www.istatistikciler.org
İstatistikçiler Dergisi 1 (2008) 1-8

İstatistikçiler Dergisi

VERİLERİN LİNEER İÇ İLİŞKİLİ OLDUĞU LOJİSTİK REGRESYONDA BAZI YANLI PARAMETRE KESTİRİCİLERİ VE HATA KARELER ORTALAMASINA GÖRE KARŞILAŞTIRMALARI

Nurkut Nuray URGAN

Dicle Üniv. Fen-Ed.Fak. Matematik Böl.
nsazli@dicle.edu.tr

Müjgan TEZ

Marmara Üniv. Fen-Ed. Fak. Matematik Böl.
mtez@marmara.edu.tr

ÖZET

En çok olabilirlik kestirimi kullanılan lojistik regresyon yaygın kullanım alanlarına sahiptir. Lojistik regresyon genellikle, medikal alanlarda yaşam olasılığı modelleme ve hastalık ve travmadaki risk faktörlerinin değerlendirme gibi alanlarda kullanılır. Lojistik regresyonda bağımsız değişkenler arasındaki lineer iç ilişki en çok olabilirlik kestiricisinin (MLE) varyansını büyük oranda etkiler. Bu çalışmada MLE' ne alternatif olarak başka kestiriciler önerildi. Belirli koşullar altında Ridge (Smidt&McDonald,1976) kestiricisi, Stein (Stein,1960) kestiricisinin toplam hata kareler ortalamalarının dolayısıyla varyanslarının en çok olabilirlik kestiricisininkinden (MLE) daha küçük olduğu biliniyor. Buna benzer olarak, Ridge ve Stein kestiricilerinin konveks bileşimi olan Liu (Liu,1993) kestiricisinin lojistik regresyondaki toplam hata kareler ortalamasının da MLE kestiricisininkinden daha küçük olabileceği üzerinde çalışılmaktadır.

Anahtar Kelimeler: *Lojistik regresyon, En çok olabilirlik kestiricisi(MLE), Yinelemeli ağırlıklandırılmış en küçük kareler(IWLS), Ridge kestirimi, Stein kestirimi, Liu kestirimi, lineer iç ilişki.*

ABSTRACT

SOME BIASED ESTIMATORS AND THEIR COMPARISONS USING MEAN SQUARED ERROR CRITATION IN LOGISTIC REGRESSION WHEN THE DATA ARE COLLINEAR

Logistic regression using maximum likelihood estimation has recently gained widespread use. Many of these applications have been in situations in which the independent variables are collinear. The collinearity among the independent variables seriously effects the variance of the maximum likelihood estimation. In this study, several alterative estimators are suggested, which are combat the collinearity and easy to obtain in practice. It has been known that under some situations of the Ridge estimator and Stein estimator are less affected by the variance of the MLE. Similarly, we are studying on Liu estimator which is the convex combining of them in logistic regression.

Keywords: *Logistic regression, Maximum Likelihood Estimation (MLE), Iterative weighted least squares (IWLS), Ridge regression, Stein regression, Liu regression, collinearity.*

1. LOJİSTİK REGRESYON MODELİ

Lojistik regresyon modeli

$$y_i = \pi_i + \varepsilon_i \quad , \quad i = 1,2,...,n \tag{1.1}$$

$$\pi = P(y = 1 | x_1, x_2, ..., x_{p-1})$$

$$= \left[1 + \exp(-\beta_0 - \beta_1 x_1 - ... - \beta_{p-1} x_{p-1}\right]^{-1}$$

$$\pi = \left[1 + \exp(-\underline{x}'\underline{\beta})\right]^{-1} \tag{1.2}$$

şeklindedir. Burada y ikili (0/1) yanıt değişken, $\underline{\beta}$ bilinmeyen parametre vektörü, $\underline{x}' = (1, x_1, ..., x_{p-1})$ bağımsız değişken vektörü yada risk faktörleri, ε_i, $E(\varepsilon_i) = 0$ ve $\mathrm{v_i} = \{\pi_i(1-\pi_i)\}$, $V = diag\{v_i\}$ dağılımlı hata vektörüdür [6]. Gruplandırılmış veriler için n_i, i-inci veri noktasındaki deneme birimi olmak üzere (1.2) modeli

$$\begin{aligned} E(y_i) &= n_i P(x_i) \\ &= n_i \frac{1}{1+e^{-x_i'\beta}} \quad , \quad i = 1, 2, ..., m \end{aligned}$$

olarak yazılır. $y_1, y_2, ..., y_m$, bağımsız binom rastgele değişkenlerinin gözlem değerleridir.

$$Var(y_i) = n_i P(x_i)\left[1 - P(x_i)\right]$$

$$\sum_{i=1}^{m} n_i = \mathrm{n}$$

n ve P için olasılık fonksiyon; $\binom{\mathrm{n}}{y} P^y (1-P)^{\mathrm{n}-y}$ dir. Log olabilirliği bulunurken $\binom{\mathrm{n}}{y}$, β içermediğinden atılır ve böylece lojistik regresyon modelinin Log olabilirliği

$$\ln\left[\mathsf{L}(P;y)\right] = \sum_{i=1}^{m} \left\{ y_i \ln\left[\frac{P(x_i)}{1-P(x_i)}\right] + n_i \ln\left[1 - P(x_i)\right] \right\} \tag{1.3}$$

olur. Burada $\ln\left[\frac{P(x_i)}{1-P(x_i)}\right]$ lojit olarak adlandırılır ve

$$\begin{aligned} \ln\left[\frac{P(x_i)}{1-P(x_i)}\right] &= x_i'\beta \\ &= \beta_0 + \sum_{j=1}^{k} x_{ij}\beta_j \quad , \quad i = 1, 2, ..., m \quad , \quad m \geq k+1 \end{aligned}$$

dir. Böylece (1.3) denklemi

$$\ln\left[\mathsf{L}(\underline{\beta};y)\right] = \sum_{i=1}^{m}\sum_{j=1}^{k} y_i x_{ij}\beta_j - \sum_{i=1}^{m} n_i \ln\left(1 + \exp\sum_{j=1}^{k} x_{ij}\beta_j\right) \tag{1.4}$$

olarak yazılabilir. (1.4) denkleminin β_j'ye göre maksimumu alınırsa, (1.4) denkleminin matris formu

$$\ln\left[L(\underline{\beta};y)\right] = \underline{\beta}'Xy - \sum_{i=1}^{m} n_i \ln\left(1+\exp(x_i'\underline{\beta}\right) \tag{1.5}$$

yazılır. (1.5) denkleminin β ' ya göre türevi

$$\frac{\partial \ln\left[L(\beta;y)\right]}{\partial \beta} = X'y - \sum_{i=1}^{m}\left[\frac{n_i}{1+e^{x_i'\beta}}\right]e^{x_i'\beta}x_i$$

$$e^{x_i'\beta}\Big/(1+e^{x_i'\beta}) = 1\Big/(1+e^{x_i'\beta}) = P(x_i) \text{ olduğundan}$$

$$\frac{\partial \ln\left[L(\beta;y)\right]}{\partial \beta} = X'y - \sum_{i=1}^{m} n_i P(x_i)x_i$$

elde edilir. $n_i P(x_i)$ binom rastgele değişkeninin ortalaması olduğundan, denklemin sağ tarafının matris yazılımı $X'(y-\mu)$ olur, burada,

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix} \qquad ve \qquad \mu_i = n_i P(x_i)$$

dir. Sonuç olarak β 'nın ML kestirimi $\hat{\beta}$,

$$X'(y-\mu) = 0 \tag{1.6}$$

"skor denkleminin" çözümü ile bulunur. X bağımsız değişkenlerin matrisi ve $\tilde{x}_i'$ matrisin i-inci sütunu olmak üzere (1.2) modeli

$$\begin{aligned}\pi_i &= 1\Big/\left(1+\exp\left[-\tilde{x}_i'\underline{\beta}\right]\right) \\ &= f\left(\tilde{x}_i', \underline{\beta}\right)\end{aligned}$$

dır. Burada $\underline{\beta}$, β_j nin $\left((p+1)\times 1\right)$ vektörüdür. Buradaki y, $(n\times 1)$ tipinde vektör ve $\hat{\pi}$, $\hat{\pi}_i = f\left(\tilde{x}_i', \hat{\beta}\right)$ nin $(n\times 1)$ tipinde vektörüdür. β 'nın kestirimi yinelemeli ağırlıklandırılmış en küçük kareler (IWLS) tekniği ile de bulunabilir. ML veya IWLS ile β 'nın kestiriminin $(l+1)$. iterasyonu

$$\hat{\beta}_{l+1} = \hat{\beta}_l + (X'\hat{V}_l X)^{-1} X'(y-\hat{\pi}_l)$$

şeklindedir [8]. $\hat{\pi}_l$ $\hat{\beta}_l$ ve $\hat{V}_l = diag\left\{\hat{\pi}_{li}(1-\hat{\pi}_{li})\right\}$ 'nin kullanılmasıyla π 'nin tahmin vektörü, $\hat{\pi}_{li}$, $\hat{\pi}_l$ 'nin i-inci elemanıdır. Bu iterasyonda kestirim yakınsak olana kadar devam eder ve sonuç kestirim $\hat{\beta}$ ile gösterilir. $\hat{\beta}$ 'nın çok değişkenli normal dağılıma yaklaşması beklenir.

2. LİNEER İÇ İLİŞKİNİN EN ÇOK OLABİLİRLİK LOJİSTİK KESTİRİCİSİNE ETKİSİ

Lineer iç ilişki için birçok ölçüm önerilmiştir ve bunlardan en çok kullanılanları:
i) R_j^2, j-inci bağımsız değişkenin regresyondaki belirleyicilik katsayısı,
ii) $(\delta_j'\delta_j)$, (i) deki regresyondaki kalan kareler toplamı,
iii) μ_j, $(X'X)$' in $\mu_1 \le \mu_2 \le ... \le \mu_p$ sıralı gizli kökleridir.
Verilen bu ölçümlerin

i) R_j^2 *'nin bazı j ler için 1 yakınsaması,*

ii) $(\delta_j'\delta_j)$ *'nin bazı j ler için 0 yakınsaması,*

iii) μ_j *'nin bazı j ler için 0 yakınsaması*

durumunda lineer iç ilişki varlığı söylenebilir. Lineer iç ilişkinin derecesi bu ölçümlerin sırasıyla 1 ve 0 yakınsaması ile saptanır. Bu ölçümlerin avantaj ve dezavantajları vardır. Örneğin; R_j^2 'nin avantajı kolay yorumlanır olması ve en küçük kareler kestiriminin değişim artışını ölçen VIF=$1/1-R_j^2$ olduğundan VIF'e yakın olmasıdır, dezavantajı ise birden fazla değişken arasındaki lineer iç ilişkiyi tanımada başarısız olmasıdır.

$\hat{\beta}$'nın kovaryans matrisi $(X'VX)^{-1}$ dir ve büyük n değeri için $Var(\hat{\beta}_j)$ yaklaşık olarak $(X'VX)^{-1}$'in (j,j)-inci elemanı olan $(X'VX)^{jj}$ dir. x_j, X matrisinin *j*-inci sütunu ve X_j, X matrisinin *j*-inci sütununun silinmesiyle kalan $(n \times p)$ tipindeki X matrisi, $\left[x_j, X_j\right]$ olacak şekilde parçalı matris olarak yazıldığında

$$\begin{aligned} Var(\hat{\beta}_j) &\approx \left(X'\hat{V}X\right)^{jj} \\ &= \left\{x_j'Vx_j - x_j'VX_j\left(X_j'VX_j\right)^{-1}X_j'Vx_j\right\}^{-1} \end{aligned} \tag{2.1}$$

şeklinde bulunur [7]. (2.1) denkleminde, a_j sabitlerin vektörü olmak üzere, bazı *j*'ler için x_j yerine $x_j = X_j a_j + \delta_j$ yazıldığında,

$$Var(\hat{\beta}_j) \approx \left(\delta_j' S \delta_j\right)^{-1}$$

elde edilir. Buradaki *S*, $S = V - VX_j\left(X_j'VX_j\right)^{-1}X_j'V$ dir. Pratikte x_j nin elemanları sonludur ve bundan dolayı $n \to \infty$ için $\lim(1/n)(X'VX)$ sonlu pozitif tanımlıdır. Ve *V* sınırlı olduğundan, *S* de sınırlı elde edilir. Böylece veriler daha fazla lineer iç ilişkili olduğunda, bazı *j* ler için $\delta_j'\delta_j \to 0$ ve $Var(\hat{\beta}_j) \to \infty$ olur. Eğer lineer iç ilişkinin derecesi büyükse, kestirimlerin biri veya daha fazlası belirsiz olacaktır ve kestirimler bağımsız değişkenlerin doğru etkilerini yansıtmayacaktır.

3. RİDGE LOJİSTİK KESTİRİMİ

Lineer iç ilişki problemi Hoerl ve Kennard(1970a.b) ın Ridge kestirim tekniğinin kullanılmasıyla ele alınabilir. Schaeffer, Roi ve Wolfe(1984)' de lojistik regresyonda kullanmak için Ridge kestirici ele alınmış ve

$$E(\hat{\beta}'\hat{\beta}) = \hat{\beta}'\hat{\beta} + \dot{I}z\left\{Var\left(\hat{\beta}'\right)\right\} + \left(Yan(\hat{\beta})\right)\left(Yan(\hat{\beta})\right)' \geq \hat{\beta}'\hat{\beta} + \Sigma\ Var\left(\hat{\beta}'\right)$$

olarak gösterilmiş ve buradan eğer veriler lineer iç ilişkili ise ençok olabilirlik kestiriminin ortalamadan çok uzun olacağı belirtilmiştir. $\hat{\beta}$, IWLS kestirici olduğunda, $\hat{\beta}$ yaklaşık olarak Ağırlıklı Hata kareler toplamını (WSSE) minimum yapar. Buradan Ridge lojistik kestiricisi, WSSE yi arttıracak minimum uzunluğa sahip kestirici olarak belirlenir [10]. Sabit kestiriciler için $\hat{\beta}$ yerine $\hat{B}$ yazılarak WSSE deki artış

$$\phi = \left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right]' V^{-1}\left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right] + 2\left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right]' V^{-1}\left[y - \hat{\pi}(\hat{\beta})\right]$$

şeklindedir. Buradaki $\hat{\pi}(\hat{B})$, β'nın kestrimi olarak $\hat{B}$ kullanılmasıyla bulunan π'nin tahminidir. $\hat{\pi}$' lar lineerleştirilerek

$$\phi = \left(\hat{B} - \hat{\beta}\right)' X'VX\left(\hat{B} - \hat{\beta}\right) \tag{3.1}$$

elde edilir ve $\hat{B}$'nın uzunluğunun minimize edilmesiyle Ridge lojistik kestiricisi

$$\hat{\beta}_R = \left(X'VX + kI\right)^{-1} X'VX\hat{\beta} \tag{3.2}$$

olarak elde edilir [1]. Buradaki k Ridge parametresi olarak adlandırılır. Pratikte $\hat{\beta}_R$, $\hat{\beta}$ nın ve $\left(X'VX\right)^{-1}$ in bir fonksiyonu olduğundan hesaplanması kolaydır. $\hat{\beta}$ ve $\left(X'VX\right)^{-1}$ in ikisi de çoğu lojistik regresyon programlarında kolayca bulunur. k'nın seçimi için kesin bir kural olmamakla birlikte verilen bazı seçim yolları aşağıdaki gibi verilmiştir.

I) $k = \dfrac{1}{\max\limits_j\left(\gamma_j'\hat{\beta}\right)^2}$, γ_j', $\left(X'\hat{V}X\right)$' in gizli vektörleri. [2;3]

II) $k = \dfrac{p+1}{\hat{\beta}'\hat{\beta}}$ [4]

III) $k = \dfrac{1}{\hat{\beta}'\hat{\beta}}$ [7]

4. STEİN KESTİRİMİ

$$E(\hat{\beta}'\hat{\beta}) \geq \dot{I}z\left(Var(\hat{\beta})\right) = \dot{I}z\left\{(X'\hat{V}X)^{-1}\right\}$$

olduğu gösterilebilir. Lineer iç ilişki varlığında $\hat{\beta}$ ortalama olarak fazla uzun olacaktır. ML kestiricisini küçültmek için,

$$\hat{\beta}_s = c\hat{\beta} \quad , \quad 0 < c < 1$$

olan Stein kestiricisi tanımlanır. c 'nin seçimi için, Stein(1960)'ın öne sürdüğü

$$c = \frac{\left(\hat{\beta}'\hat{\beta}\right)}{\left[\hat{\beta}'\hat{\beta} + \dot{I}z\left\{\left(X'\hat{V}X\right)^{-1}\right\}\right]}$$

kullanılır.

Çoğu lojistik regresyon programlarından kolaylıkla elde edilebilen niceliklere bağlı olduğundan, bu kestiriciyi yürütmek kolaydır.

5. LİU KESTİRİMİ

Lineer iç ilişki problemi için bir başka alternatif kestirici Liu Kejian(1993)'nın öne sürdüğü Liu kestiricisidir. Bu kestiricinin lojistik regresyon için gösterimi aşağıdaki gibi bulunur.

$$\pi_i = 1 \Big/ 1 + \exp\{-x_i\beta\}$$

$$E(\hat{\beta}'\hat{\beta}) = \beta'\beta + \dot{I}z\left\{Var(\hat{\beta})\right\} + \left(Yan(\hat{\beta})\right)\left(Yan(\hat{\beta})\right)' \geq \beta'\beta + \sum Var(\hat{\beta}_j)$$

$$\phi = \left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right]' V' \left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right] + 2\left[\hat{\pi}(\hat{\beta}) - \hat{\pi}(\hat{B})\right]' V^{-1}\left[y - \hat{\pi}(\hat{\beta})\right]$$

$\hat{\pi}$ ' lar lineerleştirilerek

$$\phi = \left(\hat{B} - \hat{\beta}\right)' X'VX\left(\hat{B} - \hat{\beta}\right)$$

bulunur. $\hat{B}$ 'nın uzunluğunu minimize etmek için, Lagrangian probleminden

$$F = \hat{B}'\hat{B} + 1\Big/k\left[\left(\hat{B} - \hat{\beta}\right)' X'VX\left(\hat{B} - \hat{\beta}\right) - \phi\right]$$

şeklinde yazılır, buradan da $0 = k^{1/2}\beta + \varepsilon'$, $d\hat{\beta} = \beta + \varepsilon'$ genişletilmesiyle

$$\hat{B} = \hat{\beta}_d = \left(X'VX + I\right)^{-1}\left(X'VX + dI\right)\hat{\beta} \tag{4.1}$$

Liu lojistik kestiricisi bulunur, buradaki d , 0<d<1 aralığında değer alan bir parametredir. $\hat{\beta}_d$ ' nın MSE' sini bulmak için, $O(1/n^2)$ dizilimi aşağıdaki gibi yapılır, m, ε da kuadratik form vektörü olmak üzere ve Taylor seriye açılımından

$$\hat{\beta}_d = \beta + \left(X'VX + I\right)^{-1}\left(I + d(X'VX)^{-1}\right)X'\varepsilon - \frac{1}{2}\left(X'VX + I\right)^{-1}\left(I + d(X'VX)^{-1}\right)X'm \tag{4.2}$$

şeklinde yazılır. $E(\varepsilon) = 0$ ve $Var(\varepsilon) = V$ varsayıldığından, $\hat{\beta}_d$ ' nın yanı ve varyansı (4.2) denkleminden aşağıdaki gibi bulunur.

$$\begin{aligned} MSE(\hat{\beta}_d) = İz\Big\{&\left(X'VX + I\right)^{-1}\left(I + d(X'VX)^{-1}\right)\left[X'VX + \tfrac{1}{4}X'Var(m)X - X'Cov(\varepsilon, m)X\right]\left(X'VX + I\right)^{-1}\left(I + d(X'VX)^{-1}\right)\Big\} \\ &+ \left[(1-d)\beta - \tfrac{1}{2}X'h\right]'\left(I + d(X'VX)^{-1}\right)'\left(X'VX + I\right)^{-2}\left(I + d(X'VX)^{-1}\right)\left[(1-d)\beta - \tfrac{1}{2}X'h\right] \end{aligned}$$

λ_j ler $\left(X'VX\right)$' in $\lambda_0 \le \lambda_1 \le \ldots \le \lambda_p$ şeklinde sıralı gizli kökleri ve γ_j ler de gizli vektörleri olmak üzere

$$\begin{aligned} MSE(\hat{\beta}_d) = &\sum\left(\lambda_j + 1\right)^{-2}\left(1 + d\lambda_j^{-1}\right)^2\left\{\lambda_j + \gamma_j X'\left[\tfrac{1}{4}Var(m) - Cov(\varepsilon, m)\right]X\gamma_j\right\} \\ &+ \sum\left(\lambda_j + 1\right)^{-2}\left(1 + d\lambda_j^{-1}\right)^2\left\{\gamma_j'\left[(1-d)\beta - \tfrac{1}{2}X'h\right]\right\}^2 \end{aligned} \tag{4.3}$$

yazılıabilir. Genelliği kaybetmeden veriler arasında sadece bir iç lişki olduğu ve sadece λ_0 sıfıra yaklaştığı varsayımı altında, (4.3) denkleminin d ye göre türevi alınıp, yeterince büyük n için $O(\frac{1}{n^2})$ sırası ihmal edildiğinde

$$\left[MSE(\hat{\beta}_d)\right]' = 2\sum\left(\lambda_j + 1\right)^{-2}\left(1 + d\lambda_j^{-1}\right)\left(\lambda_j\right) - 2\sum\left(\lambda_j + 1\right)^{-2}\left(1 + d\lambda_j^{-1}\right)^2\left\{\gamma_j'\left[(1-d)\beta - \tfrac{1}{2}X'h\right]\right\}\left(\gamma_j'\beta\right)$$

(4.4)

şeklinde bulunur. $d = 1$ için $\hat{\beta} = \hat{\beta}_d$ olduğundan $MSE'(\hat{\beta}) = MSE'(\hat{\beta}_{d=1})$ olur. d=1 için $MSE'(\hat{\beta}_{d=1})$

$$\left[MSE(\hat{\beta}_{d=1})\right]' = 2\sum\left(\lambda_j + 1\right)^{-1} + 2\sum\lambda_j^{-2}\left\{\gamma_j'\left[\tfrac{1}{2}X'h\right]\right\}\left(\gamma_j'\beta\right) = \left[MSE(\hat{\beta})\right]'$$

dır. $\sum\left(\lambda_j + 1\right)^{-2}\left(1 + d\lambda_j^{-1}\right)^2 \ge 0$ olduğundan $0 < d < 1$ için $MSE'(\hat{\beta}) \ge MSE'(\hat{\beta}_d)$ olur. Yani $0 < d < 1$ için $MSE(\hat{\beta}_d) - MSE(\hat{\beta}) \le 0$ dır, bu da Liu lojistik kestiriminin en çok olabilirlik kestriminden daha iyi olduğu anlamına gelir [9].

Özet olarak, Ridge, Stein ve Liu kestirimlerinin varyansları ML kestiriminin varyansından daha küçük olduğundan, lineer iç ilişkiden daha az etkilenir. (Bu üç kestiricinin ML kestiriciyi küçültüğünü göstermek kolaydır.) Böylece bu üç kestirim, iç ilişkinin varyans üzerindeki etkilerini azaltarak daha kesin kestirimle sonuçlanır.

KAYNAKLAR

[1] Gunst, R.F., Mason, R.L. (1977b). *Biased estimation in Regression: An evaluation Using Mean Square Error*. JASA, 72, pp. 616-627.
[2] Hoerl, A.E., Kennard, R.W. (1970a). *Ridge Regression: Biased estimation for Nonorthogonal problems.* Technometrics, 12, pp. 55-67.
[3] Hoerl, A.E., Kennard, R.W. (1970b). *Ridge Regression: Applications Nonorthogonal problems*. Technometrics, 12, pp. 69-82.
[4] Hoerl, A.E., Kennard, R.W. and Baldwin, K.F. (1975). *Ridge Regression: Some Simulations*. Communications in Statistics, 4(2), pp. 1105-1123.
[5] Kejian, L. (1993). *A new class of biased estimate in linear regression.* Commun. Statistics: Theory and Methods, 22(2), pp. 393-402.
[6] Myers, R.H., Montgomery, D.C., Vining, G.G. (2001). *Generalized Linear Models: with applications in engineering and sciences.* Wiley & Sons, Inc., New York.
[7] Schaeffer, R.L., Roi, L.D. and Wolfe, R.A. (1984). *A Ridge logistic estimator.* Communications in Statistics: Theory and Methods, Vol.13, No.1, pp. 99-113.
[8] Schaeffer, R.L. (1986). *Alternative estimators in logistic regression when the data are collinear.* J. Statist. Comput. Simul., Vol. 25, pp. 75-91.
[9] Urgan, N.N. (2007). *Lojistik Regresyonda Bazı Yanlı Kestiricilerin İncelenmesi.* Doktora tezi. Dicle Üniversitesi, Fen Bilimleri Enstitüsü.