

A SIMULATION STUDY OF SOME SHRINKAGE ESTIMATORS

Meral Ebegil*, Fikri Gökpinar* and Müslim Ekni*

Received 21:08:2005 : Accepted 03:02:2006

Abstract

In regression analysis, it is desired that no multicollinearity should exist between the independent (explanatory) variables. In the cases where this is not achieved, the use of the Least Square (LS) estimation method leads to mismodelling. Some methods have been developed to solve this problem; one of which is the ‘biased estimation method’. In this study, a test statistics for Ridge and Liu estimators, that are shrinkage biased estimators, is analyzed. Moreover, these estimators are compared via simulation, in terms of different correlation coefficients between the independent variables.

Keywords: Linear Admissible Estimators, Mean Square Error, Central-F Approximations, Liu Estimator.

2000 AMS Classification: 62J07

1. Introduction

Multiple linear regression is one of the statistical methods in widespread use that help one to bring out relationships between variables. Researchers working on data analysis use the multi linear regression method for forming models. The most common method used for estimating the regression coefficients is the LS method. However, in order for the LS method to give valid results, some assumptions need to be made.

In multiple linear regression analysis, there should be no relations between the independent variables. Nevertheless, in reality, this may not always be realized. Using the LS estimation method in this case may lead to an improper use of the model. Some methods have been developed to allow the analysis of the case when the independent variables depend on each other. One of these methods is the biased estimation method. The most widely used biased estimation methods are; principal components regression, ridge regression and their variations. Estimations produced by biased methods are more biased than the LS estimators are, but they produce less variable estimations. The main

*Gazi University, Faculty of Arts and Sciences, Department of Statistics, Ankara, Turkey.

purpose of biased estimation methods is to reduce the large variance area of the LS estimation method at the cost of a small bias. Therefore, more valid results can be obtained in comparison with the LS method.

One type of biased estimator is the so called *shrinkage estimator*. Basic components regression, ridge regression and their derivatives are of this type. In his study, Farebrother [3] formed a general structure for shrinkage estimators. He was able to place ridge, basic components and conditioned-minimum mean error square biased estimators within this structure by showing that each of them is a shrinkage estimator. In Liski [6], the powerful Mean Square Error (MSE) was proposed as a criteria to choose between the LS estimator and the shrinkage estimator. Liski [7] used a weaker MSE test to make a choice between the LS estimator and the shrinkage estimators. Kejian [5] suggested the Liu- Kejian estimator as an alternative to the ridge estimator. Later, this estimator was named the "Liu estimator" by Akdeniz and Kaçiranlar [2]. Sakalhoğlu, Kaçiranlar and Akdeniz [11] compared the Liu estimator with its iteration estimator. Then, Akdeniz and Erol [1] compared various shrinkage estimators, and gave a numerical example.

The second section of this study describes the basic structure of shrinkage estimators. Keeping this information in mind, necessary and sufficient conditions for the shrinkage estimators to give better results than the LS estimator are obtained by comparing their MSE matrices. A hypothesis test based on a test statistics derived from this condition is then examined.

In the last section, a simulation study is carried out using a MATLAB program. As a result of this simulation study, rejection and acceptance areas for the Ridge and Liu shrinkage estimators, constructed according to this hypothesis test were compared and the results interpreted.

2. The general structure of shrinkage estimators

In this section the general structure of shrinkage estimators is presented, inspired by the linear regression model of (Farebrother, [3]).

A multi linear regression model with n observation and k independent variables is defined as in (Farebrother, [3]):

$$(1) \quad \begin{aligned} \underline{Y} &= X\underline{\beta} + \underline{\varepsilon} \\ \underline{\varepsilon} &\sim (\underline{0}, \sigma^2 I_n) \\ \text{rank}(X_{n \times q}) &= q \leq n \end{aligned}$$

Here, \underline{Y} is the $(n \times 1)$ dimensional vector of dependent variables; X the $(n \times q)$ dimensional non-stochastic input matrix ($q = k + 1$); $\underline{\beta}$ the $(q \times 1)$ dimensional vector of unknown coefficients; and $\underline{\varepsilon}$ the error vector satisfying $E(\underline{\varepsilon}) = \underline{0}$, $E(\underline{\varepsilon}\underline{\varepsilon}') = \sigma^2 I_n$.

General linear estimators are described in the following form. Here, C and \underline{c} are a matrix and a vector, respectively.

$$(2) \quad \underline{\tilde{\beta}} = C\underline{Y} + \underline{c}.$$

This estimator is referred to as the *linear estimator of $\underline{\beta}$* [14]. In (2), if one takes $C = (X'X)^{-1}X'$ and $\underline{c} = \underline{0}$, a special case of the estimator is obtained as follows;

$$\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{Y}.$$

This estimator is called the LS estimator of $\underline{\beta}$.

Under the model (1), the form in (2) can be represented as:

$$(3) \quad \underline{\tilde{\beta}} = A\underline{\hat{\beta}} + \underline{d}.$$

Here, A is the $(q \times q)$ dimensional matrix of constants and \underline{d} a vector of dimension $(q \times 1)$. It is obvious that when $A = I$ and $\underline{d} = \underline{0}$, $\tilde{\underline{\beta}}$ is an LS estimator. It is seen that the statements (2) and (3) are the same.

When determining the best estimator from among the unbiased estimators, the one with the minimum variance is preferred. When biased estimators are concerned, the MSE is used for determining the best estimator. This is because biased and unbiased estimators can be identified by checking the MSE matrices. The MSE matrices may be written as:

$$(4) \quad \text{MSE}(\tilde{\underline{\beta}}) = E(\tilde{\underline{\beta}} - \underline{\beta})(\tilde{\underline{\beta}} - \underline{\beta})'$$

When statement (1) is used instead of (2), it takes the following form:

$$\tilde{\underline{\beta}} = C(X\underline{\beta} + \underline{c}) + \underline{\varepsilon}.$$

When $\underline{\beta}$ is eliminated from both sides of this equation and the expected value of the statement is used, the biased value becomes,

$$(5) \quad \begin{aligned} E(\tilde{\underline{\beta}}) - \underline{\beta} &= \text{Bias}(\tilde{\underline{\beta}}) \\ &= (CX - I)\underline{\beta} + \underline{c}, \end{aligned}$$

and its variance is calculated as,

$$(6) \quad \text{Var}(\tilde{\underline{\beta}}) = C\text{Var}(\underline{Y})C' = \sigma^2CC'.$$

The well-known form of the MSE matrices is $\text{MSE} = \text{Variance} + (\text{Bias})^2$. Accordingly, the MSE matrix can be written as,

$$\text{MSE}(\tilde{\underline{\beta}}) = \sigma^2CC' + [(CX - I)\underline{\beta} + \underline{c}][(CX - I)\underline{\beta} + \underline{c}]'$$

Another form uses the *scalar mean square error* (SMSE):

$$(7) \quad \text{SMSE}(\tilde{\underline{\beta}}) = E(\tilde{\underline{\beta}} - \underline{\beta})'(\tilde{\underline{\beta}} - \underline{\beta}).$$

In practice the trace operator is applied in the form:

$$(8) \quad \begin{aligned} \text{SMSE}(\tilde{\underline{\beta}}) &= \text{tr}E(\tilde{\underline{\beta}} - \underline{\beta})(\tilde{\underline{\beta}} - \underline{\beta})' \\ &= \text{trMSE}(\tilde{\underline{\beta}}). \end{aligned}$$

For unbiased estimators, $\underline{c} = \underline{0}$, and the known condition for being unbiased is $CX - I = 0$. In this case, $\text{MSE}(\tilde{\underline{\beta}}) = \text{Var}(\tilde{\underline{\beta}}) = \sigma^2CC'$. For example, this is the case for the LS estimator. In other words, the matrix of MSE is equal to the variance. Because the LS estimator is an unbiased estimator for the parameter we have in this case $\text{MSE}(\hat{\underline{\beta}}) = \text{Var}(\hat{\underline{\beta}}) = \sigma^2(X'X)^{-1}$.

The measure of loss for the estimator $\tilde{\underline{\beta}}$ related to the parameter $\underline{\beta}$ is defined as

$$(9) \quad L(\tilde{\underline{\beta}}, A) = (\tilde{\underline{\beta}} - \underline{\beta})' A (\tilde{\underline{\beta}} - \underline{\beta}).$$

Here A is a $(k \times k)$ dimensional positive definite symmetric matrix. The risk of the estimator $\tilde{\underline{\beta}}$ is defined by

$$(10) \quad R(\tilde{\underline{\beta}}, A) = E(L(\tilde{\underline{\beta}}, A)) = E(\tilde{\underline{\beta}} - \underline{\beta})' A (\tilde{\underline{\beta}} - \underline{\beta}).$$

By applying the trace operator, the relation between the risk and the MSE of the estimator can be expressed as:

$$(11) \quad R(\underline{\tilde{\beta}}, A) = \text{tr}AE(\underline{\tilde{\beta}} - \underline{\beta})(\underline{\tilde{\beta}} - \underline{\beta})' = \text{tr}AMSE(\underline{\tilde{\beta}}).$$

By considering (11) together with Theorem 2.1 below, minimization of the risk $R(\underline{\tilde{\beta}}, A)$ is equivalent to the minimization of $MSE(\underline{\tilde{\beta}})$, for any estimator $\underline{\tilde{\beta}}$ belonging to the fixed set of possible estimators [14].

2.1. Theorem. (Theobald, [13]) *A symmetric $n \times n$ matrix D is non-negative definite if and only if $\text{tr}CD \geq 0$ for all non-negative definite C .*

For example, for all $C \geq 0$ we have $D \geq 0 \iff \text{tr}CD \geq 0$.

2.2. Theorem. (Theobald, [13]): *Suppose that two estimators are given as $\underline{\beta}_1^*$ and $\underline{\beta}_2^*$. Then the following are equivalent:*

- (i) $MSE(\underline{\beta}_1^*) - MSE(\underline{\beta}_2^*) \geq 0$.
- (ii) $R(\underline{\beta}_1^*, A) - R(\underline{\beta}_2^*, A) \geq 0$ for all $A \geq 0$.

A good estimator can be found by minimizing the risk over the class of defined estimators. Such estimators are called *R-optimal* [14]. Another problem is to find the better of two proposed estimators. That is the essential problem is the problem of comparing the two estimators.

2.3. Definition. When,

$$(12) \quad \begin{aligned} &\text{for all } \underline{\beta} \text{ we have } R(\underline{\tilde{\beta}}_2, A) - R(\underline{\tilde{\beta}}_1, A) \geq 0, \text{ and,} \\ &\text{for some } \underline{\beta} \text{ we have } R(\underline{\tilde{\beta}}_2, A) - R(\underline{\tilde{\beta}}_1, A) > 0, \end{aligned}$$

the estimator $\underline{\tilde{\beta}}_1$ is said to be better than $\underline{\tilde{\beta}}_2$ with respect to the square loss measurement $L(\underline{\tilde{\beta}}, A)$.

If for the matrix differences we have,

$$(13) \quad \begin{aligned} &\text{for all } \underline{\beta} \text{ we have } MSE(\underline{\tilde{\beta}}_2) - MSE(\underline{\tilde{\beta}}_1) \geq 0, \text{ and,} \\ &\text{for some } \underline{\beta} \text{ we have } MSE(\underline{\tilde{\beta}}_2) - MSE(\underline{\tilde{\beta}}_1) \neq 0, \end{aligned}$$

then taking (11) and Theorem 1 into consideration, it can be concluded that the estimator $\underline{\tilde{\beta}}_1$ is better than $\underline{\tilde{\beta}}_2$.

The relation between the estimators $\underline{\tilde{\beta}}_1$ and $\underline{\tilde{\beta}}_2$ in (12) depends on the matrix A of the loss function in (9). In other words, the use of different weighted matrices leads to different loss functions and therefore to different risks.

If there is no better estimator, its admissibility as an estimator is described in terms of the risk $R(\underline{\tilde{\beta}}, A)$. In general, for the class of all linear estimators $\underline{\tilde{\beta}} = C\underline{Y} + \underline{c}$, admissibility depends on a given estimator class. Admissibility is defined in terms of the MSE of $\underline{\tilde{\beta}}$ in (4).

2.4. Definition. Consider the conditions:

$$(14) \quad \begin{aligned} &\text{for all } \underline{\beta} \text{ we have } \text{MSE}(\tilde{\underline{\beta}}) - \text{MSE}(\hat{\underline{\beta}}) \geq 0, \text{ and} \\ &\text{for some } \underline{\beta} \text{ we have } \text{MSE}(\tilde{\underline{\beta}}) - \text{MSE}(\hat{\underline{\beta}}) \neq 0, \end{aligned}$$

If these conditions are satisfied for all the estimators $\tilde{\underline{\beta}}$ then the estimator $\hat{\underline{\beta}}$ is said to be *admissible with respect to the MSEs*.

In view of Theorem 2.2, for all positive definite matrices, that is for all $A \geq 0$, $\hat{\underline{\beta}}$ is an admissible estimator with respect to $R(\hat{\underline{\beta}}, A)$. In order to prove this it is sufficient to display the admissibility of $\hat{\underline{\beta}}$ with respect to $R(\hat{\underline{\beta}}, I)$.

2.5. Theorem. (Toutenburg, [14], pp.12) *Let the estimator $\hat{\underline{\beta}}$ be admissible with respect to $R(\hat{\underline{\beta}}, I)$. Then, $\hat{\underline{\beta}}$ is an admissible estimator under $R(\hat{\underline{\beta}}, A)$ for each $A \geq 0$.*

For admissibility, the form of the estimator $\underline{\beta}$ is as follows [10]:

$$(15) \quad \tilde{\underline{\beta}} = A(\hat{\underline{\beta}} - \underline{b}) + \underline{b}.$$

Here, A is a $(q \times q)$ dimensional matrix and \underline{b} is a constant vector. Estimators defined this way belong to the class of linear admissible estimators. Other conditions for admissibility are that

$$(16) \quad (X'X)A \text{ or } A(X'X)^{-1} \text{ should be symmetric,}$$

and

$$(17) \quad \text{The eigenvalues of } A \text{ lie in } [0, 1].$$

Since $X'X$ and A are symmetric, there exists a $(q \times q)$ orthonormal matrix P such that $P'X'XP = \Lambda$ is a $(q \times q)$ diagonal matrix whose diagonal elements $\lambda_1, \lambda_2, \dots, \lambda_q$ are the eigenvalues of $X'X$, assumed to be in descending order. Also $P'AP$ is a diagonal matrix whose diagonal elements $\delta_1, \delta_2, \dots, \delta_q$ lie in $[0, 1]$ [6,12].

Model (1) can be written in canonical form as

$$(18) \quad \underline{Y} = XPP'\underline{\beta} + \underline{\varepsilon} = Z\underline{\alpha} + \underline{\varepsilon}$$

where $Z = XP$ and $\underline{\alpha} = P'\underline{\beta}$. Within this model (15) becomes:

$$(19) \quad \begin{aligned} \tilde{\underline{\alpha}} &= P'A(\hat{\underline{\beta}} - \underline{b}) + P'\underline{b} \\ &= P'AP(\hat{\underline{\alpha}} - \underline{a}) + \underline{a} = \Delta(\hat{\underline{\alpha}} - \underline{a}) + \underline{a}. \end{aligned}$$

Here, $\hat{\underline{\alpha}} = P'\hat{\underline{\beta}}$, $\underline{a} = P'\underline{b}$. These kind of admissible linear estimators are called shrinkage estimators.

2.1. Mean Square Error Matrices for Shrinkage Estimators. The main problem for shrinkage estimators is to determine conditions under which the risk of these estimators is less than that of the LS estimator (Liski, [6]). The MSE matrices of the estimators $\hat{\underline{\beta}}$ and $\tilde{\underline{\beta}}$ are respectively:

$$(20) \quad \begin{aligned} \text{MSE}(\hat{\underline{\beta}}) &= \sigma^2 (X'X)^{-1}, \text{ and} \\ \text{MSE}(\tilde{\underline{\beta}}) &= \sigma^2 A (X'X)^{-1} A' + (I - A)(\underline{\beta} - \underline{b})(\underline{\beta} - \underline{b})'(I - A). \end{aligned}$$

Equivalently, the canonical form may be given as:

$$(21) \quad \text{MSE}(\tilde{\underline{\alpha}}) = \sigma^2 \Delta \Lambda^{-1} \Delta + (I - \Delta)(\underline{\alpha} - \underline{a})(\underline{\alpha} - \underline{a})'(I - \Delta).$$

If the matrix difference $\text{MSE}(\hat{\underline{\beta}}) - \text{MSE}(\tilde{\underline{\beta}})$ is not negative definite then $\text{MSE}(\hat{\underline{\beta}}) - \text{MSE}(\tilde{\underline{\beta}}) \geq 0$. It follows that the difference $\text{MSE}(\hat{\underline{\beta}}) - \text{MSE}(\tilde{\underline{\beta}})$ is not negative definite if and only if the inequality

$$(22) \quad (\underline{\beta} - \underline{b})' (I + A)^{-1} X' X (I - A) (\underline{\beta} - \underline{b}) / \sigma^2 \leq 1.$$

is satisfied. Likewise by [7] the difference $\text{MSE}(\hat{\underline{\alpha}}) - \text{MSE}(\tilde{\underline{\alpha}}) \geq 0$ is not negative definite if

$$(23) \quad (\underline{\alpha} - \underline{a})' (I + \Delta)^{-1} \Lambda (I - \Delta) (\underline{\alpha} - \underline{a}) / \sigma^2 \leq 1.$$

The inequality (23) can be stated as:

$$(24) \quad \sum_{i=1}^q \gamma_i \lambda_i (\alpha_i - a_i)^2 / \sigma^2 \leq 1.$$

Here, $\gamma_i = \frac{(1 - \delta_i)}{(1 + \delta_i)}$. This result was obtained by Liski [7].

The Ridge and Liu estimators, which are known to be shrinkage estimators, can be defined respectively by,

$$(25) \quad \tilde{\underline{\beta}}_R = (X'X + kI)^{-1} X'Y,$$

$$(26) \quad \tilde{\underline{\beta}}_{Liu} = (X'X + I)^{-1} (X'Y + d\hat{\underline{\beta}}),$$

where $0 < k < 1$ and $0 < d < 1$ [5]. The necessary and sufficient condition for the Ridge estimator to have a lower risk than the LS estimator takes the form

$$(27) \quad \frac{(\underline{\beta} - \underline{b})' \left(\frac{2}{k}I + (X'X)^{-1} \right)^{-1} (\underline{\beta} - \underline{b})}{\sigma^2} \leq 1,$$

or

$$(28) \quad \frac{(\underline{\alpha} - \underline{a})' \left(\frac{2}{k}I + \Lambda^{-1} \right)^{-1} (\underline{\alpha} - \underline{a})}{\sigma^2} \leq 1.$$

In a similar way, the corresponding necessary and sufficient condition for the Liu estimator has the form

$$(29) \quad (\underline{\beta} - \underline{b})' \left(X'X + 2I + d(X'X)^{-1} \right)^{-1} \left((X'X)^2 + dI \right) (\underline{\beta} - \underline{b}) / \sigma^2 \leq 1,$$

or

$$(30) \quad (\underline{\alpha} - \underline{a})' (\Lambda + 2I + d\Lambda^{-1})^{-1} (\Lambda^2 + dI) (\underline{\alpha} - \underline{a}) / \sigma^2 \leq 1.$$

The necessary and sufficient conditions for the Ridge and $\hat{\underline{\beta}}_d = (dI)\hat{\underline{\beta}}$, $0 \leq d \leq 1$, shrinkage estimators to have a lower risk than the LS estimator are given in the works of Liski [6,7].

2.2. A Test for the Selection of Shrinkage Estimators. Using the necessary and sufficient conditions given in (22) and (23) it is possible to choose between two estimators. In this case, the form of the test statistics to be used in selecting between the shrinkage estimator $\tilde{\underline{\beta}}$ and the LS estimator $\hat{\underline{\beta}}$ is based on that inequality. Liski [6], investigated the test statistic

$$(31) \quad \tilde{F} = \hat{\underline{\beta}}' H \hat{\underline{\beta}} / m \hat{\sigma}^2,$$

where

$$\begin{aligned} H &= (I + A)^{-1} X'X(I - A), \\ \hat{\sigma}^2 &= (\underline{Y} - X\hat{\beta})'(\underline{Y} - X\hat{\beta}) / (n - q), \\ \text{rank}(H) &= m, \text{ and} \\ \underline{b} &= \underline{0}. \end{aligned}$$

The canonical form of this statistic can be written as:

$$\begin{aligned} \tilde{F} &= \hat{\underline{\alpha}}'(I + \Delta)^{-1} \Lambda(I - \Delta)\hat{\underline{\alpha}}/m\sigma^2 \\ (32) \quad &= \frac{1}{m} \sum_{i=1}^q \gamma_i (\lambda_i \hat{\alpha}_i^2 / \hat{\sigma}^2) \end{aligned}$$

Here $m = \text{rank}(I + \Delta)^{-1} \Lambda(I - \Delta)$. We may write

$$(33) \quad \tilde{F} = \frac{1}{m} \sum_{i=1}^m \gamma_i (\lambda_i \hat{\alpha}_i^2 / \hat{\sigma}^2),$$

the number of non-zero γ_i s being m . As it can be seen, m satisfies $1 \leq m \leq q$. Here, each $\frac{\lambda_i \hat{\alpha}_i^2}{\hat{\sigma}^2}$ has a non-central F distribution with degrees of freedom 1 and $(n - q)$ and non-central parameter

$$w_i = \frac{\lambda_i \alpha_i^2}{\sigma^2}.$$

When we write

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

we obtain

$$(34) \quad \tilde{F} = \frac{1}{m} \sum_{i=1}^m \gamma_i F_i.$$

In this case, the test statistic \tilde{F} is formed from a mixture of the statistics F_i . In other words, the expression (33) can be written as in (34). Here, $\lambda_i \hat{\alpha}_i^2 / \hat{\sigma}^2$ fits a chi-square distribution having non-central parameter $w_i = \frac{\lambda_i \alpha_i^2}{\sigma^2}$ and degree of freedom 1. On the other hand, $(n - q) \hat{\sigma}^2 / \sigma^2$ has a chi-square distribution with degree of freedom $(n - q)$. As previously mentioned, $\frac{\hat{\sigma}^2}{\sigma^2}$ and $\frac{\lambda_i \hat{\alpha}_i^2}{\sigma^2}$ are independent. In this case, $F_i = \frac{\lambda_i \hat{\alpha}_i^2}{\hat{\sigma}^2}$ fits an F distribution with non-central parameter $w_i = \frac{\lambda_i \alpha_i^2}{\sigma^2}$, and degrees of freedom 1 and $(n - q)$.

Unless all the weights of γ_i are one or zero in (34), it is quite hard to obtain a closed form of the distribution function of a non-central F_i . When this is the case, approximate results can be obtained [6, 7, 8].

The necessary and sufficient condition for shrinkage estimators mentioned above is $\sum_{i=1}^m \gamma_i w_i \leq 1$. Thus, the hypothesis tests can be written as:

$$(35) \quad H_0 : \sum_{i=1}^m \gamma_i w_i \leq 1 \quad \text{Alternative hypothesis } H_1 : \sum_{i=1}^m \gamma_i w_i > 1.$$

The decision rule for the proposed test is:

$$\text{If } \tilde{F} \leq \tilde{F}_\alpha(m, n - q, 1) \text{ then accept } H_0,$$

and

$$\text{If } \tilde{F} > \tilde{F}_\alpha(m, n - q, 1) \text{ then reject } H_0.$$

Here $\tilde{F}_\alpha(m, n - q, 1)$ is formed from the distribution \tilde{F} , which has non-central param-

eter $w = \sum_{i=1}^m \gamma_i w_i = 1$ and degrees of freedom m and $(n - q)$. Using the central- F approximation for \tilde{F} , described below, the values of $\tilde{F}_\alpha(m, n - q, 1)$ at the critical points are determined. Then the initial moments for the test statistic \tilde{F} are obtained from the moments of the statistics F_i using the method of moments.

The first two central moments of the test statistic \tilde{F} are, from [13],

$$E\tilde{F} = \frac{(n - q)}{(n - q - 2)m} \sum_{i=1}^m \gamma_i (1 + w_i). \quad (n - q > 2)$$

$$E\tilde{F}^2 = \frac{(n - q)^2}{(n - q - 2)(n - q - 4)m^2} \left\{ \left[\sum_{i=1}^m \gamma_i (1 + w_i) \right]^2 + 2 \sum_{i=1}^m \gamma_i^2 (1 + 2w_i) \right\} \quad (n - q > 4)$$

2.3. The Central-F Approximation for Statistics. Patnaik [9], studied a central- F approximation to a non-central distribution F . This is found by using the first two moments of the central- F distribution $F(\vartheta, n - q)$ and the non-central distribution $\tilde{F}_\alpha(m, n - q, w)$, and takes the form

$$\tilde{F}_\alpha(m, n - q, w) \approx rF(\vartheta, n - q).$$

The parameters r and ϑ are found from the first two moments of the distributions F . In other words, the two moment approximation of central- F can be achieved by equating the first two moments of central- F and \tilde{F}/r . Thus we obtain

$$\frac{(n - q)}{(n - q - 2)m} \sum_{i=1}^m \gamma_i (1 + w_i) = \frac{(n - q)}{(n - q - 2)}$$

and

$$\frac{(n - q)^2}{r^2 (n - q - 2)(n - q - 4)m^2} \left\{ \left[\sum_{i=1}^m \gamma_i (1 + w_i) \right]^2 + 2 \sum_{i=1}^m \gamma_i^2 (1 + 2w_i) \right\}$$

$$= \frac{(n - q)^2}{(n - q - 2)(n - q - 4)} \cdot \frac{\vartheta + 2}{\vartheta}.$$

Solving these equations gives

$$(36) \quad r = \frac{1}{m} \sum_{i=1}^m \gamma_i (1 + w_i), \text{ and}$$

$$\vartheta = \frac{\left[\sum_{i=1}^m \gamma_i (1 + w_i) \right]^2}{\sum_{i=1}^m \gamma_i^2 (1 + 2w_i)}$$

When $\gamma = \sum_{i=1}^q \gamma_i$ and $\sum_{i=1}^m \gamma_i w_i = 1$ are given, the measurement factor r can be determined from $r = \frac{\gamma + 1}{m}$. The values of γ_i lie in the interval $[0, 1]$, and the corrected degrees of freedom can be written as

$$\vartheta = \frac{(\gamma + 1)^2}{\sum_{i=1}^m \gamma_i^2 + 2 \sum_{i=1}^m \gamma_i^2 w_i}.$$

From this, it is easily seen that [14]:

$$(37) \quad \gamma_{\min} \leq \sum_{i=1}^m \gamma_i^2 w_i \leq \gamma_{\max}.$$

With the help of the inequality (37), the upper and lower limits of the corrected degrees of freedom ϑ are found to be:

$$(38) \quad \frac{(\gamma + 1)^2}{\sum_{i=1}^m \gamma_i^2 + 2\gamma_{\max}} \leq \vartheta \leq \frac{(\gamma + 1)^2}{\sum_{i=1}^m \gamma_i^2 + 2\gamma_{\min}}.$$

The upper limit is denoted by ϑ_{\max} , and the lower limit by ϑ_{\min} . Hence, for all $0 < \alpha < 1$ we have $F_{\alpha}(\vartheta_{\max}, n - q) \leq F_{\alpha}(\vartheta_{\min}, n - q)$, so we obtain the critical points $F_{\alpha}(\vartheta_{\max}, n - q)$ and $F_{\alpha}(\vartheta_{\min}, n - q)$. The statistic \tilde{F}/r may be compared with these values. Hence, with the help of these critical points, the following regions for the test statistic may be obtained:

Reject H_0 if $\tilde{F}/r > F_{\alpha}(\vartheta_{\min}, n - q)$,

Accept H_0 if $\tilde{F}/r < F_{\alpha}(\vartheta_{\max}, n - q)$,

Inconclusive if $F_{\alpha}(\vartheta_{\max}, n - q) \leq \tilde{F}/r \leq F_{\alpha}(\vartheta_{\min}, n - q)$.

3. A simulation study

In this section we describe a simulation that was carried out using a MATLAB package programme to compare Ridge and Liu estimators - which are shrinkage estimators - with the LS estimator.

Firstly, $n = 50$ samples of the independent variables $X = (\underline{x}_1, \underline{x}_2)$ were chosen from a normal distribution with parameters $\mu = (9, 8)$, $\sigma_1^2 = 9$, $\sigma_2^2 = 9$ and correlation coefficients $\rho = 0, 0.3, 0.6, 0.9$. Then, the error vector ε was chosen from a standard normal distribution and the dependent variable \underline{Y} determined as follows;

$$\underline{Y} = 3.5 + 3\underline{x}_1 + 2.5\underline{x}_2 + \varepsilon$$

In addition, for the Ridge and Liu estimators, $k = 0.01 (0.01) 0.99$ and $d = 0.01 (0.01) 0.99$, respectively, were used and for each of these the values of \tilde{F} , \tilde{F}/r , $F_{\alpha}(\vartheta_{\min}, n - q)$ and $F_{\alpha}(\vartheta_{\max}, n - q)$ were calculated. Finally, the regions of rejection, acceptance and inconclusiveness were found. For the Ridge estimator these are shown in Figures 1-4, and for the Liu estimator in Tables 5-8.

Figure 1. Rejected, accepted and inconclusive regions for Ridge estimators with $\rho = 0$

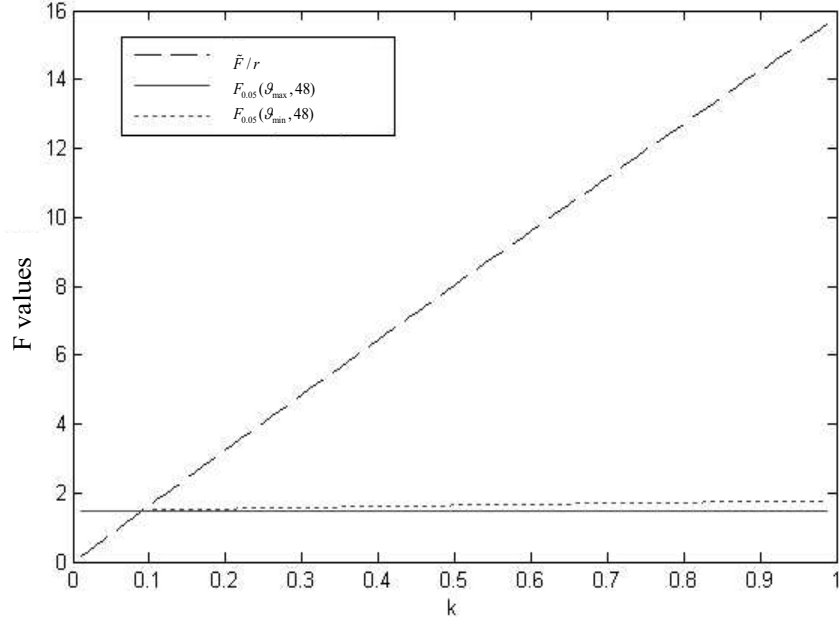


Figure 2. Rejected, accepted and inconclusive regions for Ridge estimators with $\rho = 0.3$

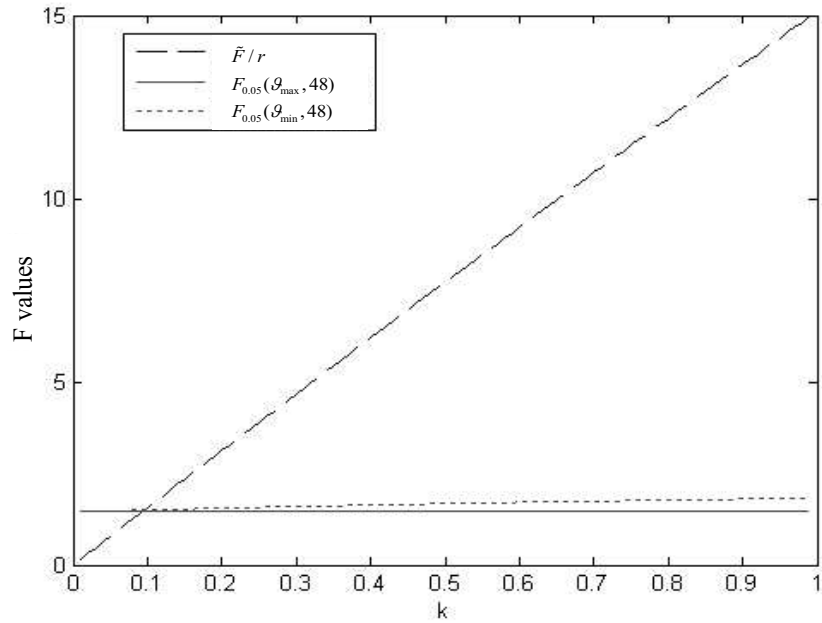


Figure 3. Rejected, accepted and inconclusive regions for Ridge estimators with $\rho = 0.6$

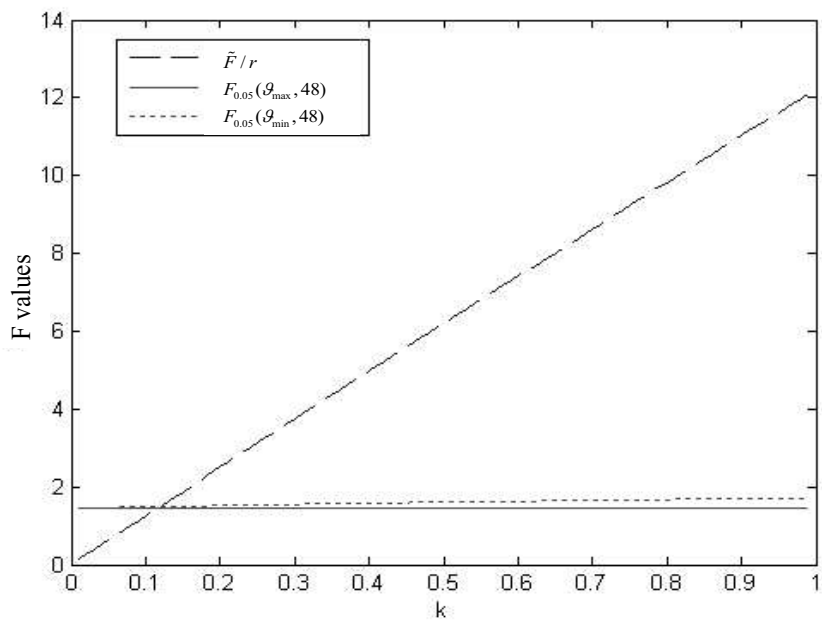


Figure 4. Rejected, accepted and inconclusive regions for Ridge estimators with $\rho = 0.9$

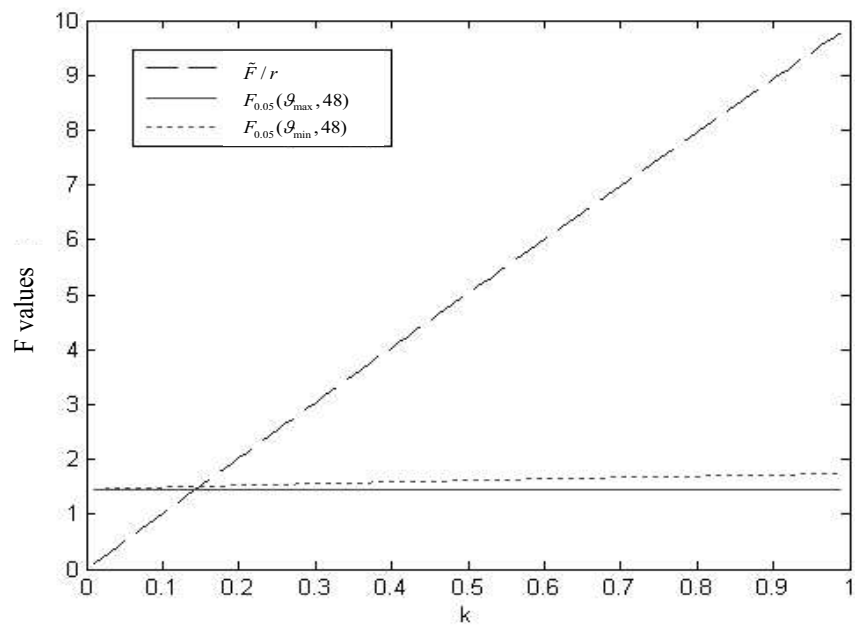


Figure 5. Rejected, accepted and inconclusive regions for Liu estimators with $\rho = 0$

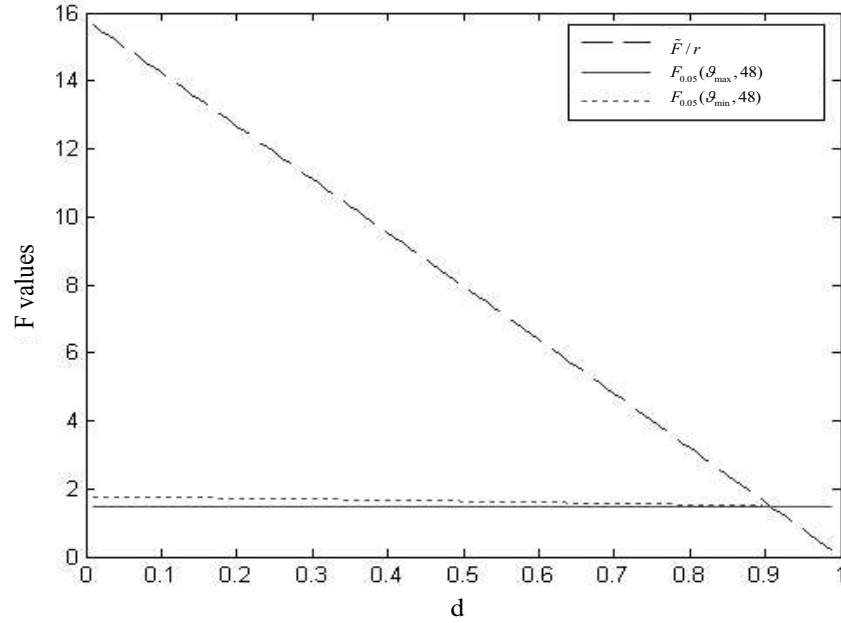


Figure 6. Rejected, accepted and inconclusive regions for Liu estimators with $\rho = 0.3$

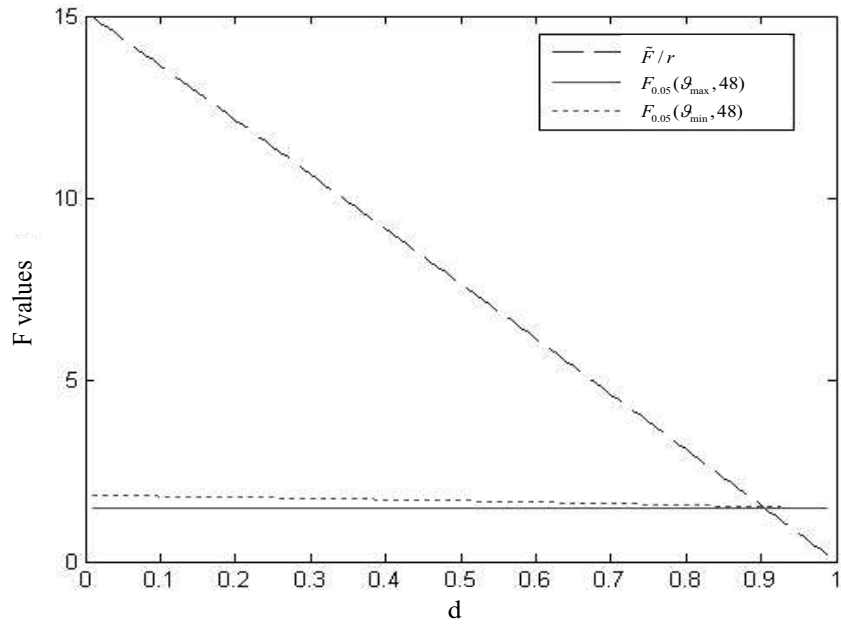


Figure 7. Rejected, accepted and inconclusive regions for Liu estimators with $\rho = 0.6$

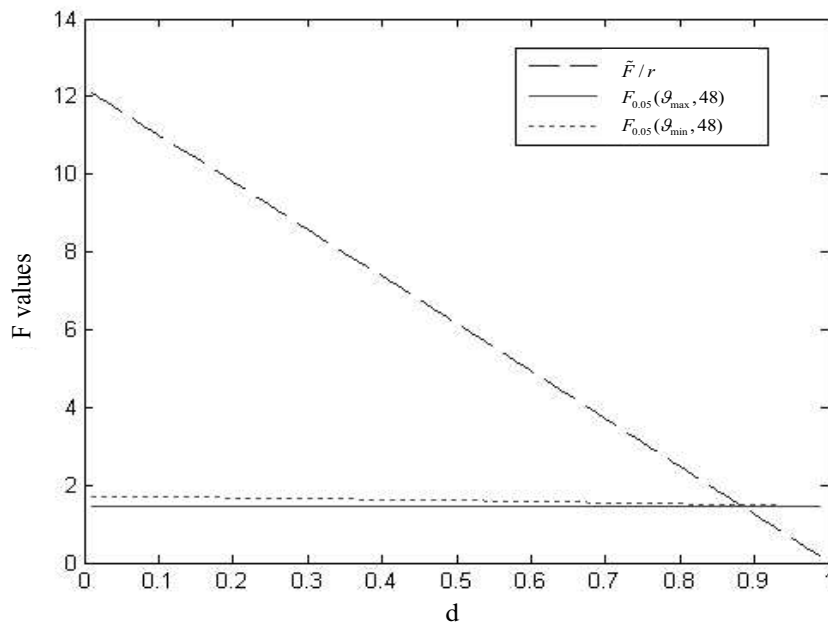
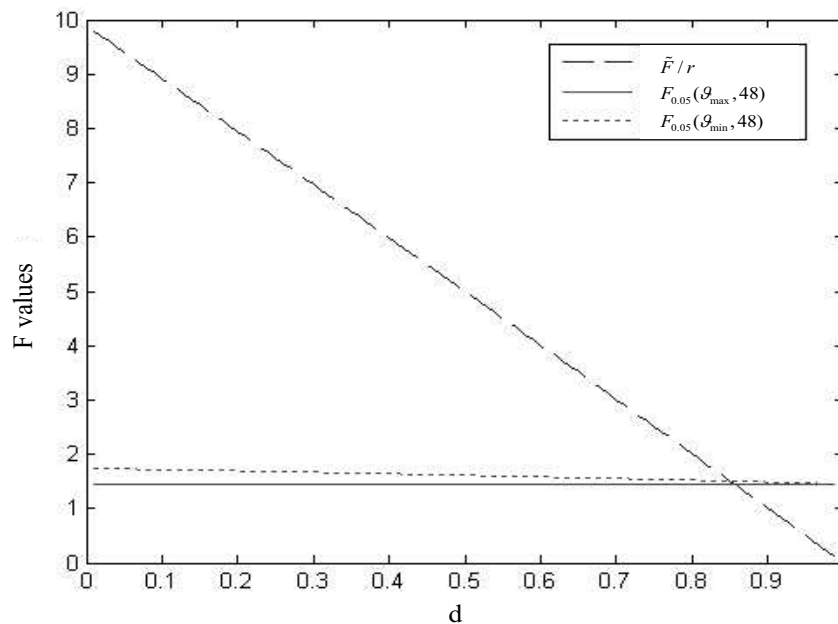


Figure 8. Rejected, accepted and inconclusive regions for Liu estimators with $\rho = 0.9$



4. Conclusion

In Figure 1, for the correlation coefficient $\rho = 0$, the Ridge estimator is seen to give better results than the LS estimator for values of k satisfying $0 < k < 0.09$. On the other hand, as the correlation between the independent variables increases, the interval in which k lies increases also. For example, when Figure 4 is analyzed, it is seen that the Ridge estimator is better than the LS estimator for values of k satisfying $0 < k < 0.15$.

Similarly, when the Liu estimator is examined, it is seen that it gives similar results to the Ridge estimator. Looking at Figure 5 we see that when $\rho = 0$, the Liu estimator gives better results than the LS estimator for d satisfying $0.91 < d < 1$, while when $\rho = 0.9$ as in Figure 8, the range for d increases to $0.85 < d < 1$.

In short, when Figures 1–8 are examined we see that when the correlation increases between the independent variables, the interval over which the Ridge and Liu estimators are preferred to the LS estimator also increases.

In conclusion, Kaçiranlar *et. al.* [4] gave a new biased estimator for β , and illustrated their findings with a numerical example. Akdeniz and Kaçiranlar [2] gave a numerical example with $n = 10$ to obtain optimal k and d values for the Ridge and Liu estimators, respectively. In this paper we also compare these estimators for the general case under different correlation coefficients with the help of a simulation study.

References

- [1] Akdeniz, F. and Erol, H. *Mean squared error matrix comparisons of some biased estimators in linear regression*, Communications in Statistics - Theory and Methods **32**, 2389–2413, 2003.
- [2] Akdeniz, F. and Kaçiranlar, S. *On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE*, Commun. Statist - Theory and Meth. **24** (7), 1789–1797, 1995.
- [3] Farebrother, R. W. *A class of shrinkage estimators*, J. R. Statist. Soc. B **40**, 47–49, 1978.
- [4] Kaçiranlar, S., Sakalhoğlu, S., Akdeniz, F., Styan, G. P. H. and Werner, H. J. *A new biased estimator in linear regression and a detailed analysis of the widely-analyzed dataset on Portland Cement*, Sankhya: The Indian Journal of Statistics Series B **61**, 443–459, 1999.
- [5] Kejian, L. *A new class of biased estimates in linear r, Regression*. Com. Statist. - Theory Meth. **22** (2), 393–402, 1993.
- [6] Liski, E. P. *A test of the mean square error criterion for shrinkage estimators*, Communications in Statistics **11** (5), 543–562, 1982.
- [7] Liski, E. P. *Choosing a shrinkage estimator - a test of the mean square error criterion*, Proc. First Tampere Sem. Linear Models, 245–262, 1983.
- [8] Norman, L. J. and Samuel, K. *Distributions in Statistics (Chapter 26)*, 1970.
- [9] Patnaik, P. B. (1949), *The non-central Ki-kare and F Distributions and their applications*, Biometrika **36**, 202–232, 1949.
- [10] Rao, C. R. *Estimation of parameters in a linear model*, The Annals of Statistics **4**, 1023–1037, 1976.
- [11] Sakalhoğlu, S., Kaçiranlar, S. and Akdeniz, F. *Bazı Yanlı Regresyon Tahmin Edicilerinin Karşılaştırılması.*, Araştırma Sempozyumu'97 Bildirileri, DİE., Ankara, 1997.
- [12] Searle S. R. *Matrix Algebra useful for Statistic* (John Wiley and Sons., New York, 1982).
- [13] Theobald, C. M. *Generalizations of mean square error applied to Ridge regression*, J. R. Statist. Soc. B **36**, 103–106, 1974.
- [14] Toutenburg, H. *Prior Information in Linear Models* (John Wiley and Sons, New York, 1982).