varsayımı önemli ve temel bir özelliktir. Bu varsayımın sağlanmaması durumunda çoklu bağlantı (multicollinearity) sorunu ortaya çıkmaktadır. Bu durum, regresyon modeli üzerinde EKK tahmin yönteminin kullanılması ile elde edilen tahminlerle yapılan çıkarsamalardan hatalı sonuçların elde edilmesine sebep olabilmektedir. Diğer bir ifadeyle, çoklu bağlantı EKK tahmin edicisinin varyans ve standart hatasını artırmakta ve bununla birlikte güven aralıklarını genişletmektedir. Bunun sonucu olarak, test istatistiği gerçek değerinden farklı bulunmakta ve dolayısıyla hipotez yanlış yorumlanabilmektedir. Çoklu bağlantının parametre tahminleri üzerindeki bu olumsuz etkilerini azaltmak veya gidermek amacıyla birçok yöntem kullanılmaktadır. Bunlardan biri olan yanlı tahmin edicilerin kullanılması, parametre tahminlerinin daha kararlı hale gelmesini sağlar. Bu tahmin yöntemlerinde yansız bir tahmin ediciye bir miktar yanlılık eklenerek çoklu bağlantı sebebiyle artan varyansın azaltılması amaçlanır. Bu durumda, daha küçük varyanslı ve daha küçük Hata Kareler Ortalaması (HKO) değerine sahip tahmin ediciler elde edilmektedir.

Literatürde yer alan çalışmalar incelendiğinde çoklu bağlantı sorunuyla karşılaşılan doğrusal regresyon modellerinde yanlı tahmin edicilerin yaygın olarak kullanıldığı görülmektedir. Çalışmalarda birçok farklı alanlara ait veriler kullanılmıştır. Hoerl ve Kennard (1970) tarafından yapılan çalışmada, Ridge tahmin edicisi HKO değerine göre

incelenmiştir. Burada, yanlılık parametresi kullanılarak parametre tahminlerinin HKO değerinin azaltılması amaçlanmıştır. Liski (1982) çalışmasında, yanlı tahmin ediciler ile EKK tahmin edicisini HKO değerine göre karşılaştırmıştır. Liu (1993) tarafından geliştirilen Liu tahmin edicisi Ridge tahmin edicisine alternatif olarak gösterilmiştir. Akdeniz ve Kaçıranlar (1995) tarafından bu tahmin edici Liu tahmin edicisi olarak tanımlanmıştır. Halawa ve El-Bassiouni (2000), Ridge tahmin edicisini kullanarak regresyon modelinin katsayılarının anlamlılığının testi için t test istatistiğini oluşturmuşlardır. EKK ve Ridge tahmin edicilerinden elde edilen test istatistiğinin bulguları, simülasyon çalışması yapılarak karşılaştırılmıştır. Kibria (2003) yaptığı çalışmada, EKK tahmin edicisi ile yanlı tahmin edicileri simülasyon tekniği kullanarak karşılaştırmıştır. Bunun sonucunda HKO ölçütüne göre; önerilen yanlı tahmin edicilerin HKO değerlerinin, EKK tahmin edicisinin HKO değerinden daha küçük olduğunu göstermiştir. Çoklu bağlantının çok yüksek çıktığı durumlarda, tek parametreli yanlı tahmin edicilere kıyasla daha iyi sonuçlar verdiği gözlemlenen iki parametreli yanlı tahmin yöntemleri ile ilgili çalışmalar yapılmıştır. Liu (2003) çalışmasında, k ve d parametrelerini kullanarak iki parametreli yanlı tahmin edici olarak Liu Tipi tahmin edicisini önermiştir.

Ebegil, Gökpınar ve Ekni (2006), simülasyon çalışması yaparak Liski'nin (1982, 1983) çalışmalarında öne sürdüğü test istatistiğini kullanarak Ridge ve Liu tahmin edicilerinden elde ettikleri bulguları karşılaştırmışlardır. Özkale ve Kaçıranlar (2007) çalışmasında, iki parametreye sahip yanlı tahmin edicileri inceleyerek yeni bir tahmin edici önermişlerdir. Sakallıoğlu ve Kaçıranlar (2008) tarafından iki parametreye sahip yanlı tahmin edici olarak k-d sınıfı tahmin edicisi önerilmiştir. Ebegil ve Gökpınar (2012) yaptıkları çalışmada, EKK ve Liu Tipi tahmin edicilerinin performanslarını gereklilik ve yeterlilik koşulu altında elde edilen test istatistiğini kullanarak incelemişlerdir. Gruber (2012) çalışmasında, EKK, Ridge ve Liu tahmin edicilerini kıyaslamak amacıyla HKO kriterini kullanarak Liu tahmin edicisinin HKO değerinin, EKK tahmin edicisinin HKO değerinden küçük çıktığı ve bazı istisnai durumların haricinde, Liu tahmin edicisinin HKO değerinin Ridge tahmin edicisinin HKO değerinden de küçük çıktığı sonucu elde edilmiştir. Gökpınar ve Ebegil (2016) çalışmasında, Halawa ve El-Bassiouni'nin (2000) çalışmasını Ridge tahmin edicisi için geliştirerek Ridge tahmin edicisi için en iyi sonucu veren k tahmin edicilerini incelemişlerdir. Wilcox (2019) tarafından yapılan çalışmada, değişen varyans söz konusu iken Ridge tahmin edicisine dayalı modelin parametrelerinin anlamlılığının testi için bir test istatistiği öne sürülmüştür. Melo ve Kibria (2021), Halawa ve El-Bassiouni'nin (2000)

çalışmasında elde edilen test istatistiğini kullanarak Liu tahmin edicisine dayalı bir test istatistiği önermişlerdir. Elde edilen test istatistiği kullanılarak yapılan simülasyon çalışmasında, farklı durumlar altında deneysel 1.tip hataları ile testin güç değerleri karşılaştırılmıştır. Bu çalışmada, Halawa ve El-Bassiouni'nin (2000) çalışması dikkate alınarak Liu Tipi tahmin edicisine dayalı bir test istatistiği önerilmiştir.

Bu çalışmanın ikinci bölümünde, doğrusal regresyon modeli ve sıklıkla karşılaşılan çoklu bağlantı sorunu ana hatları ile ele alınmış ve bu kavramlarla ilgili genel bilgilere yer verilmiştir.

Üçüncü bölümde ise çoklu bağlantı sorununun çözümüne ilişkin önerilen yanlı tahmin ediciler ele alınmıştır. Öncelikle, tek yanlılık parametresine sahip Ridge ve Liu tahmin edicileri ile ilgili teorik bilgilere yer verilmiş sonrasında çok yüksek çoklu bağlantı durumunda tek parametreli yanlı tahmin edicilerin yetersiz kalması sebebiyle önerilen iki yanlılık parametresine sahip Liu Tipi tahmin edicisi açıklanmıştır. Literatürde genellikle yanlı tahmin edicilerin parametrelerinin tahmin edilmesi üzerine çalışmalar bulunmaktadır. Ancak, bu tahmin edicilerin kullanılmasıyla elde edilen regresyon katsayılarının anlamlılığının test edilmesiyle ilgili çalışmalar sınırlı kalmıştır. Bu çalışmada, Halawa ve El-Bassiouni'nin (2000) çalışmasında geliştirdiği test istatistiği kullanılarak regresyon modelinin katsayılarının anlamlılığının test edilmesi amaçlanmış olup Liu Tipi tahmin edicisi için bir test istatistiği elde edilmiştir.

Dördüncü bölümde, yanlı tahmin edicilere ait regresyon modelinin katsayılarının anlamlılık testi, gerçek verilerin kullanıldığı bir uygulama çalışması ile yapılmıştır. Gerçek veri uygulaması olarak Portland Çimento verileri kullanılmıştır (Hald, 1952).

Çalışmanın beşinci bölümünde, tahmin edicilere ilişkin elde edilen test istatistiği ile Ridge, Liu ve Liu Tipi tahmin edicilerin farklı durumlar altında simülasyon çalışması yoluyla parametrelerinin anlamlılık testleri yapılmış ve tahmin edicilerin deneysel 1.tip hata oranları ile testin gücü performansları karşılaştırılmıştır.

2. GENEL KAVRAMLAR

Çalışmanın bu bölümünde doğrusal regresyon modeli ve çoklu bağlantı sorunu detaylı olarak incelenecektir.

2.1. Regresyon Modeli ve Varsayımlar

Regresyon analizinde amaç, bağımsız değişkenlerden herhangi birindeki değişim sonucunda bağımlı değişkende meydana gelen etkiyi bulmak ve bağımlı değişkeni en iyi tahmin eden modeli saptamaktır. Çoklu doğrusal regresyon modelinin genel formu Eş.2.1'de gösterilmiştir.

$$Y = X\beta + \varepsilon \tag{2.1}$$

Burada, Y; (nx1) boyutlu ortalama etrafında merkezileştirilmiş bağımlı değişken vektörü, q=p+1 olmak üzere, X; (nxq) boyutlu bağımsız değişkenlerin gözlem matrisi, β; (qx1) boyutlu bilinmeyen parametreler vektörü, ε; (nx1) boyutlu sıfır ortalamalı $\sigma^2 I$ varyanslı rastgele hata vektörüdür. Eş. 2.1'deki modelin kanonik formu Eş. 2.2'de gösterilmiştir.

$$Y = Z\alpha + \varepsilon \tag{2.2}$$

Burada, $Z = XP$ ve $\alpha = P'\beta$ olmak üzere $X'X$ korelasyon formunda elde etmek için ortalamadan farkı alınmış ve ölçeklendirilmiş qxq boyutlu pozitif tanımlı korelasyon matrisidir ve P matrisi, kolonları $X'X$ matrisinin özvektörlerinden oluşan ortonormal matris olmak üzere $X'X$ matrisini $P'X'XP = \Lambda$ biçiminde köşegenleştirir. Burada Λ , elemanları $X'X$ matrisinin pozitif özdeğerleri olan $(\lambda_1, \ldots., \lambda_q)$ (qxq) boyutlu köşegen bir matristir (Liski,1982,1983). Eş. 2.1'de verilen regresyon modeli için EKK tahmin edicisi Eş. 2.3'deki gibidir.

$$\hat{\beta}_{EKK} = (X'X)^{-1}(X'Y) \tag{2.3}$$

Burada, X′Y; (q-1)×1 boyutlu Y bağımlı değişkeni ile her bir bağımsız değişken arasındaki korelasyon vektörü olmak üzere merkezileştirilmeleri ve ölçeklendirilmeleri önerilmiştir. Eş. 2.3'de verilen EKK tahmin edicisinin kanonik formu Eş. 2.4'de verilmiştir.

$$\hat{\alpha}_{EKK} = \Lambda^{-1} Z'Y \tag{2.4}$$

EKK tahmin edicisi hata kareler toplamını minimum yapmayı hedefleyen bir yöntemdir. EKK tahmin edicisi yansız bir tahmin edici olmak üzere tüm yansız tahmin ediciler arasında da en küçük varyansa sahiptir. EKK yönteminin bu özellikleri sağlayabilmesi ve güvenilir tahminler elde edilebilmesi bazı varsayımların sağlanması koşuluna bağlıdır. Bu varsayımlar;

1) Y bağımlı değişkeni, X bağımsız değişkeninin doğrusal fonksiyonudur.
2) Hata terimlerinin beklenen değeri sıfır, varyansları bütün gözlemler için σ_i^2(i=1,2,….,n) olup normal dağılıma sahiptir.
3) Hata terimleri ile bağımlı değişken arasında korelasyon yoktur.
4) Bağımsız değişkenler arasında doğrusal ilişki yoktur.
5) Bağımsız değişkenler ile hata terimleri birbirinden bağımsızdır.

Bu varsayımların sağlanması durumunda EKK yöntemi model parametrelerinin tahmin edilmesinde güvenilir ve etkili sonuçlar verir. EKK tahmin edicisine ait beklenen değer ve varyans Eş. 2.5 ve Eş. 2.6'da verilmiştir.

$$E\left(\hat{\beta}_{EKK}\right) = \beta \tag{2.5}$$

$$Var\left(\hat{\beta}_{EKK}\right) = \sigma^2 (X'X)^{-1} \tag{2.6}$$

Regresyon katsayılarının anlamlılığını test etmek için kurulan hipotez Eş. 2.7' deki gibidir.

$$H_0: \beta_i = 0$$
$$H_1: \beta_i \neq 0 \ , \quad (i=1,2,….,q) \tag{2.7}$$

EKK tahmin edicisine dayalı test istatistiği Eş. 2.8' de verilmiştir.

$$t_{i(EKK)} = \frac{\hat{\beta}_{i(EKK)}}{S(\hat{\beta}_{i(EKK)})} \tag{2.8}$$

Eş. 2.8' de verilen test istatistiği H_0 hipotezinin doğruluğu altında (n-q-1) serbestlik dereceli t dağılımına sahiptir. $\hat{\beta}_{i(EKK)}$, $\hat{\beta}_{EKK}$'nın i-inci elemanını ve $S(\hat{\beta}_{i(EKK)})$, $\hat{\beta}_{EKK}$'nın varyans tahmininin i-inci köşegen elamanının karekökünü göstermektedir. Buna ilişkin denklem Eş. 2.9' da verilmiştir.

$$Var(\hat{\beta}_{EKK}) = \sigma^2(X'X)^{-1} \tag{2.9}$$

Burada σ^2 bilinmediği zaman, Eş. 2.10' de verilen EKK tahmin edicisine dayalı σ^2'nin tahmini kullanılır.

$$\hat{\sigma}^2_{EKK} = \frac{(Y - X\hat{\beta}_{EKK})'(Y - X\hat{\beta}_{EKK})}{n - q - 1} \tag{2.10}$$

2.2. Çoklu Bağlantı Sorunu

Doğrusal regresyon modelinde, bağımsız değişkenler arasında tam veya tama yakın doğrusal bir ilişki bulunması çoklu bağlantı sorunu olarak tanımlanmaktadır. Bağımsız değişkenler arasında güçlü korelasyon bulunması olarak da ifade edilebilir. Bu durumda, yansız tahmin ediciler arasında en küçük varyansa sahip olan EKK tahmin edicisi üzerinde ciddi derecede olumsuz etkiler görülmektedir. EKK tahmin edicisinin varyans ve standart hatasındaki artış, model üzerinde yapılacak tahmin ve yorumlar sonucunda yanlış bulgular elde edilmesine sebep olmaktadır. Ayrıca, regresyon katsayılarında yüksek varyans ve kovaryanslar görülmekte ve güven aralıkları genişlemektedir (Montgomery, Peck ve Vining, 2001). Dolayısıyla, hesaplanan test istatistikleri de küçülmekte ve regresyon modelinde değişken seçimleri yanıltıcı hale gelmektedir. Tam çoklu bağlantı durumunda, kötü koşulluluk problemiyle karşılaşılır. Bu durum, korelasyon matrisinin tersinin hesaplanamaması, elde edilen parametre tahminlerinin belirlenememesi ve standart hatanın sonsuz seviyeye ulaşması şeklinde de ifade edilebilir. Bu sebeplerden dolayı regresyon modelinde öncelikle çoklu bağlantının varlığı tespit edilmelidir. Eğer çoklu bağlantı saptanırsa bu durumu ortadan kaldırmaya yönelik yöntemler kullanılmalıdır.

Çoklu bağlantı durumunda, EKK parametre tahminleri mutlak değerce büyür ve gerçek değerinden uzaklaşır (Montgomery, Peck ve Vining, 2001). Yani, EKK tahmin edicisinin gerçek değeri ile tahmin değeri arasındaki uzaklık artar. $\hat{\beta}_{EKK}$ ile β arasındaki uzaklığın karesinin beklenen değeri alınarak elde edilen bu uzaklık, Hoerl ve Kennard (1970) tarafından $L^2 = (\hat{\beta}_{EKK} - \beta)'(\hat{\beta}_{EKK} - \beta)$ olmak üzere $E(L^2)$ Eş. 2.11'deki gibi ifade edilmiştir.

$$E(L^2) = E\left(\left(\hat{\beta}_{EKK} - \beta\right)'\left(\hat{\beta}_{EKK} - \beta\right)\right)$$
$$= \sigma^2 tr(X'X)^{-1} = \sigma^2 \sum_{i=1}^{q} {}^{1}/_{\lambda_i} \qquad (2.11)$$

Burada λ_i, $X'X$ matrisinin özdeğerleridir.

2.3. Çoklu Bağlantının Belirlenmesi

Korelasyon katsayısı, varyans şişirme faktörü (VIF), korelasyon matrisinin öz değerlerine göre koşul sayısı ölçütü, belirtme katsayısı ve $X'X$ matrisinin determinantı gibi birçok farklı yöntem çoklu bağlantı sorununun saptanması için kullanılabilir. Bu kısımda bu yöntemler kısaca tanıtılacaktır.

2.3.1. Korelasyon katsayısı

Korelasyon katsayısı yaklaşımı ile iki bağımsız değişken arasındaki ilişki açıklanmaktadır. $X'X$ korelasyon matrisinin köşegen elemanları dışındaki elemanları korelasyon katsayısıdır (Gujarati, 1999). X_i ve X_j vektörleri arasında güçlü bir doğrusal ilişki varsa korelasyon katsayısının mutlak değeri $|r_{ij}|$ 1'e yaklaşır. Bu yaklaşım ikiden fazla bağımsız değişken arasındaki çoklu bağlantı sorununu belirlemede yeterli olmayabilir.

2.3.2. Varyans şişirme faktörü (Variance inflation factors -VIF)

İki ya da daha fazla bağımsız değişkenler arasındaki çoklu bağlantının varlığını belirlemede etkili olan bu yöntem, Farrar ve Glauber (1967) tarafından önerilmiştir. $C = (X'X)^{-1}$ olarak ifade edilen korelasyon matrisinin tersinin j. köşegen elemanı Marquardt

(1970) tarafından varyans şişirme faktörü (VIF) olarak tanımlanmıştır. Burada, bağımsız değişken sayısı kadar VIF değeri hesaplanır. VIF Eş. 2.12'de verilmiştir.

$$VIF_j = C_{jj} = (1 - R_j^2)^{-1} \ , \ j=1,2,\ldots,q \qquad (2.12)$$

Burada R_j^2, belirtme katsayını ifade etmektedir. Çoklu bağlantıyı belirlemek için elde edilen VIF değeri için bir ölçüt belirlenememiştir. Literatürde yer alan çalışmalarda bazı yazarlar, VIF' nün 5'in üzerinde değer alması halinde, güçlü bir çoklu bağlantı sorunu olduğunu savunurken bazıları da VIF' nün 10'un üzerinde olması halinde yüksek dereceli çoklu bağlantının varlığını savunur (Montgomery, Peck ve Vining, 2001).

2.3.3. Korelasyon matrisinin özdeğerler ve özvektörleri

Bağımsız değişkenlere ilişkin X'X korelasyon matrisinin özvektör ve özdeğerleri kullanılarak çoklu bağlantının tespiti yapılmaktadır. Korelasyon matrisinin özdeğerlerinin bir veya daha fazlası sıfır ya da sıfıra yakın ise bağımsız değişkenler arasında çoklu bağlantının olduğundan bahsedilir.

Çoklu bağlantıyı belirlemede özdeğer ve özvektörleri kullanarak elde edilen başka bir yaklaşım da koşul sayısı ölçütüdür. Özdeğerlerin en büyüğü ile en küçüğünün birbirine oranı ile koşul sayısı ölçütü bulunur. $\lambda_{\min}$ ve λ_{max} , X'X matrisinin sırasıyla minimum ve maksimum özdeğerleri olmak üzere koşul sayısı Eş. 2.13'de gösterilmiştir.

$$K = \frac{\lambda_{maks}}{\lambda_{\min}} \qquad (2.13)$$

Bu değer, 100'den küçük ise çoklu bağlantının olmadığı yorumu yapılırken 100 ile 1000 arasında ise orta derecede çoklu bağlantı olduğu ve 1000'den büyük ise yüksek derecede çoklu bağlantı olduğu söylenebilir (Montgomery, Peck ve Vining, 2001).

2.3.4. Belirtme katsayısı (R^2)

Herhangi bir bağımsız değişken ile diğer bağımsız değişkenler arasındaki belirtme katsayısı R^2 ile ifade edilir. R^2 belirtme katsayısının 1'e yakın değer alması durumunda

VIF değeri de büyük değer almaktadır ve bu durumda çoklu bağlantı sorunundan söz edilmektedir. R^2 belirtme katsayısının çok küçük olduğu ve VIF değerinin de 1'e yakın değer aldığı durumlarda ise çoklu bağlantının olmadığı yorumu yapılmaktadır.

2.3.5. X′X matrisinin determinantı

Bağımsız değişkenlere ilişkin X′X korelasyon matrisinin determinantı çoklu bağlantı sorununun tespit edilmesinde kullanılabilir. Korelasyon matrisinin determinantı 0 ve 1 arasında değer alır ve bu değer 0'a eşit ise tam çoklu bağlantı olduğu, 1'e eşit ise çoklu bağlantı olmadığı anlamına gelmektedir. Determinantın 0'a yakın değer alması halinde güçlü çoklu bağlantının varlığından söz edilir. Burada, matrisin determinantı küçüldükçe çoklu bağlantının derecesinin arttığı yorumu yapılabilir. Çoklu bağlantının derecesi ile ilgili bilgi veren bu yaklaşımda, hangi değişkenin çoklu bağlantıya sebep olduğuyla ilgili bir yorum yapılamamaktadır (Farrar ve Galuber, 1967).

2.4. Çoklu Bağlantının Ortadan Kaldırılması

Çoklu bağlantıyı gidermek için geliştirilen yöntemler genellikle, çoklu bağlantıya sebep olan bağımsız değişkenlerin modelden çıkarılması, modele yeni gözlem değerlerinin eklenmesi ve gözlem sayısının artırılması ya da bağımsız değişkenlerin dönüştürülmesiyle ilgilidir. Bu yöntemlerin dışında, çoklu bağlantıyı ortadan kaldırmak için geliştirilen yöntemlerden biri de regresyon katsayılarının tahmininde yanlı tahmin edicilerin kullanılmasıdır.

EKK tahmin edicisiyle elde edilen parametre tahminlerinin çoklu bağlantı durumunda yansız fakat büyük varyanslı olmasından dolayı bu sorunu ortadan kaldırmak amacıyla yanlı tahmin yöntemleri ileri sürülmüştür.

3. YANLI TAHMİN EDİCİLER

Yanlı tahmin yöntemlerinde, yansız olan bir tahmin ediciye yanlılık parametresi eklenerek varyansın küçültülmesi amaçlanmaktadır. Yansız tahmin edicinin varyansı yanlı tahmin ediciye göre daha büyük elde edilmektedir. Buradan yola çıkarak, yanlı ve yansız tahmin edicilerinin performansları HKO değerine bakılarak karşılaştırılabilir. Bu kısımda, çalışmaya konu olmuş yanlı tahmin ediciler özetlenecektir.

3.1. Ridge Tahmin Edicisi

Ridge tahmin edicisi, ilk olarak Hoerl ve Kennard (1970) tarafından çoklu bağlantı sorununu ortadan kaldırmak amacıyla geliştirilmiştir. Buradaki amaç, $X'X$ korelasyon matrisinin köşegen elemanlarına yanlılık parametresi olarak adlandırılan k sabitini ekleyerek HKO'ları küçültmektir. k yanlılık parametresine bağlı olarak geliştirilen Ridge tahmin edicisi Eş. 3.1'de ifade edilmiştir.

$$\hat{\beta}_R = (X'X + kI)^{-1}X'Y \ , \quad k \geq 0 \tag{3.1}$$

Eş. 3.1'de verilen Ridge tahmin edicisinin Eş. 2.2'de verilen doğrusal regresyon modeline göre kanonik formu Eş. 3.2'de verilmiştir.

$$\hat{\alpha}_R = (\Lambda + kI)^{-1}\Lambda\hat{\alpha} \tag{3.2}$$

Burada k, yanlılık veya shrinkage parametresi olarak ifade edilebilir. Modeldeki regresyon katsayıları ortalamaya göre standartlaştırılmaktadır. k değerinin sıfıra eşit olması halinde Ridge tahmin edicisi EKK tahmin edicisi haline gelmektedir (Draper ve Smith, 1998).

Yanlı tahmin edicilerin HKO değeri, tahmin edicinin parametre tahminin varyansları toplamı ile yanlılığının karesinin toplamına eşittir. Aynı zamanda, bu tahmin edicinin parametre tahminine olan uzaklığıdır. $\hat{\beta}_R$ 'nin β'ya olan uzaklığının karesinin beklenen değeri olarak ifade edilen Ridge tahmin edicisinin HKO değeri, Hoerl ve Kennard (1970) tarafından Eş. 3.3'te elde edilmiştir.

$$\text{HKO}\left(\hat{\beta}_R\right) = E[(\hat{\beta}_R - \beta)'[(\hat{\beta}_R - \beta)]$$
$$= \sigma^2 \sum_{i=1}^{q} \frac{\lambda_i}{(\lambda_i+k)^2} + \sum_{i=1}^{q} \frac{\alpha_i{}^2 k_i{}^2}{(\lambda_i+k)^2}$$
$$= \Upsilon_1(k) + \Upsilon_2(k) \quad (3.3)$$

$\lambda_1, \ldots, \lambda_q$, $X'X$ matrisinin özdeğerleri olmak üzere, $\Upsilon_1(k)$ terimi Ridge tahmin edicisinin parametre tahminlerinin varyansları toplamını, $\Upsilon_2(k)$ terimi ise yanlılık karesini göstermektedir.

Ridge tahmin edicisine ait beklenen değer ve varyans değerleri sırasıyla Eş. 3.4 ve Eş. 3.5'de verilmiştir.

$$E\left(\hat{\beta}_R\right) = E[(X'X + kI)^{-1}X'Y]$$
$$= E[(X'X + kI)^{-1}X'(X\beta + \varepsilon)]$$
$$= [(X'X + kI)^{-1}X'X\beta + (X'X + kI)^{-1}X'E(\varepsilon)]$$
$$= (X'X + kI)^{-1}X'X\beta \quad (3.4)$$
$$V\left(\hat{\beta}_R\right) = V[(X'X + kI)^{-1}X'Y]$$
$$= (X'X + kI)^{-1}X'V(Y)X(X'X + kI)^{-1}$$
$$= \sigma^2(X'X + kI)^{-1}X'X(X'X + kI)^{-1} \quad (3.5)$$

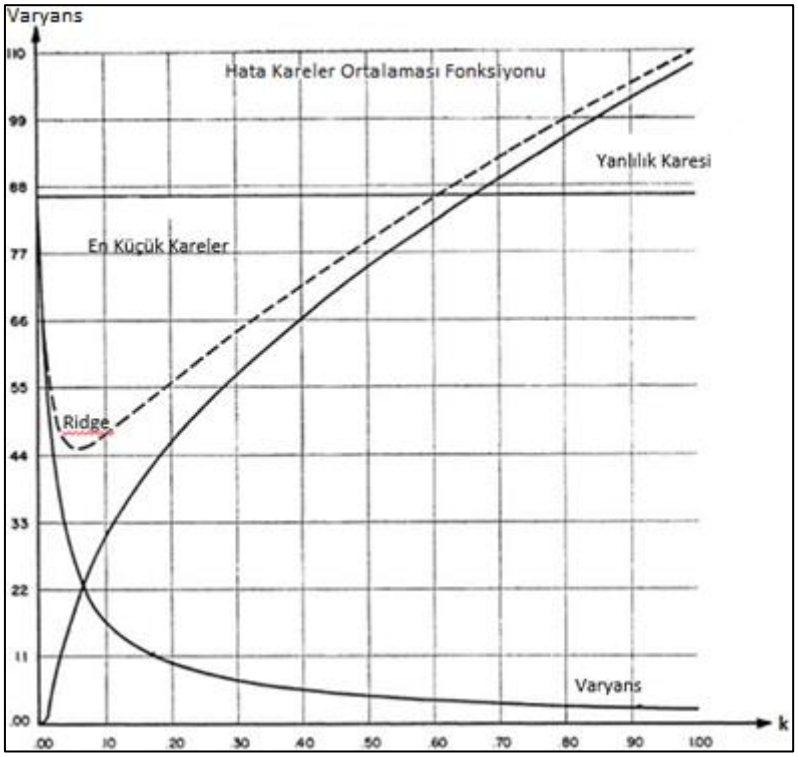
Burada σ^2 bilinmediği zaman, Eş. 3.6'da verilen Ridge tahmin edicisine dayalı σ^2'nin tahmini kullanılabilir.

$$\hat{\sigma}^2{}_R = \frac{(Y - X\hat{\beta}_R)'(Y - X\hat{\beta}_R)}{(n - q - 1)} \quad (3.6)$$

Ridge tahmin edicisine ilişkin yan değeri ise Eş. 3.7'de verilmiştir.

$$\text{yan}\left(\hat{\beta}_R\right) = E\left(\hat{\beta}_R\right) - \beta$$
$$= (X'X + kI)^{-1}X'X\beta - \beta$$
$$= [(X'X + kI)^{-1}X'X + I]\beta$$
$$= [[(X'X)^{-1}X'X + (X'X)^{-1}kI]^{-1} - I]\beta$$
$$= [I + k(X'X)^{-1})^{-1} - I]\beta$$

$$= [I - k[I + (X'X)^{-1}k]^{-1}(X'X)^{-1} - I]\beta$$
$$= [I - k[X'X + kI]^{-1} - I]\beta$$
$$= -k(X'X + kI)^{-1}\beta \quad (3.7)$$



Şekil 3.1. HKO fonksiyonları (Hoerl ve Kennard, 1970)

Hoerl ve Kennard (1970), Ridge tahmin edicisine ait HKO, varyans ve yanlılık arasındaki ilişkiyi göstermek amacıyla Şekil 3.1'deki grafiği kullanmıştır. Grafiğe göre, eksenler k parametresi ve varyanstan oluşmaktadır. Yatay eksene paralel olarak çizilen doğru EKK tahmin edicisi için HKO değerini göstermektedir. HKO, toplam varyans ve yanlılığın karesi değerlerinin toplamından elde edilmektedir. EKK tahmin edicisinin HKO değeri, parametre tahmin edicilerinin varyansına eşit olmasından dolayı k değerinin değişmesinden etkilenmemektedir. Çünkü tahmin edicilerin varyansı k'nın bir fonksiyonu değildir. Diğer eğriler ise Ridge tahmin edicisinin HKO değerine ilişkin $\Upsilon_1(k)$ ve $\Upsilon_2(k)$ terimleridir. $\Upsilon_1(k)$, Ridge tahmin edicisine ait yanlı standartlaştırılmış regresyon katsayılarının tahminlerinin toplam varyansını gösteren k'nın azalan fonksiyonudur. $\Upsilon_2(k)$ ise, ridge tahmin edicisinin yanlılığının karesini gösteren k'nın artan fonksiyonu olarak ifade edilmiştir (Hoerl ve Kennard, 1970).

Ridge tahmin edicisi ile EKK tahmin edicisi benzer özellikler gösterir. Bu tahmin ediciyi EKK tahmin edicisinden ayıran en önemli özelliği k yanlılık parametresine sahip olmasıdır. k parametresinin seçimi Ridge tahmin edicisinin performansını etkilediğinden dolayı oldukça önemlidir. Çünkü, Ridge tahmin edicisinin HKO değeri, EKK tahmin edicisinin HKO değerine kıyasla daha küçük hesaplanmaktadır ve bu durumu sağlamak için k değerinin oldukça küçük seçilmesi gerekmektedir. k parametresi hem HKO değerini küçültecek hem de yanlılığı kabul edilebilir seviyede tutacak ölçüde seçilmelidir.

En uygun k değerini belirlemek için bazı yöntemler önerilmiştir. Hoerl ve Kennard (1970) tarafından yanlı standartlaştırılmış regresyon katsayıları ile k parametresi arasındaki ilişkiyi grafiksel olarak açıklayan Ridge İzi yöntemi bunlardan biridir (Hoerl ve Kennard, 1970). Ridge izi, veride ortaya çıkan değişimlerden etkilenen parametrenin hangisi olduğunu göstermesinden dolayı hassasiyet analizi şeklinde de ifade edilmektedir (Akdeniz ve Kaçıranlar, 2001). Ridge izi grafiği, iki boyutlu bir kartezyen düzlemde yatay eksende k değerleri düşey eksende ise yanlı standartlaştırılmış regresyon katsayıları ($\hat{\beta}_R$) tahminleri olacak şekilde oluşturulmuştur. Grafikte, k değerleri arttıkça standartlaştırılmış regresyon katsayılarının tahminlerinin HKO değerinin azaldığı görülmüştür. Burada, k parametresinin belli bir değere kadar artmasından sonra, standartlaştırılmış regresyon katsayılarının değerindeki değişimin yavaşlaması ve yatay eksene paralel olduğu bir noktada katsayıların dengeye ulaşması söz konusudur. Yanlı standartlaştırılmış regresyon katsayılarının durağan olduğu konumdaki k değeri Ridge tahmin edicisi için en uygun yanlılık değeri olarak seçilmektedir (Hoerl ve Kennard, 1970). Hoerl ve Kennard (1970) tarafından k değeri $0<k<1$ aralığında önerilmiştir. En uygun k değerinin belirlenmesinde kullanılan yöntemlerden biri de VIF değeridir. Bu yaklaşıma göre, VIF değerinin 10'dan küçük olduğu durumlarda k parametresinin seçiminin yapılabileceği belirtilmiştir.

Grafik yöntemleri dışında k parametresinin seçiminde HKO değerini minimum yapmaya dayalı geliştirilen formüller de bulunmaktadır. Bunlardan ilki, Hoerl ve Kennard'ın (1970) geliştirdiği k'nın tahmin edicisi Eş. 3.8' deki gibidir.

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\max(\hat{\beta}_i^2)} \tag{3.8}$$

Burada, $\hat{\beta}_i^2$ standartlaştırılmış EKK tahmin edicisinin en büyük değerinin karesini ifade etmektedir.

Hoerl, Kennard ve Baldwin (1975) tarafından geliştirilen farklı bir k'nın tahmin edicisi ise Eş. 3.9'da verilmiştir.

$$\hat{k}_{HKB} = \frac{q\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \tag{3.9}$$

Eş. 2.7' deki yokluk hipotezini test etmek için, Ridge tahmin edicisine ait test istatistiği Eş. 3.10'da verilmiştir (Halawa ve El-Bassiouni, 2000).

$$t_{i(R)} = \frac{\hat{\beta}_{i(R)}}{S(\hat{\beta}_{i(R)})} \tag{3.10}$$

Burada, $\hat{\beta}_{i(R)}$, $\hat{\beta}_R$'nin i-inci elemanını, ve $S(\hat{\beta}_{i(R)})$, $\hat{\beta}_R$ 'nin varyans tahminin i.inci köşegen elemanının karekökünü ifade etmektedir. H_0 hipotezinin doğruluğu altında, bu test istatistiği (n-q-1) serbestlik dereceli t dağılımına sahiptir. Ridge tahmin edicisinin EKK tahmin edicisinden daha küçük HKO değerine sahip olmasından dolayı $t_{i(R)}$ test istatistiğinin mutlak değerinin $t_{(EKK)}$ test istatistiğinin mutlak değerinden daha büyük ve t test istatistiğine daha duyarlı olması beklenir (Halawa ve El-Bassiouni, 2000). Bu nedenle, H_0 hipotezinin testi için en uygun k parametresinin seçimi de önemlidir. Bu çalışmada, $\hat{k}_{HK}$ ve $\hat{k}_{HKB}$ değerleri için elde edilen test istatistikleri HK ve HKB olarak isimlendirilmiştir.

3.2. Liu Tahmin Edicisi

Ridge tahmin edicisine ait k yanlılık parametresi, tahmin edicinin doğrusal bir fonksiyonu şeklinde elde edilememektedir. Bu nedenle, yanlılık parametresi tahmin edicisinin doğrusal bir fonksiyonu olan bir tahmin edici geliştirilmiştir (Liu, 1993). Bu tahmin edici Akdeniz ve Kaçıranlar (1995) tarafından Liu (Linear Unified) tahmin edicisi olarak tanımlanmıştır. Liu (1993) tarafından elde edilen bu tahmin edici Eş. 3.11'de ifade edilmiştir.

$$\hat{\beta}_L = (X'X + I)^{-1}(X'Y + d\hat{\beta}) \ , \qquad 0 \leq d \leq 1 \tag{3.11}$$

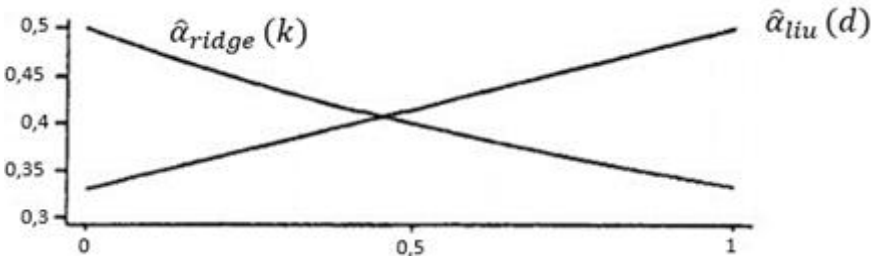
Eş. 2.3'de verilen EKK denkleminde, $X'Y = (X'X)\hat{\beta}$ olarak alındığında ve Eş. 3.11' de yerinde yazıldığında Liu tahmin edicisi Eş. 3.12'deki gibi elde edilmektedir.

$$\hat{\beta}_L = (X'X + I)^{-1}((X'X) + dI)\hat{\beta} \tag{3.12}$$

Eş. 3.11'de verilen Liu tahmin edicisinin Eş. 2.2'de verilen modele göre kanonik formu Eş. 3.13'de verilmiştir.

$$\hat{\alpha}_L = (\Lambda + I)^{-1}(Z'Y + d\hat{\alpha}) \tag{3.13}$$

Burada d, Liu tahmin edicisine ait yanlılık parametresi olup $0 \leq d \leq 1$ aralığında değerler almaktadır. Liu tahmin edicisi, Ridge tahmin edicisi ile Stein (1956) tahmin edicisinin avantajlı özelliklerinin birleşimi ile elde edilmiştir (Liu, 1993). Liu ileri sürdüğü bu tahmin edici ile EKK tahmin edicisini HKO değeri kriterine göre karşılaştırarak Liu tahmin edicisinin daha küçük HKO değerine sahip olduğu sonucuna ulaşmıştır (Liu, 1993; Kaçıranlar ve Sakallıoğlu, 2000). Liu tahmin edicisi, d parametresinin doğrusal fonksiyonu olduğu için d'nin seçimi k'nın seçimine göre daha kolay olmaktadır (Liu, 1993).



Şekil 3.2. Ridge ve Liu tahmin edicilerinin yanlılık parametreleri (Liu, 1993)

Şekil 3.2.'de, Liu ve Ridge tahmin edicilerine ait yanlılık parametrelerinin grafiği gösterilmiştir. Burada, Liu tahmin edicisinin d parametresinin doğrusal ve artan fonksiyonu olduğu görülüyor iken Ridge tahmin edicisinin k parametresinin doğrusal olmayan ve azalan bir fonksiyonu olduğu belirtilmiştir (Liu, 1993). Grafiğe göre, bu iki tahmin edicinin ortak bir noktada k ve d değeri alabildiği görülmektedir.

Regresyon katsayılarının tahmininde, Ridge ve Liu tahmin edicilerinin etkinlikleri bilinmeyen k ve d parametrelerine bağlıdır. Bilinmeyen parametrelerin yerine tahminlerinin kullanılması uygulamada daha pratiktir. (Liu, 1993). Bu bakımdan, Ridge tahmin edicisinin yanlılık parametresi olan k değerinin tahmin yöntemlerinde olduğu gibi d yanlılık parametresinin tahmini için de pek çok formül geliştirilmiştir. d yanlılık parametresinin tahmininin elde edilmesinde HKO değerini minimum yapan formüller önerilmiştir. Bunlardan bir tanesi, Liu (1993) tarafından geliştirilen d'nin tahmin edicisi Eş. 3.14'te verildiği gibidir.

$$\hat{d}_L = 1 - \hat{\sigma}^2 \left[\frac{\sum_{i=1}^{q} \frac{1}{(\lambda_i+1)}}{\sum_{i=1}^{q} \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i+1)^2}} \right] \tag{3.14}$$

Liu tahmin edicisinin HKO değeri Eş. 3.15'de verilmiştir.

$$\begin{aligned} &\text{HKO}\left(\hat{\beta}_L\right) = E[(\hat{\beta}_L - \beta)'[(\hat{\beta}_L - \beta)] \\ &= \sigma^2 \sum_{i=1}^{q} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2} + (d-1)^2 \sum_{i=1}^{q} \frac{{\alpha_i}^2}{(\lambda_i + 1)^2} \end{aligned} \tag{3.15}$$

$\lambda_1, \ldots, \lambda_q$, $X'X$ matrisinin özdeğerleri olmak üzere, denklemin ilk terimi Liu tahmin edicisinin parametre tahminlerinin varyansları toplamını, ikinci terim ise yanlılık karesini göstermektedir.

Liu tahmin edicisine ait beklenen değer ve varyansları sırasıyla Eş. 3.16 ve Eş. 3.17'de gösterilmiştir.

$$\begin{aligned} E\left(\hat{\beta}_L\right) &= E[(X'X + I)^{-1}(X'Y + d\hat{\beta})] \\ &= E[(X'X + I)^{-1}(X'X + dI)\hat{\beta}] \\ &= (X'X + I)^{-1}(X'X + dI)E(\hat{\beta}) \\ &= (X'X + I)^{-1}(X'X + dI)\beta \end{aligned} \tag{3.16}$$

$$\begin{aligned} V\left(\hat{\beta}_L\right) &= V[(X'X + I)^{-1}(X'Y + d\hat{\beta})] \\ &= V[(X'X + I)^{-1}(X'X + dI)\hat{\beta}] \\ &= \sigma^2(X'X + I)^{-1}(X'X + dI)(X'X)^{-1}(X'X + dI)(X'X + I)^{-1} \end{aligned} \tag{3.17}$$

Burada σ^2 bilinmediği zaman, Eş. 3.18'de verilen Liu tahmin edicisine dayalı σ^2'nin tahmini kullanılabilinir.

$$\hat{\sigma}^2{}_L = \frac{(Y - X\hat{\beta}_L)'(Y - X\hat{\beta}_L)}{(n - q - 1)} \tag{3.18}$$

Eş. 2.7'deki yokluk hipotezini test etmek için, Halawa ve Bassiouni'nin (2000) yaklaşımına dayanarak Liu tahmin edicisi için elde edilen test istatistiği Eş. 3.19' da verilmiştir (Melo, S.P. ve Kibria, 2021)

$$t_{i(L)} = \frac{\hat{\beta}_{i(L)}}{S(\hat{\beta}_{i(L)})} \tag{3.19}$$

Burada, $\hat{\beta}_{i(L)}$, $\hat{\beta}_L$'nin i-inci elemanını, ve $S(\hat{\beta}_{i(L)})$, $\hat{\beta}_L$ 'nın varyans tahminin i.inci köşegen elemanının karekökünü ifade etmektedir. H_0 hipotezinin doğruluğu altında, bu test istatistiği (n-q-1) serbestlik dereceli t dağılımına sahiptir.

3.3. Liu Tipi Tahmin Edicisi

Ridge tahmin edicisi ile HKO değerinin küçültülmesi ve $X'X$ matrisinin kötü koşullu olma durumunun düzeltilmesi amaçlanmaktadır (Liu,2004). Koşul sayısı ölçütüne göre, koşul sayısının yüksek olması $X'X$ matrisinin kötü koşullu olduğu anlamına gelmektedir (Liu,2003). Bu nedenle, kötü koşulluluğu ortadan kaldırmak için koşul sayısının küçük olması gerekmektedir. Bu durum, k yanlılık parametresinin büyük seçilmesi ile sağlanmaktadır. Ridge tahmin edicisi çoklu bağlantı durumunda, hem HKO değerinin azaltılmasına hem de yanlılık miktarının artmasına sebep olmaktadır. Artan yanlılık, HKO değerinin küçülmesine rağmen istenilen bir durum değildir. Bu nedenle, yanlılığın artmaması açısından k parametresinin küçük seçilmesi önemli bir husustur. Ancak bu durumda da, tahmin edici kararsız ve koşul sayısı kötü koşullu olmaya devam etmektedir. Görüldüğü üzere, k yanlılık parametresi ile $X'X + kI$ matrisinin koşul sayısı arasında zıt yönlü bir ilişki vardır (Liu, 2004). Bu nedenle, Ridge tahmin edicisi $X'X$ matrisinin kötü koşullu olma durumunu ortadan kaldırmak için yeterli değildir. Liu tahmin edicisi ise parametreyi daraltarak varyansın küçülmesine sebep olmaktadır. Ancak, Liu tahmin

edicisinin d yanlılık parametresinin doğrusal bir fonksiyonu şeklinde ifade edilmesinden dolayı, çoklu bağlantı durumunda EKK tahmin edicisi üzerindeki olumsuz sonuçlardan etkilenerek hatalı sonuçlar elde edilmesine yol açmaktadır.

Ridge ve Liu tahmin edicilerinin dezavantajları göz önünde bulundurularak, çoklu bağlantı sorununu ortadan kaldırmada yeterli olamadıkları durumların olmasından dolayı Liu (2003) tarafından iki parametreli yeni bir tahmin edici geliştirilmiştir. Bu tahmin edici Eş. 3.20'de verilmiştir.

$$\hat{\beta}_{LT} = (X'X + kI)^{-1}(X'Y - d\hat{\beta}^{*}) \qquad k > 0 \;\; , \; -\infty < d < \infty \tag{3.20}$$

Eş. 3.20'de verilen Liu Tipi tahmin edicisinin Eş.2.2'de verilen modele göre kanonik formu Eş. 3.21'de verilmiştir.

$$\hat{\alpha}_{LT} = (\Lambda + kI)^{-1}(Z'Y - d\hat{\alpha}^{*}) \tag{3.21}$$

Burada k ve d yanlılık parametreleridir. $\hat{\beta}^{*}$, β'nın herhangi bir tahmin edicisidir. k yanlılık parametresi, $X'X$ matrisinin kötü koşullu olma durumunu düzeltirken, d yanlılık parametresi HKO değerini düşürmeyi amaçlamaktadır. Bu yüzden, k ve d parametrelerinin seçimi için k'nın koşul sayısı 10 olacak biçimde ve d'nin HKO değeri minimum olacak biçimde formüller geliştirilmiştir. $(X'X + kI) < 10$ eşitliğini sağlayan k parametresinin tahmin edicisi Eş. 3.22'de gösterilmiştir.

$$\hat{k}_{LT} = \frac{\lambda_1 - 100 * \lambda_q}{99} \tag{3.22}$$

λ_1, $X'X$ matrisinin en büyük özdeğerini, λ_q ise $X'X$ matrisinin en küçük özdeğerini ifade eder. Buna bağlı olarak hesaplanan $\hat{d}$'ler Eş. 3.23 ve Eş. 3.24'deki gibidir.

$$\hat{d}_{LT1} = \frac{\sum_{i=1}^{q}(\hat{\sigma}_R^2 - \hat{k}_{LT}\,\hat{\alpha}_R^2))/(\lambda_i + \hat{k}_{LT})^2}{\sum_{i=1}^{q}(\lambda_i\hat{\alpha}_R^2 + \hat{\sigma}_R^2)/(\lambda_i + \hat{k}_{LT})^2} \tag{3.23}$$

$$\hat{d}_{LT2} = \frac{\sum_{i=1}^{q}(\lambda_i(\hat{\sigma}_R^2 - \hat{k}_{LT}\,\hat{\alpha}_R^2))/(\lambda_i + \hat{k}_{LT})^3}{\sum_{i=1}^{q}(\lambda_i(\lambda_i\hat{\alpha}_R^2 + \hat{\sigma}_R^2)/(\lambda_i + \hat{k}_{LT})^4]} \tag{3.24}$$

Halawa ve Bassiouni'nin (2000) çalışmasında Ridge tahmin edicisi için elde ettiği test istatistiği göz önüne alınarak, Eş. 2.7'deki yokluk hipotezini test etmek için Liu Tipi tahmin edicisine ait elde edilen test istatistiği Eş. 3.25'de verilmiştir.

$$t_{i(LT)} = \frac{\hat{\beta}_{i(LT)}}{S(\hat{\beta}_{i(LT)})} \tag{3.25}$$

Burada, $\hat{\beta}_{i(LT)}$, $\hat{\beta}_{LT}$'nin i-inci elemanını, ve $S(\hat{\beta}_{i(LT)})$, $\hat{\beta}_{LT}$ 'nın varyans tahminin i.inci köşegen elemanının karekökünü ifade etmektedir.

Liu (2003) çalışmasında, $\hat{\beta}^{*}$'ın hem EKK hem de ridge tahmin edicisi olduğu durumları göstermiştir. Bu çalışmada $\hat{\beta}^{*}$, $\hat{\beta}_{R}$ olarak seçilmiştir. Buna göre elde edilen Liu Tipi tahmin edicisi Eş. 3.26'da gösterilmiştir (Liu, 2003).

$$\begin{aligned}\hat{\beta}_{LT} &= (X'X + kI)^{-1}(X'Y - d\hat{\beta}_{R}) \\ &= (X'X + kI)^{-1}(X'Y) - d((X'X + kI)^{-2}X'Y) \\ &= (X'X + kI)^{-1}(I - d(X'X + kI)^{-1})X'Y\end{aligned} \tag{3.26}$$

Liu Tipi tahmin edicisine ait beklenen değer ve varyansları sırasıyla Eş. 3.27 ve Eş. 3.28'de gösterilmiştir.

$$\begin{aligned}E(\hat{\beta}_{LT}) &= E((X'X + kI)^{-1}(I - d(X'X + kI)^{-1})X'Y) \\ &= (X'X + kI)^{-1}(I - d(X'X + kI)^{-1})\,(X'X)\beta\end{aligned} \tag{3.27}$$

$$\begin{aligned}Var(\hat{\beta}_{LT}) &= \hat{s}^{2}(\hat{\beta}_{LT}) \\ &= \hat{\sigma}^{2}(X'X + kI)^{-1}(I - d(X'X + kI)^{-1})\,(X'X)(I - d(X'X + kI)^{-1})(X'X + kI)^{-1}\end{aligned} \tag{3.28}$$

Burada σ^{2} bilinmediği zaman, Eş. 3.29'da verilen Liu Tipi tahmin edicisine dayalı σ^{2}'nin tahmini kullanılabilinir.

$$\hat{\sigma}^{2}{}_{LT} = \frac{(Y - X\hat{\beta}_{LT})'(Y - X\hat{\beta}_{LT})}{(n - q - 1)} \tag{3.29}$$

4. UYGULAMA

Bu uygulamada, literatürde birçok çalışmada yaygın olarak kullanılan Hald'ın Portland Çimento verilerinden yararlanılmıştır. Söz konusu veriler, 4 bağımsız değişken ve 13 gözlemden oluşmaktadır. Bağımsız değişkenler; tricalcium aluminate (X_1), tetracalcium silicate (X_2), tetracalcium alumino ferrite (X_3), dicalcium silicate (X_4) dir. Y bağımlı değişkeni ise 1 gram çimentoya karşılık açığa çıkmakta olan ısı miktarını kalori cinsinden göstermektedir. Uygulamada, EKK, Ridge, Liu ve Liu Tipi tahmin edicilerinin parametre değerleri ve bu tahmin edicilere dayalı modelin katsayılarının test istatistikleri hesaplanarak elde edilen sonuçlar karşılaştırılmıştır. Hald'ın veri kümesindeki değişkenlerin seçilmesinde Halawa ve Bassiouni (2000) tarafından yapılan çalışmadan yararlanılmıştır.

Çizelge 4.1. Portland çimento verisi

X_1	X_2	X_3	X_4	Y
7	26	6	60	78,5
1	29	15	52	74,3
11	56	8	20	104,3
11	31	8	47	87,6
7	52	6	33	95,9
11	55	9	22	109,2
3	71	17	6	102,7
1	31	22	44	72,5
2	54	18	22	93,1
21	47	4	26	115,9
1	40	23	34	83,8
11	66	9	12	113,3
10	68	8	12	109,4

Çalışmada veriler standartlaştırılmış biçimde kullanılmıştır. Y değerleri ortalama etrafında standartlaştırılırken, X değerleri birim uzunlukta ölçeklendirme yöntemine göre standartlaştırılmıştır.

Birim uzunlukta ölçeklendirme yöntemi formülü Eş. 4.1'de verilmiştir.

$$x_s = \frac{x_{ij} - \bar{x}_j}{SST^{1/2}} \text{ , i} = 1,2, \dots, \text{n j} = 1,2, \dots, \text{q} \qquad (4.1)$$

$$SST = \sum_{i=1}^{n} y_i^2 - \left(\frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \right) \quad (4.2)$$

Ortalama etrafında standartlaştırma formülü Eş 4.3'te verilmiştir.

$$y_s = y_i - \bar{y} \quad (4.3)$$

Çizelge 4.2. Standartlaştırılmış veriler

$$Y_s = \begin{bmatrix} -16{,}92308 \\ -21{,}12308 \\ 8{,}87692 \\ -7{,}82308 \\ 0{,}47692 \\ 13{,}77692 \\ 7{,}27692 \\ -22{,}32308 \\ -2{,}32308 \\ -20{,}47692 \\ -11{,}62308 \\ 17{,}87692 \\ 13{,}97692 \end{bmatrix} \quad X_s = \begin{bmatrix} -0{,}022649 & -0{,}410983 & -0{,}260016 & 0{,}517395 \\ -0{,}317096 & -0{,}355329 & 0{,}145609 & 0{,}379423 \\ 0{,}173648 & 0{,}14557 & -0{,}169877 & -0{,}172465 \\ 0{,}173648 & -0{,}318227 & -0{,}169877 & 0{,}293191 \\ -0{,}022649 & 0{,}071351 & -0{,}260016 & 0{,}051739 \\ 0{,}173648 & 0{,}127005 & -0{,}124808 & -0{,}137972 \\ -0{,}218947 & 0{,}423827 & 0{,}235748 & -0{,}413916 \\ -0{,}317096 & -0{,}318227 & 0{,}461095 & 0{,}241451 \\ -0{,}268022 & 0{,}108454 & 0{,}280817 & -0{,}137972 \\ 0{,}664392 & -0{,}021405 & -0{,}350154 & -0{,}068986 \\ -0{,}317096 & -0{,}151265 & 0{,}506164 & 0{,}068986 \\ 0{,}173648 & 0{,}33107 & -0{,}124808 & -0{,}310437 \\ 0{,}124574 & 0{,}368173 & -0{,}169877 & -0{,}310437 \end{bmatrix}$$

Öncelikle, değişkenler arasında çoklu bağlantı olup olmadığını incelemek amacıyla korelasyon matrisi elde edilmiştir ve Çizelge 4.3'te verilmiştir.

Çizelge 4.3. X′X korelasyon matrisi

$$X'X = \begin{bmatrix} 1{,}000 & 0{,}2286 & -0{,}8241 & -0{,}2454 \\ 0{,}2286 & 1{,}000 & -0{,}1392 & -0{,}9730 \\ -0{,}8241 & -0{,}1392 & 1{,}000 & 0{,}0295 \\ -0{,}2454 & -0{,}9730 & 0{,}0295 & 1{,}000 \end{bmatrix}$$

X′X matrisi iki değişken arasındaki doğrusal ilişkiyi göstermektedir. Buna göre, X_2 ve X_4 değişkenleri arasında 0.824 büyüklüğünde ve ters yönlü ilişki var iken X_1 ve X_3 değişkenleri arasında 0.973 büyüklüğünde ve ters yönlü bir ilişki olduğu görülmektedir. Değişkenler arasındaki ilişki X′X matrisine göre yüksektir fakat çoklu bağlantıyı belirlemek için bu ölçüt yeterli olmamaktadır. Bu nedenle X′X matrisinin öz değerleri de incelenecektir.

X′X matrisinin öz değerleri büyükten küçüğe doğru sırasıyla λ_1=2.2357 , λ_2=1.5761 , λ_3=0.1866 , λ_4=0.0016 olarak elde edilmiştir.

Çizelge 4.4. X′X korelasyon matrisine ilişkin öz değer matrisi

$$\Lambda=\begin{bmatrix} 2,2357 & 0 & 0 & 0 \\ 0 & 1,5761 & 0 & 0 \\ 0 & 0 & 0,1866 & 0 \\ 0 & 0 & 0 & 0,0016 \end{bmatrix}$$

biçiminde oluşturulmuştur.

$k=\frac{\lambda_{maks}}{\lambda_{min}}$ koşul sayısı değeri yaklaşık 1397,3125 olarak hesaplanmıştır. Bu değer 1000'den büyük çıktığı için yüksek dereceli çoklu bağlantı sorunu söz konusudur. Bu nedenle, yanlı tahmin yöntemlerinin kullanılması önerilir. Çoklu bağlantının varlığında doğrusal regresyon modelinin EKK, Ridge, Liu ve Liu Tipi tahmin edicilerinden elde edilen sonuçları aşağıda incelenecektir.

4.1. EKK Tahmin Edicisi ile Elde Edilen Sonuçlar

Eş. 2.1'de verilen doğrusal regresyon modelinin Eş.2.3'de verilen EKK tahmin edicisi,

Eş. 4.4'deki gibi elde edilmiştir.

$$\hat{\beta}_{EKK}=\begin{bmatrix} 31,6060 \\ 27,4972 \\ 2,2600 \\ -8,3563 \end{bmatrix} \quad (4.4)$$

Kanonik formda elde edilen EKK tahmin edicisi Eş.4.5'de gösterilmiştir.

$$\hat{\alpha}_{EKK}=\begin{bmatrix} 20,2168 \\ 15,7784 \\ -0,4329 \\ -34,2360 \end{bmatrix} \quad (4.5)$$

Modelin σ^2 parametresinin EKK tahmin edicisi Eş. 2.10'da verilen formüle göre hesaplanmıştır ve $\hat{\sigma}^2$=5,9831 olarak elde edilmiştir. Standart hataları ise $s(\hat{\beta}_1)$ =15,1785 , $s(\hat{\beta}_2)$ = 39,0215 , $s(\hat{\beta}_3)$ = 16,7480 , $s(\hat{\beta}_4)$ = 41,1192 olarak elde edilmiştir.

Eş. 2.7'ye göre kurulan hipotez için EKK tahmin edicisine ait test istatistiği Eş. 2.8'de verilen test istatistiği kullanılarak Eş.4.6 'daki gibi elde edilmiştir.

$$t_{EKK} = \begin{bmatrix} 2{,}0823 \\ 0{,}7047 \\ 0{,}1349 \\ -0{,}2032 \end{bmatrix} \tag{4.6}$$

Eş. 2.7'de verilen H_0 hipotezinin testinde elde edilen test istatistiği ve p değerleri Çizelge 4.5'de gösterilmiştir.

Çizelge 4.5. Regresyon katsayılarının EKK tahmin edicisine dayalı test sonuçları

i	$\hat{\beta}_{i(EKK)}$	$S(\hat{\beta}_{i(EKK)})$	$t_{i(EKK)}$	$p_{i(EKK)}$
1	31,6060	15,1785	2,0823	0,0709
2	27,4972	39,0215	0,7047	0,5010
3	2,2600	16,7480	0,1349	0,8960
4	-8,3563	41,1192	-0,2032	0,8440

Çizelge 4.5' de elde edilen EKK testinin sonuçlarına göre, her bir i-inci değişken için $p \geq \alpha$=0,05 olmasından dolayı H_0 red edilemez ve buradan $\beta_1, \beta_2, \beta_3, \beta_4$ katsayılarının anlamlı olmadığı sonucuna ulaşılır.

4.2. Ridge Tahmin Edicisi ile Elde Edilen Sonuçlar

Ridge tahmin edicisine ait k yanlılık parametresinin 2 farklı değeri kullanılmıştır. İlk olarak, Hoerl, Kennard ve Baldwin (1975) tarafından önerilen ve Eş. 3.9'da verilen $\hat{k}_{HKB} = \frac{q\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ formülü kullanılarak ve $\hat{k}_{HKB}$ = 0.0131 olarak hesaplanmıştır.

Diğer yöntem olarak, Hoerl ve Kennard (1970) tarafından Eş. 3.8'de verilen $\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\max(\hat{\beta}_i^2)}$ formülü kullanılmıştır ve $\hat{k}_{HK}$ = 0.006 olarak elde edilmiştir.

Eş.3.1'de verilen Ridge tahmin edicisi, $\hat{k}_{HKB}$ ve $\hat{k}_{HK}$ değerleri için aşağıdaki gibi elde edilmiştir.

$$\hat{\beta}_{kHKB} = \begin{bmatrix} 26{,}4798 \\ 16{,}1666 \\ -3{,}1510 \\ -20{,}2144 \end{bmatrix} \quad (4.7)$$

$$\hat{\alpha}_{kHKB} = \begin{bmatrix} 2.2323 \\ 14.7450 \\ -0.4294 \\ -34.0369 \end{bmatrix} \quad (4.8)$$

$$\hat{\beta}_{kHK} = \begin{bmatrix} 27.3977 \\ 17.3914 \\ -2.2881 \\ -18.9737 \end{bmatrix} \quad (4.9)$$

$$\hat{\alpha}_{kHK} = \begin{bmatrix} 4.3109 \\ 15.2877 \\ -0.4313 \\ -34.1445 \end{bmatrix} \quad (4.10)$$

Modelin σ^2 parametresinin Ridge tahmin edicisi Eş. 3.6'da verilen formüle göre $\hat{k}_{HKB}$ ve $\hat{k}_{HK}$ değerleri için hesaplanmıştır. Buradan, $\hat{\sigma}^2(k_{HKB})$= 6,0847 olarak elde edilmiştir. Standart hataları ise $s(\widehat{\beta_1}(k_{HKB}))$=4,1524 , $s(\widehat{\beta_2}(k_{HKB}))$ = 4,8107 , $s(\widehat{\beta_3}(k_{HKB}))$ = 4,0843 ,$s(\widehat{\beta_4}(k_{HKB}))$=4,8578 olarak elde edilmiştir. $\hat{\sigma}^2(k_{HK})$=6,0424 olarak elde edilmiştir. $s(\widehat{\beta_1}(k_{HK}))$=5,0299 , $s(\widehat{\beta_2}(k_{HK}))$=8,6149 , $s(\widehat{\beta_3}(k_{HK}))$ =5,1351 , $s(\widehat{\beta_4}(k_{HK}))$ =8,9584 olarak elde edilmiştir.

Eş. 2.7'ye göre kurulan hipotez için Ridge tahmin edicisine ait Eş. 3.10'da verilen test istatistiği kullanılarak Eş.4.11 ve Eş.4.12 elde edilmiştir.

$$t_{kHKB} = \begin{bmatrix} 6{,}3770 \\ 3{,}3606 \\ -0{,}7715 \\ -4{,}1613 \end{bmatrix} \quad (4.11)$$

$$t_{kHK} = \begin{bmatrix} 5{,}4470 \\ 2{,}0188 \\ -0{,}4456 \\ -2{,}1180 \end{bmatrix} \quad (4.12)$$

Eş. 2.7'de verilen H_0 hipotezini test etmek amacıyla $\hat{k}_{HKB}$ ve $\hat{k}_{HK}$ değerleri için elde edilen test istatistiği ve p değerleri Çizelge 4.6'da gösterilmiştir.

Çizelge 4.6. Regresyon katsayılarının Ridge tahmin edicisine dayalı test sonuçları

	HK				HKB			
i	$\hat{\beta}_{i(R)}$	$S(\hat{\beta}_{i(R)})$	$t_{i(R)}$	$p_{i(R)}$	$\hat{\beta}_{i(R)}$	$S(\hat{\beta}_{i(R)})$	$t_{i(R)}$	$p_{i(R)}$
1	27,3977	5,0299	5,4470	0,0006	26,4798	4,1524	6,3770	0,0002
2	17,3914	8,6149	2,0188	0,0782	16,1666	4,8107	3,3606	0,0099
3	-2,2881	5,1351	-0,4456	0,6677	-3,1510	4,0843	-0,7715	0,4626
4	-18,9737	8,9584	-2,1180	0,0670	-20,2144	4,8578	-4,1613	0,0032

Çizelge 4.6 'da elde edilen sonuçlara göre, Ridge tahmin edicisinin her iki testinin de standart hatasının Çizelge 4.5 'de verilen EKK tahmin edicisinin standart hatasından daha küçük olduğu söylenebilir. HK testinin sonuçlarına göre, sadece β_1 'nin anlamlı olduğu görülmektedir. Yani, X_1 bağımsız değişkeni HK testinin sonuçlarına göre anlamlıdır yorumu yapılabilir. HKB testinin sonuçlarına göre, β_3 dışındaki diğer katsayılar anlamlı çıkmıştır. Burada da, X_3 değişkeninin HKB testinin sonuçlarına göre anlamlı olmadığı söylenebilir.

4.3. Liu Tahmin Edicisi ile Elde Edilen Sonuçlar

Liu tahmin edicisine ait d yanlılık parametresi için, Liu (1993) tarafından geliştirilen ve Eş.3.14'te verilen $\hat{d}_L = 1 - \hat{\sigma}^2 \left[\frac{\sum_{i=1}^{q} \frac{1}{(\lambda_i+1)}}{\sum_{i=1}^{q} \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i+1)^2}} \right]$ formülü kullanılmıştır ve $\hat{d}_L$ değeri 0,8976 olarak bulunmuştur.

Eş.3.12'de verilen Liu tahmin edicisi $\hat{d}_L$ değeri için Eş.4.13'deki gibi elde edilmiştir.

$$\hat{\beta}_{(d_L)} = \begin{bmatrix} 29,6813 \\ 25,9809 \\ 1,2534 \\ -8,8875 \end{bmatrix} \tag{4.13}$$

Modelin σ^2 parametresinin Liu tahmin edicisi Eş. 3.18'de verilen formüle göre $\hat{d}_L$ değeri için hesaplanmıştır. Buna göre, $\hat{\sigma}^2(d_L)$ değeri 6,3553 olarak hesaplanmıştır.

$s(\widehat{\beta_1}(d_L))$=14,0676 , $s(\widehat{\beta_2}(d_L))$=36,1097, $s(\widehat{\beta_3}(d_L))$=15,5168 , $s(\widehat{\beta_4}(d_L))$=38,0498 olarak elde edilmiştir.

Eş. 2.7'ye göre kurulan hipotez için Liu tahmin edicisine ait Eş. 3.19'da verilen test istatistiği kullanılarak Eş.4.14 elde edilmiştir.

$$t_{d_L} = \begin{bmatrix} 2,1099 \\ 0,7195 \\ 0,0808 \\ -0,2336 \end{bmatrix} \quad (4.14)$$

Eş. 2.7'de verilen H_0 hipotezini test etmek amacıyla Liu testi için elde edilen test istatistiği ve p değerleri Çizelge 4.7'de gösterilmiştir.

Çizelge 4.7. Regresyon katsayılarının Liu tahmin edicisine dayalı test sonuçları

i	$\hat{\beta}_{(d_L)}$	$s(\hat{\beta}(d_L))$	t_{d_L}	p_{d_L}
1	29,6813	14,0676	2,1099	0,0679
2	25,9809	36,1097	0,7195	0,4923
3	1,2534	15,5168	0,0808	0,9376
4	-8,8875	38,0498	-0,2336	0,8212

Çizelge 4.7 'da elde edilen sonuçlara göre, Liu tahmin edicisinin ait test istatistiğinin standart hatasının Çizelge 4.5 'de verilen EKK tahmin edicisinin standart hatasından daha küçük olduğu söylenebilir. ve $\beta_1, \beta_2, \beta_3, \beta_4$ katsayılarının anlamlı olmadığı görülmektedir.

4.4. Liu Tipi Tahmin Edicisi ile Elde Edilen Sonuçlar

Liu Tipi tahmin edicisine ait Eş.3.22'de verilen $\hat{k}_{LT}$ parametresi 0,0209, Eş.3.23 ve Eş.3.24'de verilen $\hat{d}_{LT1}$ ve $\hat{d}_{LT2}$ parametreleri de sırasıyla -0,00005 ve -0,0482 olarak edilmiştir.

Eş.3.20'de verilen Liu Tipi tahmin edicisi Eş.4.15 ve Eş.4.16'daki gibi elde edilmiştir. Burada, $\hat{\beta}^*$ yerine $\hat{\beta}_R$ parametresi kullanılmıştır.

$$\hat{\beta}_{LT1} = \begin{bmatrix} 25{,}8599 \\ 15{,}7756 \\ -3{,}6703 \\ -20{,}5655 \end{bmatrix} \tag{4.15}$$

$$\hat{\beta}_{LT2} = \begin{bmatrix} 29{,}6645 \\ 18{,}1758 \\ -0{,}4829 \\ -18{,}4107 \end{bmatrix} \tag{4.16}$$

Modelin σ^2 parametresinin Liu Tipi tahmin edicisi Eş. 3.29'da verilen formüle göre hesaplanmıştır. Buna göre, $\hat{\sigma}^2(d_{LT1})$ değeri 6,1419 olarak hesaplanmıştır. $s(\widehat{\beta_1}(d_{LT1}))$=3,8563, $s(\widehat{\beta_2}(d_{LT1}))$=3,4924 , $s(\widehat{\beta_3}(d_{LT1}))$=3,7466 , $s(\widehat{\beta_4}(d_{LT1}))$=3,3999 olarak elde edilmiştir. $\hat{\sigma}^2(d_{LT2})$ değeri ise 6,1484 olarak hesaplanmıştır. $s(\widehat{\beta_1}(d_{LT2}))$=5,5994, $s(\widehat{\beta_2}(d_{LT2}))$=9.2225 , $s(\widehat{\beta_3}(d_{LT2}))$=5,6808 , $s(\widehat{\beta_4}(d_{LT2}))$=9,5697 olarak elde edilmiştir.

Eş. 2.7'ye göre kurulan hipotez için Liu Tipi tahmin edicisine ait Eş. 3.25'te verilen test istatistiği kullanılarak Eş.4.17 ve Eş.4.18 elde edilmiştir.

$$t_{d_{LT1}} = \begin{bmatrix} 6{,}7059 \\ 4{,}5172 \\ -0{,}9796 \\ -6{,}0488 \end{bmatrix} \tag{4.17}$$

$$t_{d_{LT2}} = \begin{bmatrix} 5{,}2978 \\ 1{,}9708 \\ -0{,}0850 \\ -1{,}9239 \end{bmatrix} \tag{4.18}$$

Eş. 2.7'de verilen H_0 hipotezini test etmek amacıyla Liu Tipi testinin her iki değeri için elde edilen test istatistiği ve p değerleri Çizelge 4.8'de gösterilmiştir.

Çizelge 4.8. Regresyon katsayılarının Liu Tipi tahmin edicisine dayalı test sonuçları

	$\hat{d}_{LT1}$				$\hat{d}_{LT2}$			
i	$\hat{\beta}_{i(LT1)}$	$S(\hat{\beta}_{i(LT1)})$	$t_{i(LT1)}$	$p_{i(LT1)}$	$\hat{\beta}_{i(LT2)}$	$S(\hat{\beta}_{i(LT2)})$	$t_{i(LT2)}$	$p_{i(LT2)}$
1	25,8599	3,8563	6,7059	0,0002	29,6645	5,5994	5,2978	0,0007
2	15,7756	3,4924	4,5172	0,0020	18,1758	9,2225	1,9708	0,0842
3	-3,6703	3,7466	-0,9796	0,3560	-0,4829	5,6808	-0,0850	0,9343
4	-20,5655	3,3999	-6,0488	0,0003	-18,4107	9,5697	-1,9239	0,0906

Çizelge 4.8 ‘de elde edilen sonuçlara göre, Liu Tipi tahmin edicisinin her iki değeri için elde edilen testlerin standart hatasının Çizelge 4.5 ‘de verilen EKK tahmin edicisinin standart hatasından daha küçük olduğu görülmektedir. Ayrıca, LT1 testinin standart hatasının Ridge ve Liu tahmin edicilerine ait tüm testlerin standart hatasından daha küçük olduğu tespit edilmiştir. LT1 testi için $\beta_1, \beta_2, \beta_4$ anlamlı iken, LT2 testi için sadece β_1 ‘nin anlamlı olduğu görülmektedir. Yani, X_1 , X_2 ve X_4 bağımsız değişkenleri LT1 testinin sonuçlarına göre anlamlıdır ve LT2 testinin sonuçlarına göre X_1 bağımsız değişkeninin anlamlı olduğu söylenebilir.

5. SİMÜLASYON ÇALIŞMASI

Bu bölümde simülasyon çalışması yoluyla elde edilen EKK, Ridge, Liu ve Liu Tipi tahmin edicilerinin parametre tahmin değerleri iki aşamalı olarak analiz edilecektir. Elde edilen test istatistiklerinin performanslarının değerlendirilmesi için simülasyon tekniği kullanılmıştır. İlk aşamada, bu tahmin edicilere dayalı test istatistiklerinin deneysel 1.tip hata oranları elde edilmiştir. İkinci aşamada ise, nominal α değeri 0,05'e yakın değer alanlar için test istatistiklerinin güç değerleri hesaplanmıştır. Hesaplanan güç değerlerine göre karşılaştırmalar yapılmıştır. Nominal α değeri 0,05'ten ciddi derecede uzak yani 0,025 ve 0,075 aralığının dışında değer alanlar ise değerlendirmeye alınmamıştır (Bradley, 1978).

Mcdonald ve Galerneau (1975) ve Kibria (2003) çalışmalarında bağımsız değişkenleri Eş. 5.1'deki gibi ifade etmişlerdir.

$$x_{ij} = (1-\rho^2)^{1/2} z_{ij} + \rho z_{iq+1} \ , \quad i = 1,2,\dots,n \ , j = 1,2..,q \tag{5.1}$$

Burada, z_{ij} standart normal rasgele değişkenleri, ρ^2 herhangi iki bağımsız değişken arasındaki korelasyonu ifade etmektedir. $\rho = 0{,}85,\ 0{,}95$ ve $0{,}99$ değerleri için örneklem büyüklüğü 20,30,50,100 ve q=4,6,10 alınarak simülasyon çalışması yapılmıştır. Eş. 5.1'de verilen vektörlerin her biri merkezileştirilmiş ve ölçeklendirilmiştir.

y_i bağımlı değişkeni Eş.5.2'deki gibi verilmiştir.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_q x_{iq} + \varepsilon_i \ , \quad i = 1,2,\dots,n \tag{5.2}$$

Burada y_i bağımlı değişken, ε_i, 0 ortalamalı ve σ^2 varyanslı bağımsız normal dağılıma sahip hata terimi olmak üzere her biri ortalama etrafında merkezileştirilmiştir. Halawa ve Bassiouni'nin (2000) çalışmasında kullanılan Most Favorable (MF) yöntemi β_i parametresi için uygulanmıştır. Yani β_i, $X'X$ matrisinin en büyük öz değerine karşılık gelen öz vektörlerin normalleştirilmiş halidir. Simülasyon her durum için 1000 tekrar olarak çalıştırılmıştır ve standart sapma değerleri σ=0.5,1,2 olarak alınmıştır.

Simülasyon çalışmasının ilk aşaması olarak elde edilen testlerin deneysel 1.tip hata oranları Çizelge 5.1 ve Çizelge 5.9 arasında sunulmuştur.

Çizelge 5.1. ρ=0.85 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0492	0,0500	0,0504	0,0502	0,0491	0,0499	0,0501	0,0494	0,0500	0,0504	0,0499	0,0502
HK	0,0528	0,0555	0,0573	0,0590	0,0502	0,0540	0,0542	0,0552	0,0484	0,0522	0,0523	0,0539
HKB	0,0608	0,0763	0,0834	0,0861	0,0479	0,0619	0,0667	0,0710	0,0327	0,0485	0,0540	0,0572
L	0,0830	0,1152	0.1289	0,1385	0,0614	0,1042	0,1346	0,1486	0,0473	0,0738	0,1109	0,1467
LT1	0,0221	0,0214	0,0206	0,0177	0,0376	0,0416	0,0425	0,0422	0,0075	0,0426	0,0457	0,0475
LT2	0,2125	0,2314	0,2488	0,2650	0,0876	0,1082	0,0986	0,0979	0,0269	0,0724	0,0867	0,0839

Çizelge 5.2. ρ=0.95 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0493	0,0502	0,0491	0,0499	0,0495	0,0497	0,0511	0,0508	0,0494	0,0498	0,0514	0,0510
HK	0,0525	0,0563	0,0549	0,0562	0,0514	0,0528	0,0557	0,0553	0,0482	0,0518	0,0550	0,0539
HKB	0,0590	0,0771	0,0818	0,0863	0,0457	0,0549	0,0612	0,0633	0,0322	0,0452	0,0515	0,0537
L	0,0692	0,0895	0,0948	0,1036	0,0578	0,0862	0,1042	0,1164	0,0479	0,0650	0,0886	0,1152
LT1	0,0400	0,0467	0,0467	0,0469	0,0281	0,0435	0,0482	0,0493	0,0012	0,0331	0,0447	0,0499
LT2	0,0963	0,1137	0,1219	0,1168	0,0716	0,0997	0,1185	0,1262	0,0934	0,0768	0,1007	0,1241

Çizelge 5.3. ρ=0.99 ve σ=0.5 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0500	0,0496	0,0514	0,0493	0,0507	0,0510	0,0493	0,0507	0,0508	0,0498	0,0506	0,0491
HK	0,0523	0,0549	0,0573	0,0553	0,0523	0,0532	0,0529	0,0546	0,0494	0,0523	0,0528	0,0523
HKB	0,0541	0,0620	0,0665	0,0650	0,0436	0,0494	0,0511	0,0534	0,0328	0,0432	0,0472	0,0497
L	0,0549	0,0593	0,0646	0,0623	0,0529	0,0611	0,0628	0,0676	0,0503	0,0542	0,0601	0,0661
LT1	0,0304	0,0465	0,0528	0,0542	0,0144	0,0421	0,0505	0,0577	0,0000	0,0297	0,0478	0,0615
LT2	0,2871	0,3049	0,2998	0,3167	0,2998	0,2879	0,2871	0,2940	0,3379	0,3737	0,3621	0,3569

σ değerinin 0,5 olduğu farklı ρ değerleri için Çizelge 5.1, Çizelge 5.2 ve Çizelge 5.3 incelendiğinde; EKK, HK ve HKB testlerinin deneysel 1.tip hata oranlarının 0,05'e yakın

değerler aldığı görülmektedir. Örneklem büyüklüğü arttıkça ρ değerinin yüksek seviyelerinde LT1 testinin deneysel 1.tip hata oranının 0,05'e yakın çıktığı söylenebilir. Testler arasında en kötü performans gösteren test, deneysel 1.tip hata oranında önemli ölçüde sapma görülen LT2 testidir.

Çizelge 5.4. ρ=0.85 ve σ=1 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0497	0,0488	0,0505	0,0506	0,0502	0,0504	0,0499	0,0503	0,0497	0,0492	0,0508	0,0501
HK	0,0521	0,0521	0,0536	0,0555	0,0505	0,0530	0,0527	0,0543	0,0478	0,0505	0,0531	0,0527
HKB	0,0513	0,0564	0,0594	0,0627	0,0438	0,0515	0,0538	0,0574	0,0321	0,0435	0,0492	0,0518
L	0,0557	0,0662	0,0716	0,0760	0,0484	0,0629	0,0695	0,0760	0,0467	0,0489	0,0594	0,0712
LT1	0,0224	0,0201	0,0185	0,0201	0,0378	0,0411	0,0416	0,0424	0,0075	0,0412	0,0466	0,0475
LT2	0,1917	0,2484	0,2585	0,2665	0,0847	0,1071	0,1061	0,1058	0,0227	0,0692	0,0907	0,0942

Çizelge 5.5. ρ=0.95 ve σ=1 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0490	0,0501	0,0505	0,0495	0,0508	0,0497	0,0506	0,0492	0,0493	0,0488	0,0491	0,0507
HK	0,0512	0,0535	0,0547	0,0537	0,0516	0,0535	0,0545	0,0540	0,0474	0,0503	0,0521	0,0528
HKB	0,0495	0,0556	0,0589	0,0594	0,0435	0,0499	0,0526	0,0534	0,0319	0,0416	0,0473	0,0502
L	0,0521	0,0594	0,0655	0,0650	0,0497	0,0558	0,0627	0,0677	0,0480	0,0478	0,0549	0,0677
LT1	0,0410	0,0461	0,0473	0,0468	0,0266	0,0424	0,0470	0,0478	0,0007	0,0295	0,0404	0,0469
LT2	0,1103	0,1122	0,1211	0,1181	0,0592	0,0892	0,1033	0,1179	0,0973	0,0553	0,0665	0,0869

Çizelge 5.6. ρ=0.99 ve σ=1 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0493	0,0497	0,0506	0,0498	0,0490	0,0489	0,0497	0,0495	0,0494	0,0511	0,0507	0,0512
HK	0,0515	0,0538	0,0554	0,0550	0,0501	0,0529	0,0530	0,0541	0,0480	0,0525	0,0531	0,0537
HKB	0,0488	0,0529	0,0557	0,0561	0,0418	0,0476	0,0496	0,0511	0,0315	0,0437	0,0478	0,0495
L	0,0498	0,0517	0,0542	0,0543	0,0484	0,0507	0,0527	0,0545	0,0488	0,0498	0,0524	0,0565
LT1	0,0235	0,0401	0,0461	0,0491	0,0085	0,0324	0,0419	0,0477	0,0000	0,0192	0,0354	0,0448
LT2	0,2304	0,2150	0,2106	0,2168	0,2534	0,2268	0,2153	0,2012	0,3142	0,3606	0,3356	0,3155

σ değerinin 1 olduğu farklı ρ değerleri için Çizelge 5.4 , Çizelge 5.5 ve Çizelge 5.6 incelendiğinde; EKK, HK ve HKB testlerinin deneysel 1.tip hata oranlarının 0,05'e yakın değerler aldığı görülmektedir. ρ değerinin yüksek seviyelerinde L testinin deneysel 1.tip hata oranları 0,05'e yaklaşmaktadır. LT1 testinin deneysel 1.tip hata oranının, ρ değerinin ve değişken sayısının artmasından etkilenerek 0,05'e yaklaştığı gözlemlenmiştir. Burada da, LT2 testinin deneysel 1.tip hata oranı 0,05 değerinden oldukça büyük çıktığı için testler arasında en kötü performansı gösteren test olmuştur.

Çizelge 5.7. ρ=0.85 ve σ=2 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0508	0,0483	0,0505	0,0490	0,0497	0,0509	0,0500	0,0505	0,0496	0,0510	0,0498	0,0506
HK	0,0514	0,0513	0,0539	0,0531	0,0503	0,0533	0,0537	0,0539	0,0481	0,0525	0,0517	0,0532
HKB	0,0483	0,0511	0,0548	0,0551	0,0428	0,0495	0,0519	0,0531	0,0317	0,0435	0,0466	0,0500
L	0,0503	0,0543	0,0595	0,0602	0,0454	0,0521	0,0551	0,0568	0,0457	0,0452	0,0472	0,0523
LT1	0,0250	0,0191	0,0190	0,0189	0,0389	0,0428	0,0427	0,0425	0,0077	0,0418	0,0454	0,0481
LT2	0,1863	0,2398	0,2586	0,2725	0,0867	0,1071	0,1081	0,1096	0,0230	0,0732	0,0905	0,0964

Çizelge 5.8. ρ=0.95 ve σ=2 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0502	0,0508	0,0500	0,0498	0,0496	0,0485	0,0502	0,0492	0,0506	0,0489	0,0505	0,0508
HK	0,0515	0,0536	0,0541	0,0543	0,0510	0,0521	0,0540	0,0530	0,0491	0,0506	0,0526	0,0532
HKB	0,0482	0,0526	0,0534	0,0549	0,0428	0,0479	0,0506	0,0510	0,0326	0,0426	0,0460	0,0490
L	0,0497	0,0535	0,0547	0,0563	0,0471	0,0485	0,0533	0,0548	0,0487	0,0453	0,0495	0,0526
LT1	0,0403	0,0465	0,0469	0,0469	0,0278	0,0417	0,0465	0,0471	0,0009	0,0296	0,0398	0,0459
LT2	0,0860	0,1164	0,1185	0,1211	0,0571	0,0863	0,1066	0,1139	0,0860	0,0483	0,0653	0,0798

Çizelge 5.9. ρ=0.99 ve σ=2 iken testlerin deneysel 1.tip hata oranları

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0505	0,0512	0,0497	0,0516	0,0507	0,0501	0,0498	0,0508	0,0500	0,0499	0,0496	0,0495
HK	0,0523	0,0545	0,0538	0,0560	0,0508	0,0536	0,0536	0,0551	0,0493	0,0514	0,0523	0,0522
HKB	0,0486	0,0523	0,0525	0,0556	0,0426	0,0486	0,0499	0,0521	0,0325	0,0422	0,0467	0,0483
L	0,0503	0,0514	0,0505	0,0527	0,0494	0,0495	0,0500	0,0519	0,0493	0,0478	0,0484	0,0498
LT1	0,0232	0,0388	0,0439	0,0489	0,0075	0,0316	0,0400	0,0454	0,0000	0,0156	0,0309	0,0414
LT2	0,1839	0,1894	0,1900	0,2054	0,2316	0,1953	0,1894	0,1832	0,3080	0,3368	0,3196	0,3088

σ değerinin 2 olduğu farklı ρ değerleri için Çizelge 5.7 , Çizelge 5.8 ve Çizelge 5.9 incelendiğinde; EKK, HK, HKB ve L testlerinin deneysel 1.tip hata oranlarının 0,05 değerine yakın sonuçlar verdiği gözlemlenmiştir.LT1 testinin deneysel 1.tip hata oranının değişken sayısı arttıkça azaldığı söylenebilir.LT2 testi diğer durumlarla benzer şekilde 0,05 değerinden oldukça uzak değerler almıştır.

Genel olarak simülasyon sonuçları incelendiğinde; EKK, HK ve HKB testlerinin tüm durumlarda deneysel 1.tip hata oranlarının 0,05 değerine yakın çıktığı görülmektedir. L testinin deneysel 1.tip hata oranının ρ değerinin yüksek seviyelerinde daha iyi sonuçlar verdiği söylenebilir. LT1 testi ile LT2 testi karşılaştırıldığında, LT1 testinin daha üstün performans gösterdiği görülmektedir. LT1 testinin deneysel 1.tip hata oranı için, ρ değerinin düşük seviyelerinde standart sapmayı dikkate almaksızın, değişken sayısının artmasıyla 0,05'e yaklaştığı yorumu yapılabilir. Ayrıca, ρ değerinin yüksek seviyelerinde LT1 testinin deneysel 1.tip hata oranının 0,05'e yakın değerler aldığı görülmektedir.

Simülasyon çalışmasının ikinci aşamasında, deneysel 1.tip hata oranının 0,05'ten ciddi derecede uzak olduğu değerler testlerin güç değerinin hesaplanmasında göz ardı edilmiştir. 0,025 ve 0,075 aralığında elde edilen deneysel 1.tip hata değerleri testlerin güç hesaplanmasında dikkate alınmıştır (Bradley, 1978). Bu aralığın dışında bulunan değerler (*) ile ifade edilerek hesaplamaya dahil edilmemiştir.

Testlerin güç değerleri Çizelge 5.10 ve Çizelge 5.18 arasında sunulmuştur.

Çizelge 5.10. ρ=0.85 ve σ=0,5 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0895	0,1028	0,1055	0,1025	0,0696	0,0844	0,0838	0,0856	0,0559	0,0651	0,0678	0,0705
HK	0,1186	0,1462	0,1488	0,1506	0,0803	0,1030	0,1052	0,1077	0,0567	0,0724	0,0771	0,0804
HKB	0,1821	*	*	*	0,0877	0,1383	0,1459	0,1552	0,0410	0,0746	0,0869	0,0950
L	*	*	*	*	0,1004	*	*	*	0,0548	0,0999	*	*
LT1	0,1035	0,1159	0,1168	0,1197	0,0699	0,1061	0,1085	0,1120	*	0,0703	0,0861	0,0959
LT2	*	*	*	*	*	*	*	*	*	*	*	*

Çizelge 5.11. ρ=0.95 ve σ=0,5 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0923	0,0935	0,1004	0,1041	0,0707	0,0780	0,0798	0,0844	0,0558	0,0640	0,0673	0,0683
HK	0,1204	0,1332	0,1505	0,1550	0,0822	0,0970	0,1015	0,1098	0,0551	0,0724	0,0772	0,0800
HKB	0,1679	*	*	*	0,0972	0,1367	0,1516	0,1650	0,0398	0,0812	0,0954	0,1069
L	0,1510	*	*	*	0,1097	*	*	*	0,0524	0,1192	*	*
LT1	0,2474	0,2676	0,2686	0,2697	0,0604	0,0718	0,0692	0,0713	*	0,0649	0,0715	0,0732
LT2	*	*	*	*	0,0721	*	*	*	*	*	*	*

Çizelge 5.12. ρ=0.99 ve σ=0,5 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0905	0,1007	0,1048	0,1083	0,0716	0,0841	0,0844	0,0875	0,0561	0,0655	0,0697	0,0725
HK	0,1136	0,1334	0,1438	0,1499	0,0808	0,1008	0,1015	0,1050	0,0564	0,0714	0,0767	0,0796
HKB	0,1673	0,2231	0,2487	0,2631	0,0813	0,1166	0,1234	0,1302	0,0394	0,0674	0,0780	0,0846
L	0,1137	0,1437	0,1558	0,1662	0,0810	0,1135	0,1231	0,1335	0,0561	0,0761	0,0919	0,1071
LT1	0,2114	0,2617	0,2758	0,2869	*	0,1938	0,2127	0,2303	*	0,1159	0,1596	0,1825
LT2	*	*	*	*	*	*	*	*	*	*	*	*

σ değerinin 0,5 olduğu farklı ρ değerleri için Çizelge 5.10 , Çizelge 5.11 ve Çizelge 5.12 incelendiğinde; genel olarak EKK testinin diğer testlere göre daha düşük güç değerine sahip olduğu görülmektedir. Değişken sayısından bağımsız olarak incelendiğinde, örneklem büyüklüğü arttıkça testlerin güç değerleri artmaktadır. Değişken sayısı dikkate

alınırsa, değişken sayısındaki artış karşısında testlerin güç değerlerinde azalma görülmektedir. Çizelge 5.12'de LT1 testi için güç değeri, q=4 ve n=30 iken 0,2617 , q=10 ve n=30 iken 0,1159 olarak elde edilmesi bu durum için bir örnektir. ρ değeri 0,99 olduğunda LT1 testinin güç değeri diğer testlere göre daha yüksektir.

Çizelge 5.13. ρ=0.85 ve σ=1 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0923	0,0935	0,1003	0,1040	0,0687	0,0794	0,0858	0,0845	0,0561	0,0639	0,0665	0,0694
HK	0,1128	0,1193	0,1291	0,1380	0,0751	0,0936	0,1018	0,1010	0,0551	0,0697	0,0733	0,0771
HKB	0,1360	0,1654	0,1822	0,1943	0,0760	0,1117	0,1246	0,1285	0,0393	0,0685	0,0793	0,0771
L	0,1405	0,1800	0,2006	*	0,0787	0,1332	0,1539	*	0,0528	0,0782	0,1046	0,1293
LT1	*	*	*	*	0,0581	0,0672	0,0736	0,0720	*	0,0609	0,0670	0,0709
LT2	*	*	*	*	*	*	*	*	*	0,1123	*	*

Çizelge 5.14. ρ=0.95 ve σ=1 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0949	0,1011	0,1041	0,1044	0,0691	0,0824	0,0834	0,0848	0,0547	0,0661	0,0693	0,0697
HK	0,1184	0,1300	0,1408	0,1406	0,0772	0,0969	0,0990	0,1009	0,0541	0,0725	0,0750	0,0776
HKB	0,1488	0,1944	0,2182	0,2228	0,0759	0,1116	0,1179	0,1233	0,0375	0,0681	0,0777	0,0826
L	0,1372	0,1758	0,1955	0,2044	0,0796	0,1300	0,1471	0,1624	0,0529	0,0776	0,1024	0,1209
LT1	0,1019	0,1077	0,1140	0,1149	0,0585	0,0935	0,0983	0,1010	*	0,0591	0,0744	0,0822
LT2	*	*	*	*	0,1273	*	*	*	*	0,0977	0,1311	*

Çizelge 5.15. ρ=0.99 ve σ=1 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0895	0,1008	0,1077	0,1052	0,0755	0,0777	0,0855	0,0870	0,0562	0,0649	0,0681	0,0698
HK	0,1117	0,1310	0,1384	0,1390	0,0830	0,0909	0,0999	0,1017	0,0565	0,0718	0,0751	0,0771
HKB	0,1457	0,1946	0,2106	0,2183	0,0789	0,0975	0,1111	0,1162	0,0382	0,0655	0,0736	0,0786
L	0,1057	0,1339	0,1479	0,1502	0,0820	0,0957	0,1125	0,1213	0,0555	0,0699	0,0804	0,0911
LT1	*	0,2187	0,2333	0,2387	*	0,1430	0,1679	0,1804	*	*	0,1079	0,1309
LT2	*	*	*	*	*	*	*	*	*	*	*	*

σ değerinin 1 olduğu farklı ρ değerleri için Çizelge 5.13 , Çizelge 5.14 ve Çizelge 5.15 incelendiğinde; genel olarak değişken sayısına bakılmaksızın, örneklem büyüklüğü arttıkça testlerin güç değerleri de artmıştır. Değişken sayısı dikkate alındığında ise değişken sayısının artmasıyla testlerin güç değerlerinin azaldığı gözlemlenmiştir. ρ değerinin 0,85 olduğu durumlar incelendiğinde, L testinin güç değerinin diğer testlere göre daha yüksek olduğu görülmektedir. ρ değerinin 0,95 olduğu ve değişken sayısının büyük olduğu durumlarda L testinin güç değeri diğer testlerden daha yüksek çıkmıştır. ρ değeri 0,99 olduğunda ise LT1 testinin güç değeri diğer testlerden daha yüksektir.

Çizelge 5.16. ρ=0.85 ve σ=2 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0900	0,0945	0,1012	0,1035	0,0707	0,0795	0,0832	0,0863	0,0551	0,0643	0,0664	0,0699
HK	0,1042	0,1140	0,1269	0,1287	0,0777	0,0918	0,0960	0,0997	0,0546	0,0689	0,0728	0,0770
HKB	0,1154	0,1390	0,1571	0,1621	0,0743	0,1000	0,1075	0,1137	0,0387	0,0646	0,0736	0,0802
L	0,1145	0,1476	0,1651	0,1743	0,0732	0,1058	0,1191	0,1321	0,0517	0,0659	0,0800	0,0918
LT1	0,0390	*	*	*	0,0611	0,0654	0,0720	0,0745	*	0,0591	0,0665	0,0703
LT2	*	*	*	*	*	*	*	*	*	0,1057	*	*

Çizelge 5.17. ρ=0.95 ve σ=2 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0958	0,0986	0,1039	0,1075	0,0725	0,0810	0,0825	0,0853	0,0557	0,0640	0,0686	0,0697
HK	0,1136	0,1234	0,1331	0,1361	0,0804	0,0926	0,0957	0,0999	0,0544	0,0689	0,0748	0,0751
HKB	0,1325	0,1625	0,1796	0,1870	0,0766	0,0974	0,1057	0,1121	0,0378	0,0635	0,0742	0,0772
L	0,1232	0,1519	0,1694	0,1815	0,0759	0,1072	0,1205	0,1349	0,0535	0,0662	0,0828	0,0954
LT1	0,0972	0,1061	0,1136	0,1152	0,0585	0,0869	0,0929	0,0980	*	0,0521	0,0699	0,0760
LT2	*	*	*	*	0,1086	*	*	*	*	0,0817	0,1099	*

Çizelge 5.18. ρ=0.99 ve σ=2 iken testlerin güç değerleri

Test	q=4				q=6				q=10			
	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100	n=15	n=30	n=50	n=100
EKK	0,0919	0,1030	0,1039	0,1065	0,0725	0,0835	0,0869	0,0877	0,0566	0,0666	0,0702	0,0708
HK	0,1107	0,1321	0,1329	0,1383	0,0807	0,0947	0,1003	0,1012	0,0569	0,0717	0,0766	0,0771
HKB	0,1340	0,1792	0,1882	0,2016	0,0763	0,0985	0,1078	0,1103	0,0401	0,0654	0,0744	0,0768
L	0,1071	0,1322	0,1382	0,1451	0,0760	0,0992	0,1103	0,1168	0,0556	0,0693	0,0781	0,0855
LT1	*	0,2004	0,2123	0,2224	*	0,1270	0,1486	0,1604	*	*	0,0916	0,1079
LT2	*	*	*	*	*	*	*	*	*	*	*	*

σ değerinin 2 olduğu farklı ρ değerleri için Çizelge 5.16 , Çizelge 5.14 ve Çizelge 5.18 incelendiğinde; σ'nın diğer değerleriyle benzer bulgular elde edildiği söylenebilir.

Genel olarak testlerin güç değerlerine bakıldığında, değişken sayısının artması durumunda testlerin güç değerlerinin azaldığı söylenebilir. Değişken sayısının göz ardı edildiği tüm durumlar için, örneklem büyüklüğünün artmasıyla testlerin güç değerlerinin de arttığı görülmektedir. L testinin güç değeri, σ'nın büyük olduğu ve ρ değerinin 0,85 olduğu durumlarda diğer testlerden daha yüksek çıkmıştır. ρ değeri 0,95'e yükseldiğinde ve değişken sayısının küçük olduğu durumlarda, HKB testinin güç değeri daha yüksek, değişken sayısı arttığında ise L testinin güç değeri daha yüksek olmaktadır. ρ değeri 0,99'a yükseldiğinde ise en yüksek güç değeri elde edilen test LT1 olmaktadır.

Kısaca değerlendirildiğinde, ρ ve σ değerlerinin düşük seviyelerinde HKB testinin güç değerinin yüksek sonuçlar verdiği, ρ değerinin orta seviyelerinde L testinin güç değerinin yüksek olduğu ve ρ değerinin yüksek seviyelerinde LT1 testinin güç değerinin daha yüksek sonuçlar verdiği söylenebilir.

6. SONUÇ

Literatürde yer alan çalışmalar incelendiğinde, Ridge ve Liu tahmin edicileri için model katsayılarının anlamlılık testine ilişkin sınırlı sayıda çalışma olduğu görülmektedir. Bu çalışmada, regresyon modelinin katsayılarının anlamlılık testi için Liu Tipi tahmin edicisine ait bir test istatistiği elde edilmiştir. Çoklu bağlantı durumunda, tahmin edicilerden elde edilen testlerin hangisinin daha iyi performans gösterdiğini belirlemek için simülasyon çalışması yapılmıştır. Bu sebeple, Ridge, Liu ve Liu Tipi tahmin edicilerinden elde edilen testlerin farkı durumlardaki deneysel 1.Tip hataları ve güç değerleri hesaplanmıştır. Elde edilen simülasyon sonuçlarına göre, çoklu bağlantının orta seviyesinde ve σ'nın küçük değerlerinde HKB testinin diğer testlere göre daha güçlü olduğu görülmektedir. Çoklu bağlantının orta seviyesinde ve σ'nın yüksek değerlerinde ise L testinin diğer testlere göre daha güçlü olduğu gözlemlenmiştir. Çoklu bağlantının şiddetli seviyelerinde LT1 testinin diğer testlere göre daha güçlü olduğu tespit edilmiştir.

KAYNAKLAR

Akdeniz, F. and Kaçıranlar, S. (1995). On the almost unbiased generalized liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics-Theory and Methods*, 24(7), 1789-1797.

Akdeniz, F. and Kaçıranlar, S. (2001). More on the New Biased Estimator in Linear Regression, *The Indian Journal of Statistics,* 63(3), 321-325.

Bradley, J. V. (1978). Robustness?. *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.

Draper, R. N., and Smith, H. (1998). *Applied regression analysis*. US: Wiley Series in Probability & Statistics, 388.

Ebegil, M., Gökpınar, F. and Ekni, M. (2006). A simulation study on some shrinkage estimators. *Hacettepe Journal of Mathematics and Statistics,* 35, 213-226.

Ebegil, M. and Gökpınar, F. (2012) A test statistic to choose between liu-type and least squares estimator based on mean square error criteria. *Journal of Applied Statistics*, 39(10), 2081-2096.

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.

Gökpınar, E. and Ebegil, M. (2016). A study on tests of hypothesis based on ridge estimator. *Gazi University Journal of Science*, 29(4), 769-781.

Gruber, M. H. J. (2012). Liu and ridge estimators - a comparison. *Communications in Statistics - Theory and Methods* , 41(20), 3739-3749.

Gujarati, N. D. (1999). *Temel ekonometri*. İstanbul: Literatür Kitapevi.

Halawa, A. M. and Bassiouni, M.Y. (2000). Tests of regression coefficients under ridge regression models. *Journal of Statistical Simulation and Com-putation*, 65, 341-356.

Hald, A. (1952). *Statistical theory with engineering applications*. New York: Wiley.

Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). Ridge regression: Some simulations. *Communications in Statistics-Theory and Methods,* 4(2), 105-123.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.

Kaçıranlar, S., ve Sakallıoğlu, S. (2000). *Liu ve temel bileşenler tahmin edicilerinin birleştirilmesi*. İstatistik Araştırma Sempozyumu (s. 1). Ankara, Turkey.

Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Simulation and Computation,* 32(2), 419-435.

Liski, E. P. (1982). A test of the mean square error criterion for shrinkage estimators. *Communications in Statistics*, 11(5), 543-562.

Liski, E. P. (1983). Choosing a shrinkage estimator a test of the mean square error criterion. *Proc. First Tampere Sem. Linear Models*, 245-262.

Liu, K. (1993). A new class of blased estimate in linear regression. *Communications in Statistics-Theory and Methods*, 22(2), 393-402.

Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics*, 32(5), 1009-1020.

Liu, K. (2004). More on Liu-type estimator in linear regression. *Communications in Statistics-Theory and Methods*, 33(11), 2723-2733.

Mallows, C. L. (1973). Some Comments on C_p. *Technometrics*, 15(4), 661-675.

Marquardt, D.W. (1970). Generalized ınverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics*, 12, 591-612.

McDonald, G. and Galarneau, D. (1975). A monte carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, 70(350), 407-416.

Melo, S. P. and Kibria, B. M. G. (2021). Testing the regression coefficients in the liu linear regression model. *Simulation and Application*, 2-8.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2001). *Introduction to linear regression analysis* 3rd ed., New York: John Wiley & Sons, Inc.

Özkale, M. R. and Kaçıranlar, S. (2007). Superiority of the r-d class estimator over some estimators by the mean square error matrix criterion. *Statistics and Probability Letters*, 77, 438–446.

Sakallıoğlu, S. and Kaçıranlar, S. (2008). A new biased estimator based on ridge estimation. *Stat Papers*, 49, 669-689.

Stein, C. (1956). *Inadmissibility of usual estimator for the mean of a multivariate normal distribution*. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Berkeley University of California Press, 197-206.

Wilcox, R. R. (2019). Multicolinearity and ridge regression: Results on type I errors, power and heteroscedasticity. *Journal of Applied Statistics*, 46(5), 946-957.



Gazili olmak ayrıcalıktır