# RIDGE TAHMİNİNE DAYALI YANLI TAHMİN EDİCİ İÇİN BİR TEST İSTATİSTİĞİ

Meral EBEGİL
Gazi Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, 06500, ANKARA
mdemirel@gazi.edu.tr

***ÖZET***

*Bu çalışmada, Ridge tahmin edicisine dayalı yanlı tahmin edici ile en küçük kareler tahmin edicisi arasında bir seçim yapmak için ortalama hata kriterine bağlı bir test istatistiği önerilmiştir.*

***Anahtar Kelimeler:*** *Ortalama Hata Kare, Ridge Regresyon, Liu-Tipi Tahmin Edici, Doğrusal Kabul Edilebilir Tahmin Ediciler, Shrinkage Tahmin Edicileri*

# A TEST STATISTIC FOR BIASED ESTIMATOR BASED ON RIDGE ESTIMATOR

***ABSTRACT***

*In this study, a test statistic of the mean square error criterion has been proposed to choose between a biased estimator based on Ridge estimator and the least squares estimator.*

***Key Words:*** *Mean Square Error, Ridge Regression, Liu-Type Estimator, Linear Admissible Estimators, Shrinkage Estimators.*

## 1. GİRİŞ

Regresyon katsayılarını tahmin etmek için yaygın olarak kullanılan yöntem en küçük kareler (EKK) yöntemidir. Ancak, EKK yönteminin doğru sonuçlar vermesi için çeşitli varsayımların sağlanması gerekmektedir. Bunlardan biri bağımsız değişkenler arasında ilişki olmamasıdır. Ama gerçekte bu durum her zaman sağlanmayabilir. Böyle durumlarda, EKK tahmin yönteminin kullanılması yanlış model bulgularına ve kullanımına neden olur. Bu tür birbiriyle bağımlılık gösteren bağımsız değişkenlerle analiz yapmak için çeşitli yöntemler geliştirilmiştir. Bu yöntemlerden biri yanlı tahmin yöntemleridir. Yanlı yöntemlere ilişkin tahmin ediciler, EKK tahmin edicilerine göre yanlı, ancak çok daha küçük varyanslı tahminler verirler. Yanlı tahmin yöntemlerinde genel amaç, EKK tahmin yönteminde büyük olan varyans alanını küçük bir yan karşılığında daraltmaktır. Böylece EKK yöntemine göre daha doğru sonuçlar elde edilir.

Yanlı tahmin ediciler sınıfının içinde yer alan bazı tahmin ediciler, Shrinkage tahmin edicileri olarak adlandırılır. Temel Bileşenler regresyonu, Ridge regresyonu ve bunların türevleri bu sınıfın birer üyesidirler. Farebrother (1978) yaptığı çalışmasında, Shrinkage tahmin edicileri için genel bir yapı oluşturmuştur. Ridge, Temel Bileşenler ve Koşullu-minimum hata kare ortalamalı yanlı tahmin edicilerinin birer Shrinkage tahmin edicisi olduklarını göstermiştir [1]. Liski (1982) çalışmasında, EKK tahmin edicisi ile Shrinkage tahmin edicisi arasında seçim yapmak için güçlü Ortalama Hata Kare (OHK) ölçütünü vermiştir [2]. Yine Liski (1983) çalışmasında EKK tahmin edicisi ile Shrinkage tahmin edicisi arasında seçim yapmak için daha zayıf OHK test işlemini kullanmıştır [3]. Kejian (1993) çalışmasında ridge tahmin edicisine alternatif olarak Liu-Kejian tahmin edicisini önermiştir [4]. Daha sonra bu tahmin edici Akdeniz ve Kaçıranlar (1995) tarafından "Liu tahmin edicisi" olarak adlandırılmıştır [5]. Sakallıoğlu, Kaçıranlar ve Akdeniz (1997) tarafından Liu tahmin edicisinin iterasyon tahmin edicileri ile karşılaştırılması incelenmiştir [6]. Demirel (1999) çalışmasında, Shrinkage tahmin edicilerinin genel yapısını vermiştir [7]. Ebegil, Gökpınar ve Ekni (2005) çalışmalarında, yanlı tahmin ediciler sınıfı içinde yer alan Shrinkage tahmin edicilerinden, Ridge ve Liu tahmin edicileri için bir test istatistiği vermiş, ayrıca bu tahmin ediciler bağımsız değişkenler arasındaki farklı korelasyon yapılarına göre simülasyon yoluyla karşılaştırılmıştır [8]. Sakallıoğlu ve Kaçıranlar (2003) tarafından Ridge tahmin edicisine dayalı yanlı bir tahmin edici önerilmiştir [9]. Ebegil (2006) çalışmasında Liu-Tipi ve Ridge tahmin edicisine dayalı yanlı tahmin edicileri de Shrinkage tahmin ediciler yapısı içinde ifade ederek bu tahmin edicilerin birer Shrinkage tahmin edicisi olduğunu göstermiştir [10]. Ebegil (2007) tarafından, Ridge tahmin edicisine dayalı yanlı tahmin edicinin en az, EKK tahmin edicisi kadar etkin olabilmesi için gereklilik ve yeterlilik koşullarını elde etmiştir [11].

Bu çalışmanın ikinci bölümünde Shrinkage tahmin edicilerinin temel yapısı kısaca verildikten sonra, Ridge tahmin edicisine dayalı yanlı tahmin edicinin en az, EKK tahmin edicisi kadar etkin olabilmesi için gereklilik ve yeterlilik koşulları üzerinde durulmuştur. Üçüncü bölümde, öncelikle bir Shrinkage tahmin edicisi ile EKK tahmin edicisinden birini seçmek için bir test istatistiği tanıtılmıştır. Daha sonra bu test istatistiği ve Ebegil (2007)'de oluşturulan gereklilik ve yeterlilik koşuluna bağlı olarak Ridge tahmin edicisine dayalı yanlı tahmin edicisi ile EKK tahmin edicisinden birini seçmek için bir test istatistiği önerilmiştir.

## 2. SHRINKAGE TAHMİN EDİCİLER SINIFININ GENEL YAPISI

Bu bölümde doğrusal regresyon modelinden yola çıkarak Shrinkage tahmin edicilerinin genel yapısı verilmiştir.

n gözlemli, k bağımsız değişkenli bir çoklu doğrusal regresyon modeli,

$$Y = X\beta + \varepsilon \ , \qquad \varepsilon \sim \left(0 \ , \ \sigma^2 I_n\right) \qquad (1)$$
$$rank\left(X_{n\times q}\right) = q \leq n$$

biçiminde tanımlanır [1]. $Y$, $(n\times 1)$ boyutlu bağımlı değişken vektörü; $q = k+1$ olmak üzere, $X$, $(n\times q)$ boyutlu stokastik olmayan girdi matrisi; $\beta$, $(q\times 1)$ boyutlu bilinmeyen katsayılar vektörü; $\varepsilon$, $E(\varepsilon)=0$ ve $E(\varepsilon\varepsilon') = \sigma^2 I_n$ koşullarını sağlayan hata vektörüdür.

$(X^{'}X)$ matrisi; $\left(q\times q\right)$ boyutlu, pozitif tanımlı, simetrik bir matris olduğundan dolayı öyle bir ortonormal $P$ matrisi vardır ki $(X^{'}X)$ matrisini $P^{'}X^{'}XP = \Lambda$ biçiminde köşegenleştirir. $\Lambda$, elemanları $(X^{'}X)$ matrisinin pozitif öz değerleri $\lambda_1 > \lambda_2 > \ldots > \lambda_q$ olan $\left(q\times q\right)$ boyutlu köşegen bir matristir [2,3].

$Z = XP$ ve $\alpha = P^{'}\beta$ olmak üzere, (1) eşitliğindeki modelin kanonik formu,

$$Y = XPP^{'}\beta + \varepsilon = Z\alpha + \varepsilon \tag{2}$$

şeklindedir.

Bir regresyon parametresinin doğrusal kabul edilebilir tahmin ediciler sınıfı kapsamı içindeki $b \in R^{q}$ olmak üzere kabul edilebilir doğrusal bir tahmin edici,

$$\tilde{\beta} = A(\hat{\beta} - b) + b \tag{3}$$

biçimindedir [12]. Burada $\hat{\beta}$; $\beta$'nın EKK tahmin edicisi, $A$; $(q \times q)$ boyutlu bir matris ve $b$; $(q \times 1)$ boyutlu sabit bir vektördür. Bu şekilde gösterilen tahmin ediciler, doğrusal kabul edilebilir tahmin ediciler (linear admissible estimators) sınıfındandır. Bir tahmin edicinin kabul edilebilir (admissible) bir tahmin edici olması için diğer koşullar aşağıdaki gibi sıralanabilir.

$$(X^{'}X)A \text{ veya } A(X^{'}X)^{-1} \tag{4}$$

simetrik ve $A$ matrisinin öz değerleri [0,1] aralığında olmalıdır.

(3) ve (4) eşitliklerindeki koşullara ek olarak, kabul edilebilirlik için $A$ matrisinin simetrik olduğu varsayılır. $A$ ve $\left(X'X\right)$ matrisleri aynı ortonormal $P$ matrisi tarafından köşegenleştirilebilir. $P'AP = \Delta$, elemanları $A$ matrisinin [0,1] aralığına düşen öz değerleri $\delta_1, \delta_2, .....,\delta_q$ olan $(q \times q)$ boyutlu köşegen bir matristir [2,3].

(2) eşitliğindeki model altında (3) eşitliği,

$$\tilde{\alpha} = P'A(\hat{\beta} - b) + P'b$$

$$\begin{aligned} &= P^{'}AP(\hat{\alpha}-a)+a \\ &= \Delta(\hat{\alpha}-a)+a \end{aligned} \tag{5}$$

olur. Burada $\hat{\alpha}=P^{'}\hat{\beta}$, $a=P^{'}b$'dır. Bu tür doğrusal kabul edilebilir tahmin ediciler, Shrinkage tahmin ediciler olarak isimlendirilir [2,3].

## 2.1. Shrinkage ve En Küçük Kareler Tahmin Edicilerinin Ortalama Hata Kare Matrisleri

EKK tahmin edicileri $\hat{\beta}$ ve Shrinkage tahmin edicileri $\tilde{\beta}$'nın OHK matrisleri sırasıyla,

$$OHK(\hat{\beta})=\sigma^2(X^{'}X)^{-1}$$

ve

$$OHK(\tilde{\beta})=\sigma^2A(X^{'}X)^{-1}A^{'}+(I-A)(\beta-b)(\beta-b)^{'}(I-A) \tag{6}$$

dır [2,3]. (6) eşitliği ile verilen bu matrislerin eşdeğer kanonik formları,

$$OHK(\hat{\alpha})=\sigma^2\Lambda^{-1}$$

ve

$$OHK(\tilde{\alpha})=\sigma^2\Delta\Lambda^{-1}\Delta+(I-\Delta)(\alpha-a)(\alpha-a)^{'}(I-\Delta) \tag{7}$$

biçiminde gösterilir. Bu iki matris arasındaki farkın, yani $OHK(\hat{\beta})-OHK(\tilde{\beta})$'nın negatif olmayan tanımlı bir matris olması, $OHK(\hat{\beta})-OHK(\tilde{\beta})\geq 0$ ile mümkündür. Buradan hareketle $OHK(\hat{\beta})-OHK(\tilde{\beta})$ farkının negatif olmayan tanımlı olmasının gerek ve yeter koşulu

$$(\beta-b)^{'}(I+A)^{-1}X^{'}X(I-A)(\beta-b)/\sigma^2\leq 1 \tag{8}$$

eşitliğinin sağlanmasıdır. Aynı şekilde, $OHK(\hat{\alpha}) - OHK(\tilde{\alpha})$ farkının negatif olmayan tanımlı bir matris olması ise ancak ve ancak

$$(\alpha - a)^{'}(I+\Delta)^{-1}\Lambda(I-\Delta)(\alpha - a)/\sigma^2 \leq 1 \tag{9}$$

ile mümkündür [2]. Eş (8)'deki koşul altında bir Shrinkage tahmin edicisi olan $\tilde{\beta}$'nın, en az $\hat{\beta}$ EKK tahmin edicisi kadar etkin olduğu söylenebilir [2].

## 2.2. Shrinkage Tahmin Edicileri Yardımıyla Ortalama Hata Karelerin Azaltılması

Liski (1982), çalışmasında genel yapısını verdiği Shrinkage tahmin edicilerinin OHK matrisleri ile, bilinen EKK tahmin edicilerinin OHK matrislerini karşılaştırarak bir gereklilik ve yeterlilik koşulu elde etmiştir. Daha sonra bu gereklilik ve yeterlilik koşulunu kullanarak Ridge ve $\hat{\beta}_d = (d\mathrm{I})\hat{\beta}$, $0 \leq d \leq 1$, tahmin edicilerinin her biri için gereklilik ve yeterlilik koşullarını oluşturmuştur [2]. Bu koşullar sağlandığında Shrinkage tahmin edicileri en az, EKK tahmin edicileri kadar etkindir. Ekni (1999) çalışmasında Liu tahmin edicisi için gereklilik ve yeterlilik koşulunu elde etmiştir [13].

Sakallıoğlu ve Kaçıranlar (2003), Ridge tahminine dayalı bir tahmin ediciyi, k>0 ve $-\infty$<d<$+\infty$ olmak üzere,

$$\tilde{\beta}(k,d) = (X'X + \mathrm{I})^{-1}\left(X'Y + d\tilde{\beta}(k)\right) \tag{10}$$

şeklinde tanımlamışlardır. Burada $\tilde{\beta}(k) = (X'X + k\mathrm{I})^{-1}X'Y$, Ridge regresyon tahmin edicisidir [9]. Ebegil (2007) çalışmasında, Eş (10)'da verilen tahmin edici için gereklilik ve yeterlilik koşullarını,

$$(\alpha - a)'\left(\frac{(\Lambda + k\mathrm{I}) - d\Lambda}{(\mathrm{I} + k\Lambda^{-1})(\mathrm{I} + 2\Lambda) + d}\right)(\alpha - a)/\sigma^2 \leq 1 \tag{11}$$

veya

$$(\beta - b)'\left(\frac{\left(X'X + k\mathrm{I}\right) - dX'X}{\left(\mathrm{I} + k\left(X'X\right)^{-1}\right)\left(\mathrm{I} + 2\left(X'X\right)\right) + d}\right)(\beta - b)/\sigma^2 \leq 1 \qquad (12)$$

biçiminde elde etmiştir [11]. (11) ve (12) eşitliklerindeki gereklilik ve yeterlilik koşulunun sağlanması durumunda Ridge tahminine dayalı yanlı tahmin edici, en az EKK tahmin edicisi kadar etkindir sonucu çıkarılabilir.

## 3. SHRINKAGE TAHMİN EDİCİSİNİN SEÇİMİ İÇİN BİR TEST

Liski (1982) çalışmasında, OHK matrislerine dayalı olarak, bir Shrinkage tahmin edicisi ile EKK tahmin edicisinden birini seçmek için bir test istatistiği oluşturmuştur [2]. Bu çalışmada, bu test istatistiği ve Ebegil (2007) tarafından oluşturulan gereklilik ve yeterlilik koşullarına bağlı olarak Ridge tahminine dayalı yanlı tahmin edici ile EKK tahmin edicisinden birini seçmek için bir test istatistiği önerilmiştir.

(8) ve (9) eşitliklerinde verilen gereklilik ve yeterlilik koşulları altında iki tahmin edici arasından seçim yapılması mümkündür. Bu durumda Shrinkage tahmin edicisi $\tilde{\beta}$ ile EKK tahmin edicisi $\hat{\beta}$ arasında bir seçim yapılması için kullanılacak test istatistiğinin yapısı, bu eşitsizliğe dayandırılır. Liski (1982) Toro-Vizcarrondo ve Wallace'nin test istatistiğini kullanarak aşağıdaki test istatistiği üzerinde çalışmıştır [2]:

$$\tilde{F} = \hat{\beta}'H\hat{\beta}\big/ m\hat{\sigma}^2 \ , \qquad (13)$$

burada

$$H = (I + A)^{-1} X^{'} X (I - A),$$
$$\hat{\sigma}^2 = (Y - X\hat{\beta})^{'}(Y - X\hat{\beta})\big/(n - q),$$
$$rank(H) = m \ \text{ve} \ b = 0$$

olarak alınmıştır. Bu istatistiğin kanonik formu,

$$\begin{aligned}\tilde{F} &= \hat{\alpha}'(I+\Delta)^{-1}\Lambda(I-\Delta)\hat{\alpha}/m\hat{\sigma}^2 \\ &= \frac{1}{m}\sum_{i=1}^{q}\gamma_i(\lambda_i\hat{\alpha}_i^2/\hat{\sigma}^2) \qquad (14)\end{aligned}$$

şeklinde yazılabilir. Burada $m = rank(I+\Delta)^{-1}\Lambda(I-\Delta)$ dır.
$m$, sıfırdan farklı $\gamma_i$'lerin sayısı olmak üzere ,

$$\tilde{F} = \frac{1}{m}\sum_{i=1}^{m}\gamma_i(\lambda_i\hat{\alpha}_i^2/\hat{\sigma}^2) \qquad (15)$$

şeklinde de ifade edilebilir. $m$'nin $1 \le m \le q$ olduğu görülmektedir.

$$F_i = \lambda_i\hat{\alpha}_i^2/\hat{\sigma}^2$$

olarak ifade edildiğinde,

$$\tilde{F} = \frac{1}{m}\sum_{i=1}^{m}\gamma_i F_i \qquad (16)$$

olur. Bu durumda $\tilde{F}$ test istatistiği $F_i$ istatistiklerinin karışımından oluşur. Diğer bir ifadeyle, (15) eşitliği, (16) eşitliği gibi de yazılabilir. Burada $\lambda_i\hat{\alpha}_i^2/\hat{\sigma}^2$, serbestlik derecesi 1 ve merkezi olmama parametresi $w_i = \lambda_i\alpha_i^2/\sigma^2$ olan bir ki-kare dağılımına uyar. $(n-q)\hat{\sigma}^2/\sigma^2$ ise serbestlik derecesi $(n-q)$ olan ki-kare dağılımına sahiptir. $\frac{\hat{\sigma}^2}{\sigma^2}$ ve $\frac{\lambda_i\hat{\alpha}_i^2}{\sigma^2}$ bağımsızdır. Bu durumda, $F_i = \frac{\lambda_i\hat{\alpha}_i^2}{\hat{\sigma}^2}$, serbestlik derecesi 1 ve $(n-q)$, merkezi olmama parametresi $w_i = \lambda_i\alpha_i^2/\sigma^2$ olan $F$ dağılımına uymaktadır.

(16) eşitliği, $\gamma_i$'lerin bütün ağırlıkları 1'ler veya sıfırlar olmadıkça, merkezi olmayan $F_i$'nin dağılım fonksiyonun kapalı formunun ifade edilmesi oldukça zordur. Bununla birlikte yaklaşık çözümlere ulaşılabilir [2,3,14].

Shrinkage tahmin edicileri için gereklilik ve yeterlilik koşulu, $\sum_{i=1}^{m} \gamma_i \omega_i \leq 1$ idi. Bundan dolayı hipotez testleri,

$$H_0: \sum_{i=1}^{m} \gamma_i \omega_i \leq 1 \text{ karşıt hipotez } H_1: \sum_{i=1}^{m} \gamma_i \omega_i > 1 \quad (17)$$

şeklinde yazılabilir. Önerilen test işleminde karar kuralı,

$$\tilde{F} \leq \tilde{F}_\alpha(m, n-q, 1) \text{ ise } H_0 \text{ kabul,}$$

ve

$$\tilde{F} \rangle \tilde{F}_\alpha(m, n-q, 1) \text{ ise } H_0 \text{ red}$$

şeklindedir. $\tilde{F}_\alpha(m, n-q, 1)$, serbestlik derecesi $m$ ve $(n\text{-}q)$, merkezi olmayan parametresi $w = \sum_{i=1}^{m} \gamma_i \omega_i = 1$ olan $\tilde{F}$'nın dağılımından elde edilir. $\tilde{F}$ için merkezi-F yaklaşımı kullanarak $\tilde{F}_\alpha(m, n-q, 1)$ kritik noktaların değeri belirlenebilir. Bunun için momentler yaklaşımı ile $F_i$ istatistiklerinin momentleri kullanılarak $\tilde{F}$ test istatistiğinin başlangıç momentleri elde edilir. Bunu için Theobald (1974)'e bakılabilir [15].

### 3.1. $\tilde{F}$ İstatistiği İçin Merkezi-F Yaklaşımı

Patnaik (1949), merkezi olmayan F dağılımı için merkezi-F yaklaşımı üzerinde çalışmıştır [16]. Bu yaklaşımda merkezi-F dağılımı

$F(\vartheta, n-q)$'nın ve merkezi olmayan $\tilde{F}(m, n-q, \omega)$ dağılımının ilk iki momentleri kullanılarak

$$\tilde{F}(m, n-q, \omega) \approx r\ F(\vartheta, n-q)$$

yazılabilir. $r$ ve $\vartheta$ parametreleri F dağılımlarının ilk iki momentleri ile bulunur. Diğer bir ifadeyle, merkezi-F'in iki moment yaklaşımı, merkezi-F ve $\tilde{F}/r$'nin ilk iki momentleri için ifadelerin eşitlenmesiyle,

$$r = \frac{1}{m}\sum_{i=1}^{m}\gamma_i(1+\omega_i) \quad \text{ve} \quad \vartheta = \frac{\left[\sum_{i=1}^{m}\gamma_i(1+\omega_i)\right]^2}{\sum_{i=1}^{m}\gamma_i^2(1+2\omega_i)} \tag{18}$$

biçiminde elde edilir. Burada $\gamma = \sum_{i=1}^{q}\gamma_i$ ve $\sum_{i=1}^{m}\gamma_i\omega_i = 1$ verilmişken $r = \frac{\gamma+1}{m}$, şeklinde de ifade edilebilir ve $\gamma_i$'ler [0,1] aralığında olmak üzere düzeltilmiş serbestlik derecesi, $\vartheta = (\gamma+1)^2/(\sum_{i=1}^{m}\gamma_i^2 + 2\sum_{i=1}^{m}\gamma_i^2\omega_i)$ biçiminde yazılabilir [2]. Buradan kolaylıkla

$$\gamma_{\min} \leq \sum_{i=1}^{m}\gamma_i^2\omega_i \leq \gamma_{\max} \tag{19}$$

olduğu görülür [17]. Bu eşitsizlik yardımıyla $\vartheta$ düzeltilmiş serbestlik derecesi için alt ve üst sınırları,

$$(\gamma+1)^2/(\sum_{i=1}^{m}\gamma_i^2 + 2\gamma_{\max}) \leq \vartheta \leq (\gamma+1)^2/(\sum_{i=1}^{m}\gamma_i^2 + 2\gamma_{\min}) \tag{20}$$

şeklinde belirlenir [2]. Üst sınır $\vartheta_{\max}$ ve alt sınır $\vartheta_{\min}$ olarak ifade edilebilir. Bu halde bütün $0 < \alpha < 1$ için

$F_\alpha(\vartheta_{\max}, n-q) \le F_\alpha(\vartheta_{\min}, n-q)$ iken $F_\alpha(\vartheta_{\max}, n-q)$ ve $F_\alpha(\vartheta_{\min}, n-q)$ kritik noktaları elde edilir. $\tilde{F}/r$ istatistiği, bu kritik değerlerle karşılaştırılabilir. Buradan hareketle, test istatistiği için bu kritik noktalardan,

$\tilde{F}/r > F_\alpha(\vartheta_{\min}, n-q)$ ise $H_o$ hipotezi red ,

$\tilde{F}/r < F_\alpha(\vartheta_{\max}, n-q)$ ise $H_o$ hipotezi kabul ve

$F_\alpha(\vartheta_{\max}, n-q) \le \tilde{F}/r \le F_\alpha(\vartheta_{\min}, n-q)$ kararsızlık (belirsizlik) durumu

şeklinde bölgeler oluşturulabilir.

### 3.2. Ridge Tahminine Dayalı Yanlı Tahmin Edicinin Seçimi İçin Bir Test

Bu bölümde, daha önceki bölümlerde açıklanan test istatistiğinin Ridge tahminine dayalı yanlı tahmin ediciye nasıl uygulanacağı açıklanmıştır.

(13) eşitliğinde verilen test istatistiğinde gerekli ifadeler yerine konularak, Ridge tahminine dayalı yanlı tahmin edici için test istatistiğini oluşturmak mümkün olacaktır. Bunun için, (10) eşitliğinde verilen Ridge tahminine dayalı bir tahmin edicisinde, Ridge regresyon tahmin edicisi $\tilde{\beta}(k) = (X'X + kI)^{-1} X'Y$ yerine yazılır ve gerekli düzenlemeler yapılırsa,

$$\begin{aligned}\tilde{\beta}_{k,d} &= (X'X + I)^{-1}\left(I + d(X'X + kI)^{-1}\right)X'X\hat{\beta} \\ &= A\hat{\beta} \qquad (21)\end{aligned}$$

şeklinde ifade edilebilir. Buradan test istatistiğinde yer alan, $H = (I + A)^{-1} X' X (I - A)$ ifadesindeki $A$ matrisi yerine yazılarak, Ridge tahminine dayalı yanlı tahmin edici için test istatistiği,

$$\tilde{F} = \hat{\beta}' H \hat{\beta} \big/ m\hat{\sigma}^2$$

$$=\frac{\hat{\beta}'\left[\dfrac{\left(X'X+k\mathrm{I}\right)-dX'X}{\left(\mathrm{I}+k\left(X'X\right)^{-1}\right)\left(\mathrm{I}+2\left(X'X\right)\right)+d}\right]\hat{\beta}}{m\hat{\sigma}^2} \tag{22}$$

şeklinde oluşturulur. Burada $m$ ve $\hat{\sigma}^2$ daha önceki bölümlerde tanımlandığı gibi hesaplanır. Bu test istatistiğinin kanonik formu,

$$\tilde{F}=\frac{\hat{\alpha}'\left(\dfrac{\left(\Lambda+k\mathrm{I}\right)-d\Lambda}{\left(\mathrm{I}+k\Lambda^{-1}\right)\left(\mathrm{I}+2\Lambda\right)+d}\right)\hat{\alpha}}{m\hat{\sigma}^2} \tag{23}$$

biçiminde belirlenir. Diğer kritik noktalar, karar kuralları ve tanımlamalar önceki bölümlerde açıklandığı gibi, buradaki ifadelere bağlı kalınarak hesap edilir.

## 4. SONUÇ

Model kurmak amacıyla, regresyon katsayılarını tahmin etmek için yaygın olarak kullanılan yöntem En Küçük Kareler (EKK) yöntemidir. EKK tahmin yönteminde amaç hatayı en küçüklemektir. Ancak bağımsız değişkenler arasında şiddetli bir ilişki bulunuyorsa, bu tür veriler hatada, dolayısıyla varyansta yanıltıcı bir büyümeye sebep olur. Bu büyüme parametre tahminlerine ve kestirim sonuçlarına olumsuz şekilde yansır. Diğer bir deyişle, bağımsız değişkenler arasında şiddetli bir ilişki bulunduğu durumlarda, varyanstaki yanıltıcı büyümeden dolayı, EKK tahmin yönteminin kullanılması yanlış model bulgularına ve kullanımına neden olur. Bu olumsuz etkiyi yok etmek için yanlı tahmin yöntemlerine başvurulur. Bu yöntemler küçük bir yan karşılığı varyans alanını dolayısıyla hatayı küçültür. Yanlı yöntemlere ilişkin bu tahmin ediciler, EKK tahmin edicilerine göre yanlı, ancak çok daha küçük varyanslı tahminler verirler. Diğer bir ifadeyle, yanlı tahmin yöntemlerinin kullanılmasındaki genel amaç, EKK tahmin yönteminde büyük olan varyans alanını küçük bir yan karşılığında daraltmaktır. Böylece EKK yöntemine göre daha doğru sonuçlar elde edilir.

Bu amaçla, Liski (1982) çalışmasında, OHK matrislerine dayalı olarak, bir Shrinkage tahmin edicisi ile EKK tahmin edicisinden birini seçmek için bir test istatistiği oluşturmuştur. Bu çalışmada, bu test istatistiği ve Ebegil (2007) tarafından, Ridge tahminine dayalı yanlı tahmin edicinin en az EKK tahmin edicisi kadar etkin olması için oluşturulan gereklilik ve yeterlilik koşulları kullanılarak, Ridge tahminine dayalı yanlı tahmin edici ile EKK tahmin edicisinden birini seçmek için bir test istatistiği oluşturulmuştur. Bu test istatistiği kullanılarak, Ridge tahminine dayalı yanlı tahmin edicinin, EKK tahmin edicisine tercih edileceği durumlar belirlenebilecektir.

**KAYNAKLAR**


[1]. Farebrother, R. W., (1978), A Class of Shrinkage Estimators., Journal of the Royal Statistical Society B, 40, 47-49.

[2]. Liski, E. P., (1982), A Test of the Mean Square Error Criterion for Shrinkage Estimators. Communications in Statistics 11(5), 543-562.

[3]. Liski, E. P., (1983), Choosing a Shrinkage Estimator-a test of the Mean Square Error Criterion. Proc. First Tampere Sem. Linear Models, 245-262.

[4]. Kejian, L., (1993). A New Class of Biased Estimate in Linear Regression., Communications. in Statistics: Theory and Methods.,22(2), 393-402.

[5]. Akdeniz, F. and Kaçıranlar, S. (1995). On the Almost Unbiased Generalized Liu Estimator and Unbised Estimation of the Bias and MSE. Commun. Statist-Theory and Meth., 24(7), 1789-1797.

[6]. Sakallıoğlu, S., Kaçıranlar, S. ve Akdeniz, F. (1997). Bazı Yanlı Regresyon Tahmin Edicilerinin Karşılaştırılması., Araştırma Sempozyumu'97 Bildirileri, DİE., Ankara.

[7]. Demirel, M., (1999), Bazı Shrinkage Tahmin Edicileri ile En Küçük Kareler Tahmin Edicisinin Bir Test İstatistiği ile Karşılaştırılması, Yüksek Lisans Tezi, Ankara, 11-18.

[8]. Ebegil, M.,Gökpınar, F. and Ekni, M. (2006), A Simulation Study on Some Shrinkage Estimators, Hacettepe Journal of Mathematics and Statistics, 35(2):213-226.

[9]. Sakallıoğlu, S., Kaçıranlar, S., (2005), Yeni Bir Yanlı Tahmin Edici ve Liu-Tipi Tahmin Edici ile Karşılaştırmalar. 4. İstatistik Kongresi, 08-12 Mayıs 2005, Bildiri ve Poster Özetleri Kitabı. Antalya. 250-251.

[10]. Ebegil, M. (2006). Shrinkage Tahmin Ediciler Sınıfı Üzerine Bir Çalışma, Sakarya Üniversitesi, Fen Edebiyat Fakültesi, Fen Edebiyat Dergisi 8(2), 53-58.

[11]. Ebegil, M. (2007). En Küçük Kareler Tahmin Edicisi İle Bir Shrınkage Tahmin Edicisinin Etkinlik Karşılaştırması, Anadolu Üniversitesi Bilim ve Teknoloji Dergisi 8(1), 79-83.

[12]. Rao, C. R., (1976). Estimation of Parameters in a Linear Models., The Annals of Statistics, 4, 1023-1037.

[13]. Ekni, M., (1999). EKK Tahmin Edicisi ile Bir Shrinkage Tahmin Edicisinin etkinlik karşılaştırması, Journal of the Institute of Science and Technology,Gazi Universitesi, 12(3), 509-514.

[14]. Norman, L. J. and Samuel, K.(1970), Distributions in Statistics, Chapter 26.

[15]. Theobald, C. M. (1974). Generalizations of Mean Square Error Applied to Ridge Regression, Journal of the Royal Statistical Society B,36, 103-106.

[16]. Patnaik, P. B. (1949), The Noncentral Ki-Kare and F Distributions and their Applications, Biometrika 36, 202-232.

[17]. Toutenburg, H. (1982), Prior Information in Linear Models. John Wiley and Sons. New York.