Stochastic and Deterministic Study of Ridge Regression

Noha E. Abd Al-Rahman[1], Salwa A. Mousa[2] and Amina E. Abo-Hussien[3]

[1] Department of Statistics, Faculty of Commerce, Al-Azhar University, Girls' Branch, Dakahlia , Egypt.

[2,3] Department of Statistics, Faculty of Commerce, Al-Azhar University, Girls' Branch, Nasr city, Cairo, Egypt.

The present paper is concerned with studying the ridge parameters k through deterministic and stochastic approach in the case of ordinary ridge regression (ORR). In the deterministic approach, some new formulas for the ridge parameters are proposed and compared with the formula suggested by Hoerl and Kennard (1970a). The performance of the proposed ridge parameters is evaluated through a simulation study in the presence of multicollinearity. An application using real data is given. The evaluation is based on the mean square error (MSE) and relative MSE (RMSE). In the stochastic approach, the main properties of the proposed new formulas and the formula suggested by Hoerl and Kennard (1970a) are studied. The probability density function (pdf) and distribution function of the formulas are derived. The empirical distributions of these formulas are derived using Pearson's method. It is found that the performance of the new formulas are better than ordinary least squares (OLS) and the formula suggested by Hoerl and Kennard (1970 a) for all selected distributions.

***Keywords*:** Multicollinearity; Ridge Regression; Deterministic approach; Stochastic approach; Monte Carlo simulation; Mean square error; Relative mean square error; Sampling distribution.

1. Introduction

In multiple regression analysis, it is usually assumed that the explanatory variables are independent, but in most applications there is a correlation among the explanatory variables which is called multicollinearity. In the presence of multicollinearity OLS regression produce estimates having a large MSE.

Hoerl (1962) introduced the ridge regression (RR) estimators as an alternative to OLS estimators in the presence of multicollinearity by adding a small value k to the diagonal elements of the correlation matrix (k is a positive quantity less than one). Much of the discussions on RR is concerned with the problem of finding good empirical value of k. In the literature, various methods are suggested to choose the ridge parameter k in ridge regression in both situations of ORR and general ridge regression (GRR) that allows separate ridge parameter for each regressor.

A general discussion of these methods are given in Hoerl and Kennard (1970 a), Hoerl and Kennard (1970 b), Marquardt (1970), Marquardt and Snee (1975), Hoerl *et al* (1975), Lawless and Wang (1976), Golub *et al* (1979), Firinguetti and Rubio (2000), Rubio and Firinguetti (2002), Kibria (2003), Khalaf and Shukur (2005), Alkhamisi *et al* (2006), Batah *et al* (2009), Muniz and Kibria (2009), Al-Hassan and Yazid (2010), Dorugade and Kashid (2010), Mansson *et al* (2010), Abd El-Salam (2011), El-Dereny and Rashwan (2011), Abd- Eledum and Alkhalifa (2012), Muniz *et al* (2012), Khalaf et al (2013), Dorugade (2014), Dorugade (2014), Abd El-Salam (2015) and Khalaf and Iguernane (2016).

The present study is concerned with studying the ridge parameters k through deterministic and stochastic approach in the case of ORR.

Researchers use OLS method to estimate the parameters of the regression model. The multiple linear regression model is the common and it is formulated as follows:

$$Y = X\beta + \epsilon \quad , \qquad (1)$$

where: Y is an $(n \times 1)$ vector of responses variable, X is an $(n \times p)$ design matrix of explanatory variables ,

β is a $(p \times 1)$ vector of unknown parameters and
$\boldsymbol{\epsilon}$ is an $(n \times 1)$ vector of random errors.

The estimators of the OLS method is best linear unbiased estimators. The OLS estimators are given by:

$$\hat{\beta}_{OLS} = (X^{\prime}X)^{-1}X^{\prime}Y , \qquad (2)$$

where: $X^{\prime}$ is the transpose of the matrix X.

One of the assumptions of the OLS method is the independence between the explanatory variables, however if they are not independent this means that a problem of multicollinearity exists. The main properties of $\hat{\beta}$ are unbiased and have minimum variance. When the cross-product matrix, $X^{\prime}X$ is ill-conditioned, the OLS estimatse of β has a large variance and multicollinearity is said to be present. Since MSE is equal to the variance plus $(\text{bias})^2$, variance and bias have the same effect on MSE.

One of several methods that have been proposed to remedy multicollinearity problems is RR by modifying the method of least squares to allow biased estimators with smaller variance of the regression coefficients, than the OLS estimators. The method for estimating β with smaller variance, but with some bias, than the OLS estimator have been studied by many authors. Hoerl and Kennard (1970a) suggested a small positive number to be added to the diagonal elements of $X^{\prime}X$ which is the use of $(X^{\prime}X + kI)^{-1}$, $k \geq 0$, k is the ridge parameter , I is the identity matrix , and this method is called RR .

The first step in ridge regression is to standardize the explanatory and response variables. The OLS estimators to the standardized model is:

$$\widehat{\beta}^*{}_{OLS} = (X^{*'}X^*)^{-1}X^{*'}Y^* \qquad (3)$$

The ridge standardized regression estimators are obtained using the following formula:

$$\widehat{\beta}^*{}_{RR} = (X^{*'}X^* + kI)^{-1} \quad X^{*'}Y^* \;, \qquad (4)$$

where :$\widehat{\beta^*}_{RR}$ is the vector of the standardized ridge regression coefficients and X^* , Y^* are standard values.

The objective of the present paper is to study the ridge parameter k through stochastic and deterministic approach. In the deterministic approach modified formulas for the ridge regression parameters k are proposed, the performance of the ridge regression parameters k of the formula suggested by Hoerl and Kennard (1970a) and the proposed modified formulas are studied through a simulation study. In the stochastic approach the density and distribution functions for the formulas of RR parameters are derived and the empirical distributions of these formulas are derived using Pearson's method. An application using real data is given. The deterministic approach is introduced in section (2). The stochastic approach is introduced in section (3). The results and dissection are discussed in section (4).

2. The Deterministic Approach

Some new formulas for the ridge parameter are proposed and compared with some of the existing formulas in the literature. The performance of the proposed ridge parameters is evaluated through a simulation study in the presence of multicollinearity. The evaluation is based on the MSE and RMSE. The proposed new formulas of the ridge parameter k and the existing formula used in the comparison are introduced in sub- section (2.1). The simulation study is illustrated in sub-section (2.2). Results and discussion of the simulation study are given in sub-section (2.3). Application using real data is given in sub-section (2.4).

2.1 The proposed new formulas of the ridge parameter

Hoerl and Kennard (1970a) were the first researchers who have solved the problem of multicollinearity by ridge regression and suggested the following formula for the ridge parameter $\boldsymbol{k}$:

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\beta}^2_{max}} \quad , \tag{5}$$

where: $\hat{\sigma}^2 = \sum_{i=1}^{n} u_i^2 \ / (n-p-1)$, $(i = 1, \dots, n)$

$\boldsymbol{p}$: is the number of the explanatory variables and $\boldsymbol{u_i}$ are the residuals obtained from the OLS regression.

$\boldsymbol{\hat{\beta}^2_{max}}$:is the maximum element of the estimators of OLS $(\boldsymbol{\hat{\beta}^*_{OLS}})$.

The proposed new formulas are derived based on the formula suggested by Hoerl and Kennard (1970a). The suggested estimators are as follows:

$$\hat{k}_{N1} = [\frac{\hat{\sigma}^2}{\hat{\beta}^2_{max}}]^{\frac{1}{P}} \tag{6}$$

$$\hat{k}_{N2} = [\frac{\hat{\sigma}^2}{\hat{\beta}_{max}^2}]^{-1} \quad (7)$$

A simulation study is conducted to compare the performance of the proposed new formulas with the formula suggested by Hoerl and Kenard (1970 a) and OLS.

2.2 The simulation study

The objective of this sub-section is to investigate the performance of the ridge regression estimator $\hat{\boldsymbol{k}}_{\boldsymbol{HK}}$ and the proposed ridge regression estimators $\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$ and $\hat{\boldsymbol{k}}_{\boldsymbol{N2}}$ against OLS under several degrees of multicollinearity using Monte Carlo simulation. The performance of these estimators can be evaluated using the MSE and relative MSE (RMSE) of the estimated regression coefficients which are given by:

$$MSE_1(OLS) = E(\hat{\beta}^*_{OLS} - \hat{\beta})'(\hat{\beta}^*_{OLS} - \hat{\beta}) = \frac{1}{r}\sum_{i=0}^{r}(\hat{\beta}^*_{i_{OLS}} - \hat{\beta}_i)^2, \quad (8)$$

$$MSE_1(RR) = E(\hat{\beta}^*_{RR} - \hat{\beta})'(\hat{\beta}^*_{RR} - \hat{\beta}) = \frac{1}{r}\sum_{i=0}^{r}(\hat{\beta}^*_{i_{RR}} - \hat{\beta}_i)^2. \quad (9)$$

where: r is the number of replications .

$$\text{RMSE}_1 = \frac{\text{MSE}_1(\text{RR})}{\text{MSE}_1(\text{OLS})}. \quad (10)$$

The following factors are investigated in the simulation study:

i. Sample size (n):

Different sample sizes are taken (n=10, 30, 50, 75 and 100) to study the effect of small, moderate and large samples on the properties of the estimators.

ii. The correlation among explanatory variables (ρ):

To study the properties of the estimators in the presence of multicollinearity ρ is taken as $\rho = (0.2, 0.7, 0.8 \text{ and } 0.9)$.

iii. Distribution of random errors (e):

To study the effect of the distribution of random errors on the estimators, different distributions of random error are considered: ($N(0,1)$, ${\chi^2}_{(1)}$, $T_{(4)}$ and $F(2,10)$).

iv. Number of explanatory variables (p) :

Different values of the explanatory variables are taken (p= 3, 4 and 5).

Steps of the simulation study using R programming:

The computation of the simulation study is developed using R program (version 3.2.2 and package MASS). Some functions in R program such as glm and ridge packages are used to compare the performance of different ridge estimators under different sample sizes, different distributions of random errors, different values of the explanatory variables and strength of correlation in the following steps:

1) Generate the values of $\boldsymbol{Z}$ from the standard normal distribution.
2) Obtain the values of $\boldsymbol{x}$ using the equation:

$$x_{ij} = (1-\rho^2)^{\frac{1}{2}} z_{ij} + \rho\, z_{ip} \,, i = 1,2,\dots,n \,, j = 1,2,\dots,p \quad (11)$$

where :$\boldsymbol{z_{ij}}$ are independent standard normal and $\boldsymbol{\rho}$ is the correlation among explanatory variables

[See Gibbons (1981)].

3) The initial values of the coefficients are chosen such that $\beta_0 = 0$ and $\sum \beta_i^2 = 1$ which is a common restriction in simulation studies [See Hoerl *et al* (1975), Kibria (2003)].
4) The other factors [the sample size (n), the degree of the correlation among regressors ($\boldsymbol{\rho}$), the distribution of error (e_i), the number of the explanatory variables (p) and (σ^2)] are given for each replicate (r =1000).
5) Generate the values of error (e_i) using the selected distribution.
6) Generate the response variables $\boldsymbol{y}$ using the following equation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i \,, i = 1,2,\dots,n \,. \quad (12)$$

7) Obtain the OLS estimators, $\hat{\beta}^*_{OLS}$, using equation (3) .
8) Find the RR estimators to the proposed new formulas and the formula suggested by Horel and Kenard in (1970a) using equation (4).
9) Repeat steps (1-9), r =1000 times.

10) Calculate the MSE to OLS estimators, using equation (8).
11) Obtain the MSE to RR estimators using equation (9).
12) Calculate the relative MSE using equation (10).

Simulation results are summarized in tables (2.1) – (2.12) in Appendix A.

2.3 Results and discussion of the simulation study

The main results of the simulation study are as follows:

- It is found that formulas $\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$ and $\hat{\boldsymbol{k}}_{\boldsymbol{N2}}$ are better than OLS and $\hat{\boldsymbol{k}}_{\boldsymbol{HK}}$ for all selected distributions and all values of ρ , p and n.
- It is found that formula $\hat{\boldsymbol{k}}_{\boldsymbol{N2}}$ is better than $\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$ for all distributions and all values of ρ , p and for large values of n whereas $\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$ is the best when the sample size is too small.
- It is noticed that as ρ and p increases, RR gives better results than OLS.
- As n increases the MSE decreases for all formulas which agree with the theoretical results.

- It is found that for small values of ρ , the least square method is better than RR which agree with the theoretical studies.

- All factors chosen to vary in the design of the experiment affects the estimated MSE. As expected, it is noticed that increasing the degree of correlation leads to a higher estimated MSE. This increase is much greater for OLS than for the RR estimators.
- As n and p increase, the RMSE of the proposed ridge parameters $\hat{\boldsymbol{k}}_{N1}, \hat{\boldsymbol{k}}_{N2}$ are better than $\hat{\boldsymbol{k}}_{HK}$ **,**for different selected distributions.
- As ρ increases the RMSE decreases for all formulas in the different selected distributions.

2.4 Application Using Real Data

The Household Income, Expenditure and Consumption Survey (HIECS) carried out in Egypt in 2012- 2013 is used to select a sample of 20 households from Cairo Governorate. It is desired to examine the relationship between income and spending for the family according to some variables such as housing conditions and characteristics of the head of the family.

In the present application, the response variable is taken to be the net annual household income in pounds Y, and eight explanatory variables are selected as follows: X_1:Total annual household expenditure in pounds, X_2: Household expenditure on housing in pounds, X_3: Housing space, X_4 : Number of rooms, X_5 : Number of household members, X_6 : Total expenses on garments, X_7:Number of working days for household head in a weak,

X_8: Average number of working hours for household head in a day.

The multiple linear regression model is as follows:

$$Y = \sum_{i=1}^{8} \beta_i X_i + e_i \quad , i = 1,2,\dots,8 \qquad (13)$$

Table (2.13) describes the data used in the application .To study this relationship using OLS method, there is a problem because it is expected that some of the independent variables are correlated, this problem is called multicollinearity.

Diagnostics of multicollinearity is investigated for the data using correlation matrix, methods based on the eigenvalues, condition number and the variance inflation factor (VIF). The investigation reveals that:

The explanatory variables are moderate to highly correlated, at least one eigenvalue of the explanatory variables is close to zero, the condition number is calculated as 747.98, the VIF = 10.38 . The parameters of the ridge regression are computed using equations (5), (6) and (7) as follows:

$$\hat{k}_{HK} = 0.179176, \hat{k}_{N1} = 0.806603, \quad \hat{k}_{N2} = 5.581117$$

Table (2.13): The distribution of 20 households according to income and the selected variables on housing conditions and characteristics of the head of the family in Cairo in 2012-2013.

Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
25600	24442	856	90	3	4	310	0	0
68900	74535	894	84	3	5	715	5	8
21200	24319	312	65	3	4	835	6	8
33600	36348	585	65	3	5	425	0	0
30624	31606	530	80	3	5	1090	6	8
130120	126311	1955	120	4	4	1860	5	8
31200	29809	580	48	3	4	1405	5	8
33790	18855	653	60	3	4	860	0	0
18808	19048	436	50	2	2	330	0	0
33192	26970	618	60	3	5	1695	5	8
30735	30570	456	68	2	7	1480	7	9
35700	26414	526	30	2	5	991	6	9
23600	20401	806	70	4	3	707	7	10
20400	19391	1090	75	4	5	600	6	8
24000	21619	654	55	4	3	420	0	0
18800	17959	589	50	3	2	390	0	0
10000	8084	123	55	3	1	34	0	0
52860	27690	467	75	3	5	1123	5	7
515000	170440	1849	90	4	4	3000	6	10
11363	8277	402	40	2	3	200	6	10

Source: Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt, Arab Rep. (HIECS) 2012-2013.

Hoerl and Kenard (1970 a,b) proposed the following formulas to the MSE to assess the performance of OLS, RR estimators , this formula is given by:

$$\text{MSE}_2(\text{RR}) = \sigma^2 \sum_{i=1}^{r} \frac{\delta_i}{(\delta_i+\widehat{kI})^2} + \sum_{i=1}^{r} \frac{\widehat{kI}^2\, \widehat{\beta_{i}}_{RR}^{*\,2}}{(\delta_i+\widehat{kI})^2}, \quad (14)$$

$MSE_2(OLS)$ can be obtained using equation (14) when k=0 as follows:

$$MSE_2(OLS) = \boldsymbol{\sigma^2} \sum_{\boldsymbol{i=1}}^{r} \frac{\mathbf{1}}{\delta_i}, \quad (15)$$

where: δ_i are the eigen values of $X^{`}X$.

$$RMSE_2 = \frac{MSE_2(RR)}{MSE_2(OLS)}. \quad (16)$$

Using equations (14), (15) and (16), the MSE and the RMSE of $\hat{k}_{HK}$,$\hat{k}_{N1}$,$\hat{k}_{N2}$ and OLS are calculated as follows:

$\hat{\boldsymbol{k}}$'s	MSE	RMSE
$\hat{\boldsymbol{k}}_{\boldsymbol{HK}}$	0.6503	0.2994
$\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$	**0.5243**	**0.2414**
$\hat{\boldsymbol{k}}_{\boldsymbol{N2}}$	0.6505	0.2995
OLS	2.1714	

Notice that ridge regression estimates are better than OLS estimates and $\hat{\boldsymbol{k}}_{\boldsymbol{N1}}$ is the best formula.

3. The Stochastic Approach

The present section is concerned with studying the main properties of the formulas suggested in section 2 and the formula proposed by Hoerl and Kennard (1970 a). The probability density function (pdf) of the formulas are derived and in addition, the empirical distributions of these formulas are derived using Pearson's method. The probability density and distribution functions of RR parameters are derived in sub-section (3.1).Empirical sampling distributions of ridge parameters are introduced in sub-section (3.2).

3.1 The Probability Density and Distribution Functions of RR Parameters

In this sub-section, the pdf and the distribution function of the formula suggested by Hoerl and Kennard (1970a), $\hat{k}_{HK}$, and the pdf and the distribution functions of the new proposed formulas $\hat{k}_{N1}$, $\hat{k}_{N2}$ are derived.

The probability density function and the distribution function of $\hat{\mathbf{k}}_{\mathbf{HK}}$

To derive the density function of $\hat{k}_{HK}$ under the assumption of the linear regression model in equation (1), and the normality condition can be written as:

$$\hat{k}_{HK} = \frac{\lambda}{(n-p)} \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \frac{\sigma^2}{\lambda\, \hat{\alpha}^2_{max}} \;, \; i = 1,2, \dots, p \qquad (17)$$

where: p is the number of the explanatory variables, n is the number of observations, λ is the shrinkage parameter.

$$\text{Put } u = \frac{(n-p)\hat{\sigma}^2}{\sigma^2}, \quad u \sim \chi^2_{(n-p)}(0) \tag{18}$$

$\boldsymbol{u}$ is the central Chi-square distribution with $(\boldsymbol{n}-\boldsymbol{p})$ degrees of freedom ,

$$\boldsymbol{v} = \frac{\lambda\, \hat{\boldsymbol{\alpha}}^2_{\boldsymbol{max}}}{\boldsymbol{\sigma}^2} \tag{19}$$

$$\boldsymbol{v} \sim \chi^2_{(1)}(\theta)\ ,$$

where:

$$\theta = \frac{\hat{\alpha}^2_{\max}\lambda}{\sigma^2}\ ,$$

$\boldsymbol{v}$ is the non-central Chi-square distribution with one degree of freedom and non-central parameter θ. Since α_{max} ,σ^2 are independent and normally distributed, then

$$\hat{\alpha}_{\max} \sim N(\alpha_{\max}\ , \sigma^2\lambda^{-1})\ ,$$

Substitute (18) and (19) in (17), then

$$\hat{k}_{HK} = \frac{\lambda}{(n-p)}\frac{u}{v} = \frac{\lambda}{y} \tag{20}$$

where:

$$\boldsymbol{y} = \frac{v\,(n-p)}{u}$$

[See Johnson and Kotz (1970)].

The probability density function

The pdf of the ridge parameter suggested by Hoerl and Kennard (1970 a), $\boldsymbol{\hat{k}_{HK}}$, is derived as follows:

$$f(\hat{k}_{HK})=\boldsymbol{e}^{-\frac{\theta}{2}}\sum_{j=0}^{\infty}\left[\frac{(\theta/2)^{j}}{j!}\right]\cdot\left[\frac{(\lambda/(n-p))^{j+\frac{1}{2}}\left(\frac{1}{\hat{k}_{HK}}\right)^{j+\frac{3}{2}}}{\left(1+\frac{\lambda}{(n-p)}\left(\frac{1}{\hat{k}_{HK}}\right)\right)^{\frac{1+(n-p)}{2}+j}}\right]\left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right], \quad (21)$$

$$\hat{k}_{HK}>0 \quad .$$

[See Appendix B]

The distribution function

The distribution function of $\boldsymbol{\hat{k}_{HK}}$ is derived as follows:

$$F(X_i)=1-\boldsymbol{e}^{-\frac{\theta}{2}}\sum_{j=0}^{\infty}\left[\frac{(\theta/2)^{j}}{j!}\right]\cdot\left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right]\cdot\ \boldsymbol{\beta_r}\left(\frac{1}{2}+j\,,\frac{(n-p)}{2}\right). \quad (22)$$

where:

β_r is the incomplete beta $\int_0^{\frac{\lambda}{X_i(n-p)+\lambda}}(Z)^{\,j+\frac{1}{2}-1}(1-Z)^{\frac{(n-p)}{2}-1}\ dz.$

[See Appendix B]

The probability density function and the distribution function of $\hat{\mathbf{k}}_{\mathbf{N1}}$

To drive the pdf of the first formula suggested as a modification of the formula of Hoerl and Kennard (1970a) , $\hat{\mathbf{k}}_{\mathbf{N1}}$, the assumption of the linear regression model in equation (1), and the normality condition can be written as:

$$\hat{k}_{N1} = \left[\frac{\lambda}{(n-p)} \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \frac{\sigma^2}{\lambda\, \hat{\alpha}^2_{max}}\right]^{\frac{1}{p}} \qquad (23)$$

Substitute (18) and (19) in (23) then

$$\hat{\mathrm{k}}_{\mathrm{N1}} = \left[\frac{\lambda}{(\mathrm{n-p})} \frac{u}{v}\right]^{\frac{1}{\mathrm{p}}} = \left[\frac{\lambda}{\mathrm{y}}\right]^{\frac{1}{\mathrm{p}}} \qquad (24)$$

The probability density function

To drive the pdf of the ridge parameter suggested as a modification of the formula of Hoerl and Kennard (1970a) , $\hat{\mathbf{k}}_{\mathbf{N1}}$,is derived as follows: :

$$f\left(\hat{k}_{N1}\right)=e^{-\frac{\theta}{2}}\sum_{j=0}^{\infty}\left[\frac{(\theta/2)^j}{j!}\right]\left[\frac{p\,(\lambda/(n-p))^{j+\frac{1}{2}}\left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]^p}\right]^{j+\frac{1}{2}}\left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]}\right]}{\left(1+\frac{\lambda}{(n-p)}\left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]^p}\right]\right)^{\frac{1+(n-p)}{2}+j}}\right]\left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right], \qquad (25)$$

$$\hat{\mathrm{k}}_{\mathrm{N1}} > 0 \ .$$

[See Appendix B]

The distribution function :

The distribution function of $\hat{\mathbf{k}}_{\mathbf{N1}}$ is derived as follows:

$$F(X_i) = 1 - e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right] . \ \beta_q\left(\frac{1}{2}+j\,,\frac{(n-p)}{2}\right). \quad (26)$$

where:

β_q is the incomplete beta $\int_{\mathbf{0}}^{\frac{\lambda}{[X_i]^p(n-p)+\lambda}} (\boldsymbol{Z})^{\boldsymbol{j}+\frac{1}{2}-\mathbf{1}} (\mathbf{1}-\boldsymbol{Z})^{\frac{(\boldsymbol{n}-\boldsymbol{p})}{2}-\mathbf{1}} \boldsymbol{dz}.$ [See Appendix B]

The probability density function and the distribution function of $\hat{\mathbf{k}}_{\mathbf{N2}}$

To drive the pdf of the second formula suggested as a modification of the formula of Hoerl and Kennard (1970a) ,$\widehat{\mathbf{k}}_{\mathbf{N2}}$, the assumption of the linear regression model in equation (1), and the normality condition can be written as:

$$\hat{k}_{N2} = \left[\frac{\lambda}{y}\right]^{-1}, \quad (27)$$

The probability density function

The pdf of the ridge parameter suggested as a modification of the formula of Hoerl and Kennard (1970a) ,$\widehat{\mathbf{k}}_{\mathbf{N2}}$, is derived as follows:

$$f(\hat{\mathrm{k}}_{\mathrm{N2}}) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{(\lambda/(n-p))^{\frac{1}{2}+j} \left[[\hat{\mathrm{k}}_{\mathrm{N2}}]\right]^{j+\frac{1}{2}-1}}{\left(1+\frac{\lambda}{(n-p)}\left[[\hat{\mathrm{k}}_{\mathrm{N2}}]\right]\right)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right], \quad (28)$$

$$\hat{\mathrm{k}}_{\mathrm{N2}} > 0 \ .$$

[See Appendix B]

The distribution function:

The distribution function of $\hat{\boldsymbol{k}}_{\boldsymbol{N2}}$ is derived as :

$$F(X_i) = 1 - e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] . \beta_w \left(j + \frac{1}{2}, \frac{(n-p)}{2}\right). \quad (29)$$

where:

β_w is the incomplete beta $\int_{\mathbf{0}}^{\frac{\lambda[X_i]}{(n-p)+\lambda[X_i]}} (\boldsymbol{Z})^{j+\frac{1}{2}-1} (\mathbf{1} - \boldsymbol{Z})^{\frac{(n-p)}{2}-1} \boldsymbol{dz}.$

[See Appendix B]

3.2 Empirical Sampling Distribution of Ridge Parameters:

In this sub-section, the empirical sampling distribution of the ridge parameters suggested by Horel and Kennard (1970a), $\hat{\mathbf{k}}_{\mathbf{HK}}$ and the new proposed formulas $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ are derived using Person's system approach.

Pearson's system approach

The selected approach is based on computing the following criterion (D) which is a function of the first four central moments to determine the distribution family:

$$D = \frac{\beta_1(\beta_2+3)^2}{4(4\beta_2-3\beta_1)(2\beta_2-3\beta_1-6)} \quad , \qquad (30)$$

where: The two moment ratios $\beta_1 = \mu_3{}^2 / \mu_2{}^3$, $\beta_2 = \mu_4 / \mu_2{}^2$, denote the skewness and kurtosis measures respectively , μ_r is the rth central moments.

Pearson's system approach can be summarized in the following steps:

1. Estimate the first four central moments from the resulting 1000 replication of the ridge parameters suggested by Hoerl and Kennard (1970 a) ($\hat{k}_{HK}$) and the proposed new formulas $\hat{k}_{N1}$ and $\hat{k}_{N2}$.

2. Use the central moment estimates to compute β_1, β_2 and D for each of $\hat{k}_{HK}$ and the proposed ridge parameters $\hat{k}_{N1}$ and $\hat{k}_{N2}$.

3. Select the appropriate distribution from Pearson's family according to the values of β_1 , β_2 and D

The investigation of the simulation results, two Pearson distribution were fitted to the $\hat{k}_{HK}$, $\hat{k}_{N1}$ and $\hat{k}_{N2}$ using Pearson's approach reveals that:

1- $\hat{k}_{HK}$ and $\hat{k}_{N2}$ may be well described using Pearson's type I system of frequency curves given by:

$$y = y_0 \left[1 + \frac{x}{a_1}\right]^{m_1} \left[1 - \frac{x}{a_2}\right]^{m_2}, (-a_1 < x < a_2), \quad (31)$$

where: $\frac{m_1}{a_1} = \frac{m_2}{a_2}$,

$$a_1 + a_2 = \frac{1}{2}\sqrt{\mu_2}\sqrt{\{\, \beta_1(r+2)^2 + 16(r+1)\ \}}, \; r = \frac{6(\beta_2 - \beta_1 - 1)}{(6 + 3\,\beta_1 - 2\,\beta_2)}$$

2- $\hat{k}_{N1}$ may be well described using Person's type VI system of frequency curves given by:

$$y = y_0(x-a)^{q_2}x^{-q_1}, a < x < \infty \quad (32)$$

where: $a = a_1 + a_2$,q_1 and q_2 are given by

$$\frac{r-2}{2} \pm \frac{2(r-2)}{2}\sqrt{\frac{\beta_1}{\{\beta_1(r+2)^2+16(r+1)\}}} .$$

[See Elderton and Johnson (1969)].

The graphical representation:

A simulation study [discussed in (2.2)] is conducted to generate the sampling distributions of $\hat{\mathbf{k}}_{\mathbf{HK}}$, $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ for different sample sizes [$n = 10, 30, 50, 75, 100$], different values for the number of explanatory variables [P = 3,4,5] , different values for the correlation among regressors [$\rho = 0.2, 0.7, 0.8, 0.9$], different assumptions for the distribution of error (e_i)~[$N(0,1), F(2,10), T_{(4)}$, $\chi^2_{(1)}$]and repetitions 1000.

The sampling distributions are illustrated in Figures (3.1) to (3.4) of the values [$n = 100, p = 5, \rho = 0.9$] .

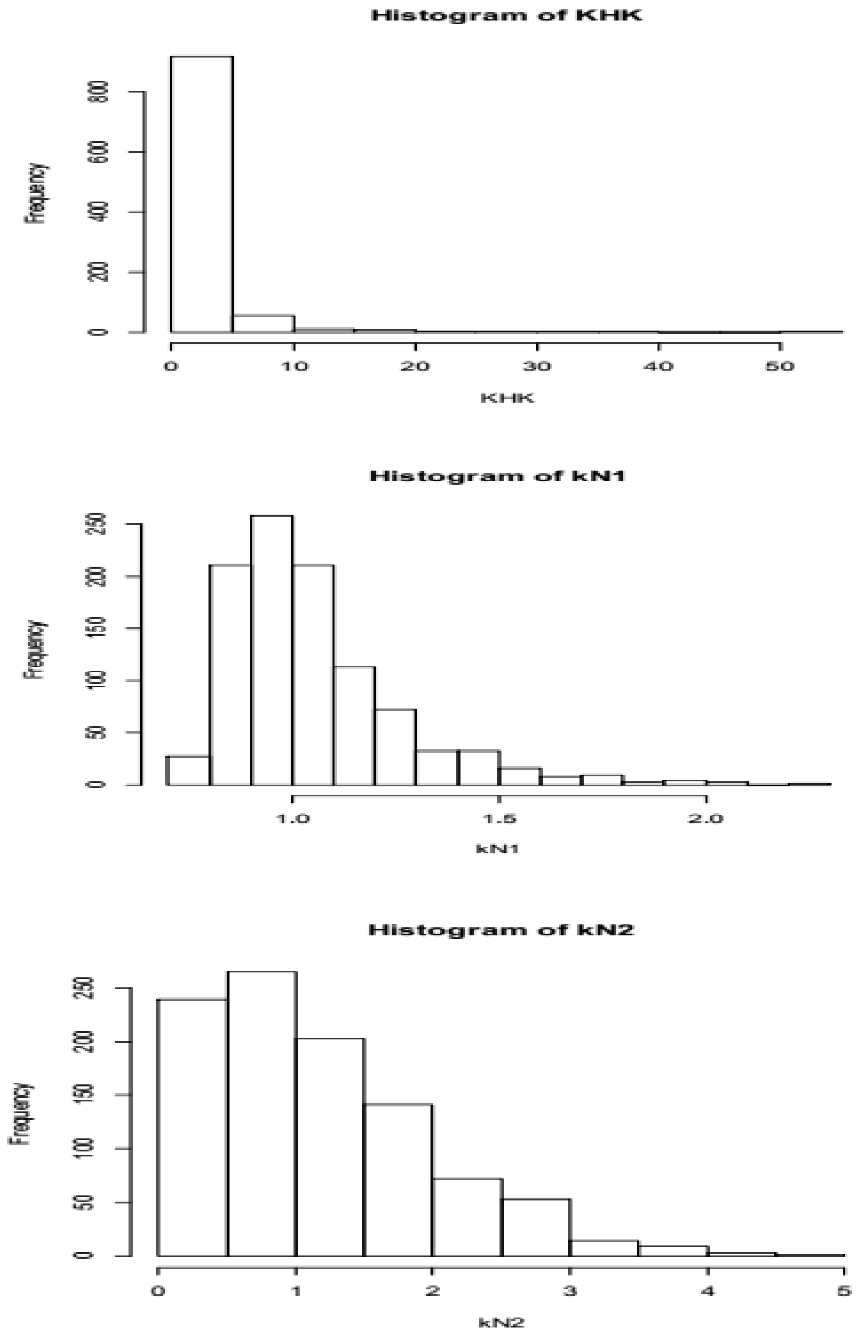

Figure (3.1) the histogram of the sampling distribution of $\hat{\mathbf{k}}_{\mathbf{HK}}$, $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ when e ~N(0,1).

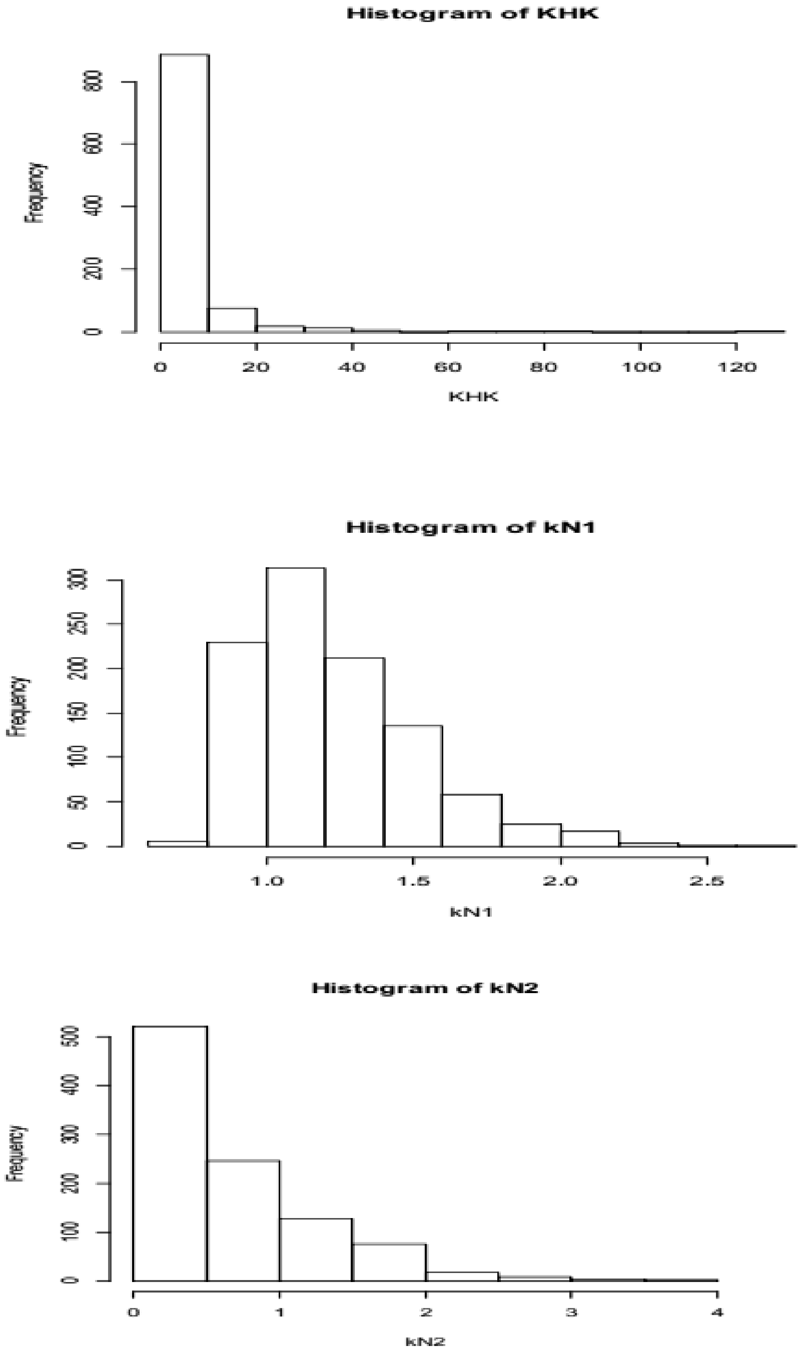

Figure (3.2) the histogram of the sampling distribution of $\hat{\mathbf{k}}_{\mathbf{HK}}$, $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ when $e \sim F(2,10)$.

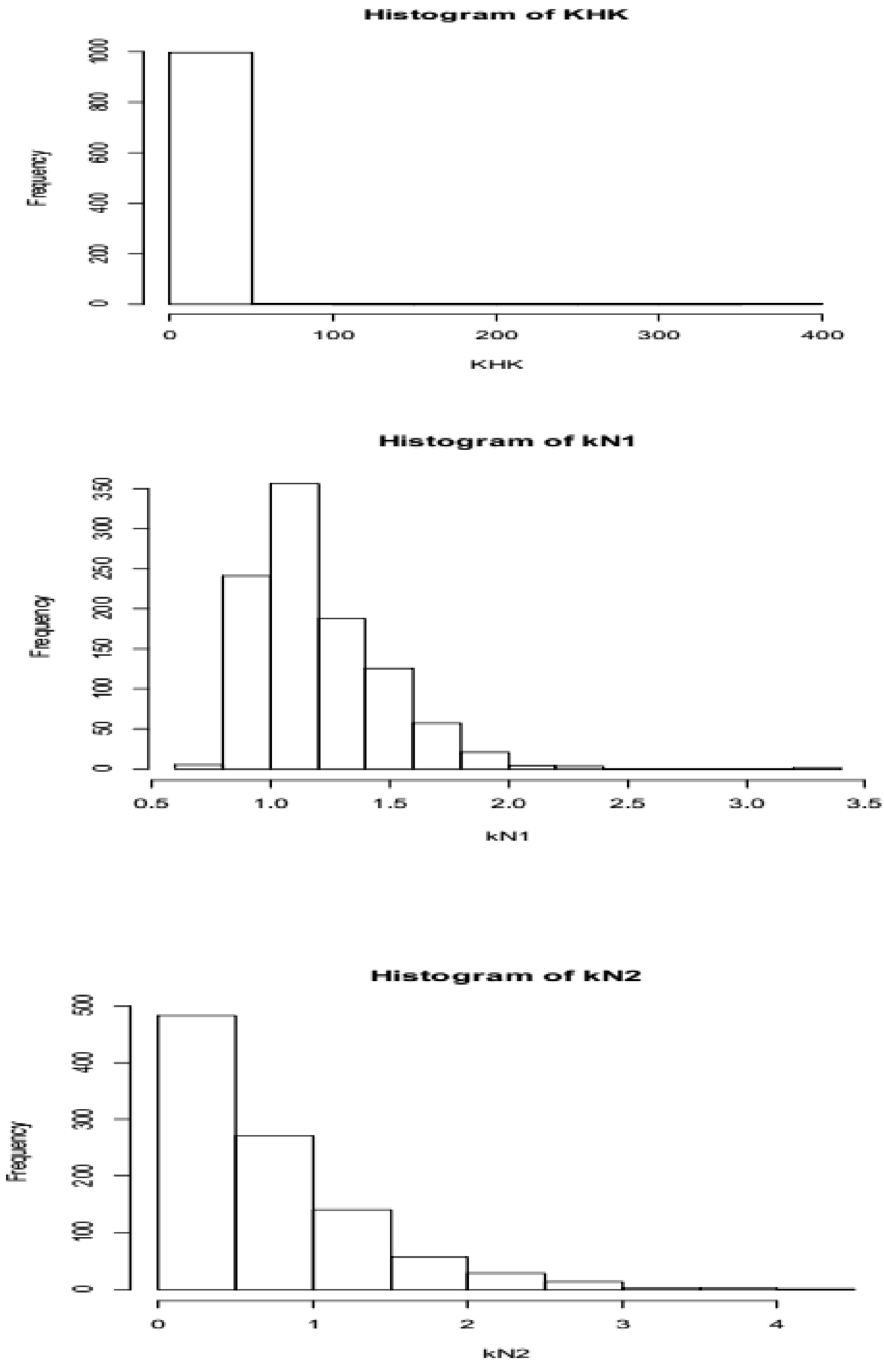

Figure (3.3) the histogram of the sampling distribution of $\hat{\mathbf{k}}_{\mathbf{HK}}$, $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ when $e \sim T_{(4)}$.

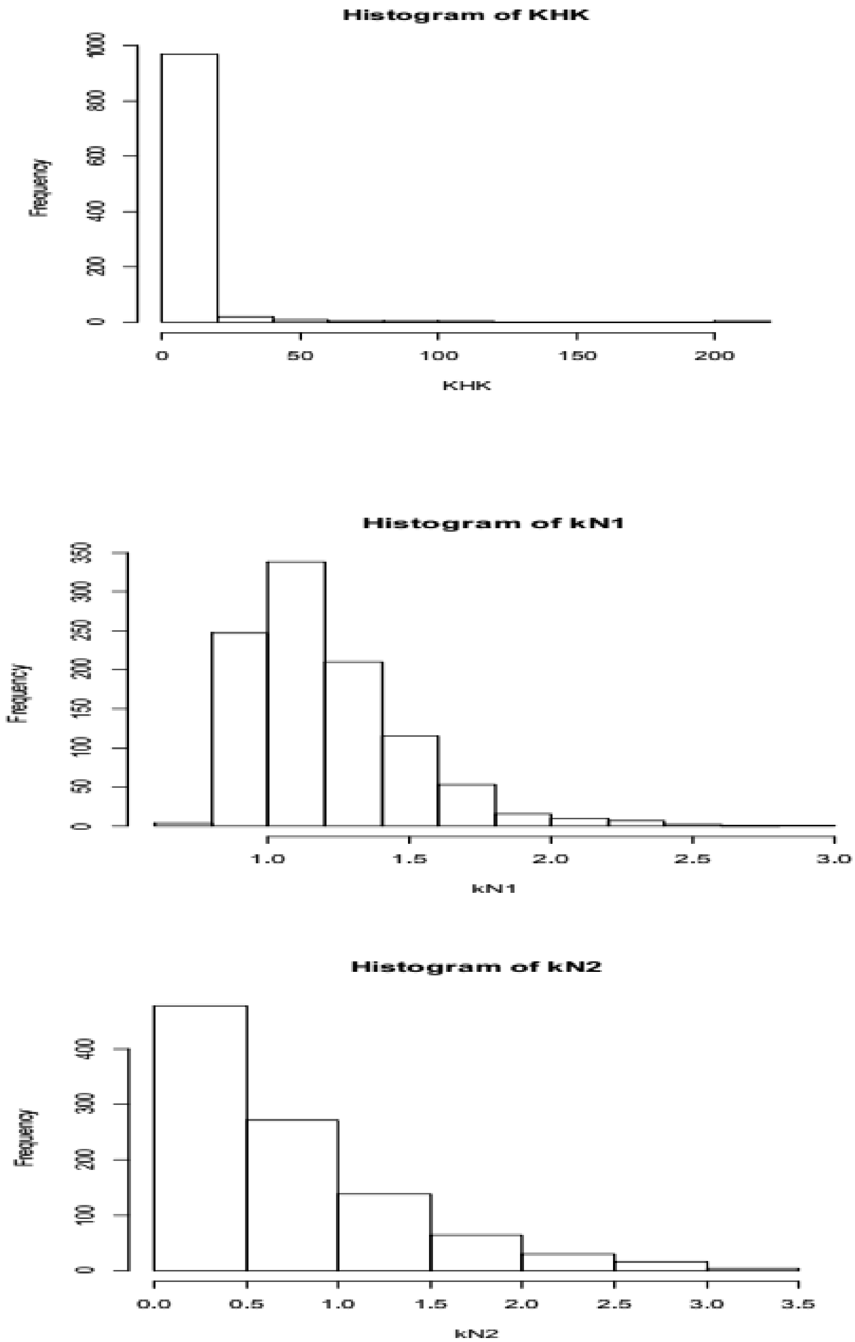

Figure (3.4) the histogram of the sampling distribution of $\hat{\mathbf{k}}_{\mathbf{HK}}$, $\hat{\mathbf{k}}_{\mathbf{N1}}$ and $\hat{\mathbf{k}}_{\mathbf{N2}}$ when $e \sim \chi^2_{(1)}$.

The investigation of the graphical representation reveals that:

- The histograms of the formula $\hat{k}_{HK}$ for different distributions are almost similar. It is similar to the gamma distribution, a special case of the beta and also a non-central F distribution.

- The histograms of the formula $\hat{k}_{N1}$ for different distributions are almost similar. It is similar to the beta distribution and F distribution.

- The histograms of the formula $\hat{k}_{N2}$ for different distributions are almost similar. It is similar to the gamma distribution, a special case of the beta and also a non-central F distribution.

- When comparing the formulas ($\hat{k}_{HK}$, $\hat{k}_{N1}$ and $\hat{k}_{N2}$), found that:

  - The formulas $\hat{k}_{HK}$ and $\hat{k}_{N2}$ have almost the same distribution form, a special case of the beta.

  - The formula $\hat{k}_{N1}$ followed form beta distribution and F distribution and this was confirmed by a numerical study of the person's system approach.

4. Results and Discussion

The present paper is concerned with studying the ridge parameters k through deterministic and stochastic approach in the case of ordinary ridge regression (ORR). In the deterministic approach, some new formulas for the ridge parameters are proposed and compared with the formula suggested by Hoerl and Kennard (1970a). The performance of the proposed ridge parameters is valuated through a simulation study in the presence of multicollinearity. The evaluation is based on the mean square error (MSE) and relative MSE (RMSE). The results found that formulas $\hat{k}_{N1}$ and $\hat{k}_{N2}$ are better than OLS and $\hat{k}_{HK}$ for all selected distributions and all values of ρ , p and n. The formula $\hat{k}_{N2}$ is better than $\hat{k}_{N1}$ for all selected distributions and all values of ρ , p and for large values of n whereas k_{N1} is the best when the sample size is too small and the application used real data confirmed these results.

In the stochastic approach, the main properties of the proposed new formulas and the formula suggested by Hoerl and Kennard (1970a) are studied. The probability density function (pdf) of the formulas are derived and the empirical distributions of these formulas are derived using Pearson's method. It is found that the formulas $\hat{k}_{HK}$ and $\hat{k}_{N2}$ have almost the same distribution form, a special case of the beta [Pearson's type I system of frequency curves]. The formula $\hat{k}_{N1}$ followed the beta distribution and F distribution [Person's type VI system of frequency curves]. [See Lahcene. (2013)].

References

1. Abd. Eledum, H. Y. and Alkhalifa, A .A. (2012). Generalized Two Stages Ridge Regression Estimator GTR for Multicollinearity and Autocorrelated Errors. *Canadian Journal on Science and Engineering Mathematics* ,vol. 3, no. 3, pp.79-85.

2. Abd El-Salam, M.E. (2011). A Modification of the Ridge Type Regression Estimators. *American Journal of Applied Sciences*, vol. 8, no. 1, pp. 97-102.

3. Abd El-Salam, M.E. (2015). Alternative Ridge Robust Regression Estimator for Dealing with Collinear Influential Data Points. *International Journal of Contemporary Mathematical Sciences*, vol.10, no.3, pp.119 – 130.

4. Alkhamisi, M. A., Khalaf, G. and Shukur, G. (2006). Some Modifications for Choosing Ridge Parameter. *Communication in Statistics- Theory and Methods,* vol.35, pp.1–16.

5. Al-Hassan and Yazid. M. (2010). Performance of a New Ridge Regression Estimator. *Journal of Applied Sciences*, vol. 9,pp. 43-50.

6. Batah, F.S.M , Ozkale, M.R, Gore, S.D.(2009) .Combining Unbiased Ridge and Principal Component Regression Estimators. *Communication in statistic-theory and methods*, vol.38, pp. 2201–2209.

7. Dorugade, A.V, Kashid ,D.N. (2010). Alternative Method for Choosing Ridge Parameter for Regression. *Applied Mathematical Sciences*, Vol. 4, no. 9, pp.447 – 456.

8. Dorugade, A.V.(2014). On Comparison of Some Ridge Parameters in Ridge Regression. *Sri Lankan Journal of Applied Statistics*, vol.15, no.1, pp.31-45.

9. Dorugade, A.V.(2014). New Ridge Parameters for Ridge Regression. *Journal of Association of Arab Universities for Basic and Applied Sciences,* vol. 15, pp.94–99.

10. Elderton.W.P and Johnson.N.L (1969). *Systems of Frequency Curves*. Cambridge University Press.

11. El-Dereny, M.E, Rashwan N. I.(2011). Solving Multicollinearity Problem Using Ridge Regression Models. *International Journal of Contemporary Mathematical Sciences.* vol. 6, no. 12, pp.585 – 600.

12. Firinguetti.L and Rubio.H.(2000)**.**A note on the Moments of Stochastic Shrinkage Parameters in Ridge Regression **.** *Communications in Statistics: Simulation and Computation,* vol. 29, no.3, pp.955-970

13. Golub, G. H., Heath, M. and Wahba, G. (1979). Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, vol.21,no.12,pp. 215–23.

14. Hoerl, A. E. (1962). Application of Ridge Analysis to Regression Problems.*Chemical Engineering Progress*, vol 58, no 3, pp. 54-59.

15. Hoerl, A. E., and Kennard, R. W. (1970a). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*. vol. 12, no. 1. pp. 55-67.

16. Hoerl, A. E.,and Kennard, R. W. (1970b). Ridge Regression: Application to Nonorthogonal Problems. *Technometrics,*vol. 12, no. 3, pp. 69-82.

17. Hoerl, A. E, Kennard, R. W. and Baldwin, K. F. (1975). Ridge Regression: Some Simulation. *Communication in Statistics- Theory and Methods,* vol.4, pp.105–123.

18. Johnson, N. and Kotz, S. (1970). *Continuous Univariate Distributions*.vol.1,2nd edition John Wiley &Sons, New York.

19. Kibria, B. M. G. (2003). Performance of Some New Ridge Regression Estimators. *Communications in Statistics-Simulation and Computation,* vol. 32, no. 2, pp. 419-435

20. Khalaf, G., Shukur, G. (2005). Choosing Ridge Parameters for Regression Problems. *Communication in Statistics- Theory and Methods,* vol. 34, pp.1177–1182.

21. Khalaf G, Mansson K and Shukur .G (2013). Modified Ridge Regression Estimators. *Communication in statistics- theory and methods*, vol.42,pp.1476–1487.

22. Khalaf, G and Iguernane, M. (2016) "Multicollinearity and A Ridge Parameter Estimation Approach," *Journal of Modern Applied Statistical Methods*, vol. 15, no. 2, pp.400-410.

23. Lahcene, B (2013) On Pearson Families of Distributions and its Applications. *African Journal of Mathematics and Computer Science Research* ,vol. 6, no. 5, pp. 108-117

24. Lawless, J.F. and Wang, P.(1976).A simulation Study of Ridge and other Regression Estimators. *Communications in Statistics—Simulation and Computation*, A, vol.5, pp.307-324.

25. Mansson .k ,Shukur .G, Kibria B. M. G. (2010). A Simulation Study of Some Ridge Regression Estimatorsunder Different Distributional Assumptions. *Communications in Statistics—Simulation and Computation*, vol.39, pp. 1639–1670.

26. Marquardt, D.W. (1970) Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics*, vol. 12, no. 3. pp.591-612

27. Marquardt, D.W and Snee, R.D. (1975). Ridge Regression in Practice. *The American Statistician*, vol. 29, no. 1. 3-20.

28. Muniz, G., Kibria, B. M. G. (2009). On Some Ridge Regression Eestimators: An Empirical Comparisons. *Communications in Statistics—Simulation and Computation*, vol.38,pp.621–630.

29. Muniz, G., Kibria, B. M. G.,Mansson, K., Shukur,G. (2012).On Developing Ridge Regression Parameters: A Graphical Investigation. *Sort: Statistics and Operations Research Transactions*, vol. 36, no. 2, pp.115–138.

30. Rubio.H and Firinguetti.L (2002). The Distribution of Stochastic Shrinkage Parameters in Ridge Regression**.** *Communication in Statistics- Theory and Methods,* vol.31, no.9, pp.1531-1547.

Appendix A

Table (2.1) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=3, e ~N(0,1)

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	e ~N(0,1) ,10 Observation			
0.2	0.4022	0.3363	0.2926	0.2877
	(1.39)	(1.16)	(1.02)	
0.7	0.5000	0.4519	0.4031	0.4785
	(1.04)	(0.95)	(0.84)	
0.8	0.6202	0.5928	0.5722	0.7442
	(0.83)	(0.79)	(0.77)	
0.9	0.8268	0.7748	0.7966	1.3050
	(0.63)	(0.59)	(0.61)	
	e ~N(0,1) ,30 Observation			
0.2	0.1509	0.1251	0.1089	0.1041
	(1.44)	(1.20)	(1.04)	
0.7	0.1969	0.1689	0.1509	0.1827
	(1.08)	(0.92)	(0.83)	
0.8	0.2760	0.2386	0.2023	0.2862
	(0.96)	(0.83)	(0.71)	
0.9	0.5002	0.4272	0.3422	0.6204
	(0.81)	(0.69)	(0.55)	
	e ~N(0,1) ,50 Observation			
0.2	0.1224	0.0990	0.0865	0.0839
	(1.45)	(1.18)	(1.03)	
0.7	0.1412	0.1089	0.0914	0.0941
	(1.50)	(1.16)	(0.97)	
0.8	0.1936	0.1494	0.1208	0.1359
	(1.42)	(1.09)	(0.89)	
0.9	0.3338	0.2696	0.2060	0.2724
	(1.23)	(0.99)	(0.75)	
	e ~N(0,1) ,75 Observation			
0.2	0.0794	0.0676	0.0607	0.0590
	(1.34)	(1.14)	(1.02)	
0.7	0.1034	0.0951	0.0892	0.1157
	(0.89)	(0.82)	(0.77)	
0.8	0.1636	0.1416	0.1136	0.1874
	(0.87)	(0.76)	(0.61)	
0.9	0.3067	0.2427	0.1534	0.3399
	(0.90)	(0.71)	(0.45)	
	e ~N(0,1) ,100 Observation			
0.2	0.0833	0.0756	0.0705	0.0689
	(1.20)	(1.09)	(1.02)	
0.7	0.0784	0.0727	0.0701	0.0794
	(0.99)	(0.92)	(0.88)	
0.8	0.1154	0.1024	0.0915	0.1189
	(0.97)	(0.86)	(0.77)	
0.9	0.2291	0.1856	0.1323	0.2279
	(1.01)	(0.81)	(0.58)	

Table (2.2) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=3, $e \sim F(2,10)$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim F(2,10)$, 10 Observation			
0.2	0.5139	0.4657	0.4385	0.4378
	(1.17)	(1.06)	(1.00)	
0.7	0.6115	0.5930	0.5830	0.6759
	(0.90)	(0.88)	(0.86)	
0.8	0.6987	0.6839	0.6936	0.8805
	(0.79)	(0.78)	(0.79)	
0.9	0.9552	0.9105	0.9241	1.5816
	(0.60)	(0.58)	(0.58)	
	$e \sim F(2,10)$, 30 Observation			
0.2	0.3101	0.2605	0.2363	0.2320
	(1.33)	(1.12)	(1.01)	
0.7	0.3708	0.3197	0.2888	0.3037
	(1.22)	(1.05)	(0.95)	
0.8	0.4413	0.3934	0.3572	0.3997
	(1.10)	(0.98)	(0.89)	
0.9	0.6544	0.5988	0.5410	0.7520
	(0.87)	(0.79)	(0.72)	
	$e \sim F(2,10)$, 50 Observation			
0.2	0.2966	0.2482	0.2294	0.2272
	(1.30)	(1.09)	(1.00)	
0.7	0.3519	0.2815	0.2465	0.2437
	(1.44)	(1.16)	(1.01)	
0.8	0.4073	0.3350	0.2920	0.2932
	(1.39)	(1.14)	(0.99)	
0.9	0.5172	0.4607	0.4116	0.4404
	(1.17)	(1.05)	(0.93)	
	$e \sim F(2,10)$, 75 Observation			
0.2	0.2411	0.2062	0.1922	0.1905
	(1.26)	(1.08)	(1.00)	
0.7	0.2560	0.2075	0.1818	0.1841
	(1.39)	(1.13)	(0.99)	
0.8	0.3118	0.2555	0.2180	0.2346
	(1.32)	(1.09)	(0.93)	
0.9	0.4641	0.3997	0.3336	0.4029
	(1.15)	(0.99)	(0.83)	
	$e \sim F(2,10)$, 100 Observation			
0.2	0.2446	0.2193	0.2094	0.2081
	(1.17)	(1.05)	(1.00)	
0.7	0.2328	0.1881	0.1674	0.1653
	(1.41)	(1.14)	(1.01)	
0.8	0.2752	0.2206	0.1915	0.1923
	(1.43)	(1.15)	(0.99)	
0.9	0.3944	0.3238	0.2705	0.2906
	(1.36)	(1.11)	(0.93)	

Table (2.3) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=3, $e \sim T_{(4)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim T_{(4)}$, 10 Observation			
0.2	0.5507	0.4994	0.4686	0.4652
	(1.18)	(1.07)	(1.00)	
0.7	0.6570	0.6382	0.6308	0.7039
	(0.93)	(0.91)	(0.89)	
0.8	0.7293	0.7128	0.7109	0.8886
	(0.82)	(0.80)	(0.80)	
0.9	1.0123	0.9664	0.9825	1.6271
	(0.62)	(0.59)	(0.60)	
	$e \sim T_{(4)}$, 30 Observation			
0.2	0.2830	0.2347	0.2102	0.2056
	(1.37)	(1.14)	(1.02)	
0.7	0.3332	0.2824	0.2515	0.2613
	(1.28)	(1.08)	(0.96)	
0.8	0.4134	0.3677	0.3321	0.3753
	(1.10)	(0.98)	(0.88)	
0.9	0.6159	0.5672	0.5205	0.7080
	(0.87)	(0.80)	(0.74)	
	$e \sim T_{(4)}$, 50 Observation			
0.2	0.2534	0.2105	0.1923	0.1898
	(1.33)	(1.10)	(1.01)	
0.7	0.3045	0.2363	0.2014	0.1979
	(1.54)	(1.19)	(1.02)	
0.8	0.3621	0.2888	0.2442	0.2449
	(1.48)	(1.18)	(0.99)	
0.9	0.5081	0.4408	0.3819	0.4110
	(1.24)	(1.07)	(0.93)	
	$e \sim T_{(4)}$, 75 Observation			
0.2	0.1953	0.1660	0.1531	0.1511
	(1.29)	(1.09)	(1.01)	
0.7	0.2009	0.1618	0.1402	0.1420
	(1.41)	(1.14)	(0.99)	
0.8	0.2530	0.2019	0.1676	0.1842
	(1.37)	(1.09)	(0.91)	
0.9	0.4115	0.3465	0.2813	0.3463
	(1.19)	(1.00)	(0.81)	
	$e \sim T_{(4)}$, 100 Observation			
0.2	0.1991	0.1794	0.1703	0.1688
	(1.17)	(1.06)	(1.00)	
0.7	0.1843	0.1491	0.1313	0.1290
	(1.43)	(1.16)	(1.02)	
0.8	0.2121	0.1690	0.1453	0.1474
	(1.44)	(1.15)	(0.99)	
0.9	0.3263	0.2658	0.2176	0.2441
	(1.34)	(1.09)	(0.89)	

Table (2.4) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=3, $e \sim {\chi^2}_{(1)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim {\chi^2}_{(1)}$,10 Observation			
0.2	0.4766	0.4258	0.3962	0.3957
	(1.20)	(1.07)	(1.00)	
0.7	0.5627	0.5377	0.5195	0.6174
	(0.91)	(0.87)	(0.84)	
0.8	0.6548	0.6363	0.6398	0.8310
	(0.79)	(0.77)	(0.76)	
0.9	0.9039	0.8452	0.8566	1.4867
	(0.61)	(0.57)	(0.58)	
	$e \sim {\chi^2}_{(1)}$,30 Observation			
0.2	0.2793	0.2340	0.2115	0.2073
	(1.34)	(1.12)	(1.02)	
0.7	0.3348	0.2887	0.2595	0.2787
	(1.20)	(1.04)	(0.93)	
0.8	0.4215	0.3745	0.3351	0.3929
	(1.07)	(0.95)	(0.85)	
0.9	0.6364	0.5737	0.5041	0.7425
	(0.86)	(0.77)	(0.67)	
	$e \sim {\chi^2}_{(1)}$,50 Observation			
0.2	0.2554	0.2121	0.1940	0.1917
	(1.33)	(1.10)	(1.01)	
0.7	0.3076	0.2397	0.2059	0.2032
	(1.51)	(1.18)	(1.01)	
0.8	0.3661	0.2917	0.2477	0.2495
	(1.47)	(1.17)	(0.99)	
0.9	0.4910	0.4293	0.3738	0.4132
	(1.19)	(1.04)	(0.90)	
	$e \sim {\chi^2}_{(1)}$,75 Observation			
0.2	0.1984	0.1689	0.1562	0.1543
	(1.28)	(1.09)	(1.01)	
0.7	0.2114	0.1712	0.1497	0.1545
	(1.36)	(1.11)	(0.97)	
0.8	0.2856	0.2283	0.1888	0.2121
	(1.34)	(1.08)	(0.89)	
0.9	0.4128	0.3543	0.2936	0.3721
	(1.11)	(0.95)	(0.79)	
	$e \sim {\chi^2}_{(1)}$,100 Observation			
0.2	0.2052	0.1846	0.1754	0.1740
	(1.17)	(1.06)	(1.00)	
0.7	0.1905	0.1555	0.1383	0.1367
	(1.39)	(1.14)	(1.01)	
0.8	0.2240	0.1787	0.1537	0.1578
	(1.42)	(1.13)	(0.97)	
0.9	0.3407	0.2771	0.2259	0.2567
	(1.33)	(1.08)	(0.88)	

Table (2.5) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=4, e ~N(0,1)

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	e ~N(0,1) ,10 Observation			
0.2	0.5404	0.4980	0.4996	0.5367
	(1.0)	(0.92)	(0.93)	
0.7	0.7522	0.7383	0.7625	1.2593
	(0.59)	(0.59)	(0.61)	
0.8	0.9640	0.8880	0.8522	1.8814
	(0.51)	(0.47)	(0.45)	
0.9	1.6013	1.1621	0.9578	3.7369
	(0.43)	(0.31)	(0.26)	
	e ~N(0,1) ,30 Observation			
0.2	0.1834	0.1511	0.1375	0.1343
	(1.36)	(1.12)	(1.02)	
0.7	0.2827	0.2507	0.2236	0.3180
	(0.89)	(0.79)	(0.70)	
0.8	0.4248	0.3670	0.2997	0.5286
	(0.80)	(0.69)	(0.57)	
0.9	0.7984	0.6489	0.5617	1.1633
	(0.69)	(0.56)	(0.48)	
	e ~N(0,1) ,50 Observation			
0.2	0.1506	0.1218	0.1112	0.1093
	(1.37)	(1.11)	(1.01)	
0.7	0.1963	0.1682	0.1456	0.2008
	(0.98)	(0.84)	(0.73)	
0.8	0.2984	0.2436	0.1835	0.3212
	(0.93)	(0.76)	(0.57)	
0.9	0.5560	0.4420	0.3457	0.6739
	(0.83)	(0.66)	(0.51)	
	e ~N(0,1) ,75 Observation			
0.2	0.1051	0.0908	0.0843	0.0828
	(1.26)	(1.09)	(1.01)	
0.7	0.1169	0.1065	0.1004	0.1142
	(1.02)	(0.93)	(0.88)	
0.8	0.1743	0.1523	0.1335	0.1738
	(1.00)	(0.88)	(0.77)	
0.9	0.3467	0.2850	0.2178	0.3459
	(1.00)	(0.82)	(0.63)	
	e ~N(0,1) ,100 Observation			
0.2	0.0913	0.0828	0.0781	0.0765
	(1.19)	(1.08)	(1.02)	
0.7	0.0947	0.0891	0.0859	0.0988
	(0.96)	(0.90)	(0.87)	
0.8	0.1431	0.1286	0.1142	0.1531
	(0.93)	(0.84)	(0.75)	
0.9	0.2872	0.2358	0.1733	0.3001
	(0.96)	(0.79)	(0.58)	

Table (2.6) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=4, $e \sim F(2,10)$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim F(2,10)$, 10 Observation			
0.2	0.6319	0.6172	0.6406	0.6890
	(0.91)	(0.89)	(0.92)	
0.7	0.8486	0.8444	0.8771	1.4085
	(0.60)	(0.59)	(0.62)	
0.8	1.0500	0.9812	0.9497	2.0073
	(0.52)	(0.49)	(0.47)	
0.9	1.6476	1.2316	1.0298	3.7763
	(0.44)	(0.33)	(0.27)	
	$e \sim F(2,10)$, 30 Observation			
0.2	0.3383	0.2849	0.2664	0.2635
	(1.28)	(1.08)	(1.01)	
0.7	0.4426	0.3968	0.3676	0.4158
	(1.06)	(0.95)	(0.88)	
0.8	0.5665	0.5192	0.4743	0.6065
	(0.93)	(0.86)	(0.78)	
0.9	0.8917	0.8091	0.7705	1.2081
	(0.74)	(0.67)	(0.64)	
	$e \sim F(2,10)$, 50 Observation			
0.2	0.3337	0.2781	0.2622	0.2604
	(1.28)	(1.06)	(1.00)	
0.7	0.3616	0.3036	0.2716	0.2901
	(1.25)	(1.05)	(0.94)	
0.8	0.4505	0.3909	0.3422	0.4033
	(1.12)	(0.97)	(0.85)	
0.9	0.6855	0.6167	0.5622	0.7757
	(0.88)	(0.79)	(0.72)	
	$e \sim F(2,10)$, 75 Observation			
0.2	0.2696	0.2362	0.2260	0.2246
	(1.20)	(1.05)	(1.00)	
0.7	0.2908	0.2429	0.2229	0.2213
	(1.31)	(1.09)	(1.01)	
0.8	0.3560	0.2982	0.2699	0.2735
	(1.30)	(1.09)	(0.99)	
0.9	0.5218	0.4621	0.4170	0.4518
	(1.16)	(1.02)	(0.92)	
	$e \sim F(2,10)$, 100 Observation			
0.2	0.2533	0.2264	0.2179	0.2166
	(1.16)	(1.04)	(1.00)	
0.7	0.2455	0.2025	0.1856	0.1840
	(1.33)	(1.10)	(1.01)	
0.8	0.2957	0.2434	0.2193	0.2224
	(1.33)	(1.09)	(0.99)	
0.9	0.4328	0.3739	0.3307	0.3632
	(1.19)	(1.03)	(0.91)	

Table (2.7) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=4, $e \sim T_{(4)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim T_{(4)}$, 10 Observation			
0.2	0.6736	0.6494	0.6630	0.7001
	(0.96)	(0.92)	(0.94)	
0.7	0.9011	0.8946	0.9335	1.4435
	(0.62)	(0.62)	(0.65)	
0.8	1.1195	1.0419	1.0107	2.0735
	(0.54)	(0.50)	(0.49)	
0.9	1.7744	1.3053	1.0662	3.9388
	(0.45)	(0.33)	(0.27)	
	$e \sim T_{(4)}$, 30 Observation			
0.2	0.3157	0.2621	0.2426	0.2392
	(1.31)	(1.09)	(1.01)	
0.7	0.3974	0.3479	0.3197	0.3550
	(1.12)	(0.98)	(0.90)	
0.8	0.5150	0.4666	0.4254	0.5295
	(0.97)	(0.88)	(0.80)	
0.9	0.8224	0.7563	0.7276	1.0940
	(0.75)	(0.69)	(0.66)	
	$e \sim T_{(4)}$, 50 Observation			
0.2	0.2875	0.2357	0.2198	0.2178
	(1.31)	(1.08)	(1.00)	
0.7	0.3241	0.2661	0.2345	0.2478
	(1.31)	(1.07)	(0.95)	
0.8	0.4138	0.3518	0.3034	0.3519
	(1.18)	(0.99)	(0.86)	
0.9	0.6414	0.5785	0.5273	0.7009
	(0.92)	(0.83)	(0.75)	
	$e \sim T_{(4)}$, 75 Observation			
0.2	0.2206	0.1913	0.1815	0.1799
	(1.22)	(1.06)	(1.00)	
0.7	0.2273	0.1888	0.1718	0.1712
	(1.33)	(1.10)	(1.00)	
0.8	0.2867	0.2391	0.2141	0.2213
	(1.29)	(1.08)	(0.97)	
0.9	0.4464	0.3901	0.3446	0.3930
	(1.14)	(0.99)	(0.88)	
	$e \sim T_{(4)}$, 100 Observation			
0.2	0.2060	0.1852	0.1775	0.1760
	(1.17)	(1.05)	(1.00)	
0.7	0.1976	0.1636	0.1488	0.1476
	(1.34)	(1.11)	(1.01)	
0.8	0.2538	0.2065	0.1833	0.1882
	(1.35)	(1.09)	(0.97)	
0.9	0.4130	0.3454	0.2948	0.3340
	(1.24)	(1.03)	(0.88)	

Table (2.8) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=4, $e \sim \chi^2_{(1)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
		$e \sim \chi^2_{(1)}$,10 Observation		
0.2	0.5881	0.5701	0.5916	0.6389
	(0.92)	(0.89)	(0.92)	
0.7	0.8097	0.8022	0.8302	1.3855
	(0.58)	(0.58)	(0.59)	
0.8	1.0234	0.9424	0.8948	2.0088
	(0.51)	(0.47)	(0.45)	
0.9	1.6624	1.1989	0.9792	3.8516
	(0.43)	(0.31)	(0.25)	
		$e \sim \chi^2_{(1)}$,30 Observation		
0.2	0.3064	0.2570	0.2393	0.2363
	(1.29)	(1.08)	(1.01)	
0.7	0.4038	0.3560	0.3240	0.3776
	(1.07)	(0.94)	(0.86)	
0.8	0.5271	0.4755	0.4239	0.5649
	(0.93)	(0.84)	(0.75)	
0.9	0.8518	0.7620	0.7215	1.1545
	(0.74)	(0.66)	(0.62)	
		$e \sim \chi^2_{(1)}$,50 Observation		
0.2	0.2798	0.2313	0.2165	0.2146
	(1.30)	(1.07)	(1.00)	
0.7	0.3482	0.2872	0.2527	0.2724
	(1.28)	(1.05)	(0.93)	
0.8	0.4406	0.3785	0.3265	0.3868
	(1.14)	(0.98)	(0.84)	
0.9	0.6743	0.6146	0.5685	0.7617
	(0.89)	(0.81)	(0.75)	
		$e \sim \chi^2_{(1)}$,75 Observation		
0.2	0.2341	0.2033	0.1935	0.1920
	(1.21)	(1.05)	(1.00)	
0.7	0.2518	0.2095	0.1915	0.1924
	(1.31)	(1.09)	(0.99)	
0.8	0.3149	0.2640	0.2373	0.2474
	(1.27)	(1.07)	(0.96)	
0.9	0.4747	0.4215	0.3765	0.4294
	(1.11)	(0.98)	(0.88)	
		$e \sim \chi^2_{(1)}$,100 Observation		
0.2	0.2085	0.1871	0.1794	0.1780
	(1.17)	(1.05)	(1.00)	
0.7	0.2076	0.1707	0.1550	0.1544
	(1.35)	(1.11)	(1.00)	
0.8	0.2631	0.2138	0.1893	0.1948
	(1.35)	(1.09)	(0.97)	
0.9	0.4171	0.3510	0.3008	0.3381
	(1.23)	(1.04)	(0.89)	

Table (2.9) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=5, e ~N(0,1)

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	e ~N(0,1) ,10 Observation			
0.2	0.8065	0.8206	0.8579	1.2859
	(0.62)	(0.63)	(0.66)	
0.7	1.4770	1.2331	1.2194	3.7418
	(0.39)	(0.33)	(0.33)	
0.8	2.0305	1.4216	1.2549	5.5855
	(0.36)	(0.25)	(0.22)	
0.9	3.6126	1.6886	1.1765	10.797
	(0.33)	(0.16)	(0.11)	
	e ~N(0,1) ,30 Observation			
0.2	0.19531	0.1658	0.1545	0.1519
	(1.28)	(1.09)	(1.01)	
0.7	0.3434	0.2998	0.2733	0.3284
	(1.05)	(0.91)	(0.83)	
0.8	0.4734	0.4219	0.3760	0.5070
	(0.93)	(0.83)	(0.74)	
0.9	0.8088	0.7272	0.6809	1.0440
	(0.77)	(0.69)	(0.65)	
	e ~N(0,1) ,50 Observation			
0.2	0.1476	0.1236	0.1163	0.1152
	(1.28)	(1.07)	(1.00)	
0.7	0.3056	0.2542	0.1898	0.3702
	(0.83)	(0.69)	(0.51)	
0.8	0.4696	0.3610	0.2451	0.5887
	(0.79)	(0.61)	(0.42)	
0.9	0.8623	0.6201	0.5487	1.2154
	(0.71)	(0.51)	(0.45)	
	e ~N(0,1) ,75 Observation			
0.2	0.1138	0.0955	0.0893	0.0879
	(1.29)	(1.08)	(1.01)	
0.7	0.1616	0.1489	0.1382	0.1764
	(0.92)	0.84)	(0.78)	
0.8	0.2484	0.2166	0.1785	0.2774
	(0.89)	(0.78)	(0.64)	
0.9	0.4870	0.3979	0.3009	0.5558
	(0.88)	(0.72)	(0.54)	
	e ~N(0,1) ,100 Observation			
0.2	0.1151	0.1014	0.0964	0.0953
	(1.20)	(1.06)	(1.01)	
0.7	0.1288	0.1216	0.1161	0.1354
	(0.95)	(0.89)	(0.86)	
0.8	0.1954	0.1740	0.1496	0.2089
	(0.94)	(0.83)	(0.71)	
0.9	0.3898	0.3171	0.2315	0.4142
	(0.94)	(0.77)	(0.56)	

Table (2.10) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=5, $e \sim F(2,10)$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim F(2,10)$, 10 Observation			
0.2	0.8950	0.9401	1.0308	1.5213
	(0.58)	(0.61)	(0.67)	
0.7	1.5817	1.3230	1.3382	4.0786
	(0.39)	(0.32)	(0.33)	
0.8	2.1428	1.4965	1.3366	5.9741
	(0.36)	(0.25)	(0.22)	
0.9	3.7559	1.7450	1.2075	11.304
	(0.33)	(0.15)	(0.11)	
	$e \sim F(2,10)$, 30 Observation			
0.2	0.3774	0.3287	0.3129	0.3100
	(1.21)	(1.06)	(1.00)	
0.7	0.5064	0.4547	0.4267	0.4525
	(1.12)	(1.00)	(0.94)	
0.8	0.6182	0.5851	0.5609	0.6344
	(0.97)	(0.92)	(0.88)	
0.9	0.8997	0.8771	0.8740	1.1475
	(0.78)	(0.76)	(0.76)	
	$e \sim F(2,10)$, 50 Observation			
0.2	0.3314	0.2770	0.2627	0.2608
	(1.27)	(1.06)	(1.00)	
0.7	0.4295	0.3679	0.3192	0.3953
	(1.09)	(0.93)	(0.81)	
0.8	0.5722	0.4972	0.4245	0.5974
	(0.96)	(0.83)	(0.71)	
0.9	0.9330	0.8226	0.7936	1.2439
	(0.75)	(0.66)	(0.64)	
	$e \sim F(2,10)$, 75 Observation			
0.2	0.2888	0.2443	0.2338	0.2325
	(1.24)	(1.05)	(1.00)	
0.7	0.3118	0.2648	0.2453	0.2513
	(1.24)	(1.05)	(0.98)	
0.8	0.3946	0.3423	0.3119	0.3383
	(1.17)	(1.01)	(0.92)	
0.9	0.6057	0.5587	0.5139	0.6233
	(0.97)	(0.89)	(0.82)	
	$e \sim F(2,10)$, 100 Observation			
0.2	0.2877	0.2562	0.2490	0.2481
	(1.15)	(1.03)	(1.00)	
0.7	0.2875	0.2432	0.2274	0.2269
	(1.27)	(1.07)	(1.00)	
0.8	0.3544	0.3023	0.2780	0.2859
	(1.24)	(1.06)	(0.97)	
0.9	0.5353	0.4854	0.4440	0.4974
	(1.08)	(0.98)	(0.89)	

Table (2.11) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=5, $e \sim T_{(4)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim T_{(4)}$,10 Observation			
0.2	0.9243	0.9602	1.0270	1.4451
	(0.63)	(0.66)	(0.71)	
0.7	1.5917	1.3549	1.3336	3.9082
	(0.40)	(0.35)	(0.34)	
0.8	2.1413	1.5287	1.3339	5.7637
	(0.37)	(0.27)	(0.23)	
0.9	3.7155	1.7705	1.2116	11.0069
	(0.34)	(0.16)	(0.11)	
	$e \sim T_{(4)}$,30 Observation			
0.2	0.3420	0.2886	0.2705	0.2669
	(1.28)	(1.08)	(1.01)	
0.7	0.4778	0.4307	0.4057	0.4274
	(1.12)	(1.01)	(0.95)	
0.8	0.5916	0.5544	0.5274	0.5887
	(1.00)	(0.94)	(0.89)	
0.9	0.8780	0.8591	0.8630	1.1092
	(0.79)	(0.77)	(0.78)	
	$e \sim T_{(4)}$,50 Observation			
0.2	0.2801	0.2278	0.2132	0.2110
	(1.32)	(1.07)	(1.01)	
0.7	0.3906	0.3343	0.2922	0.3605
	(1.08)	(0.93)	(0.81)	
0.8	0.5360	0.4661	0.4011	0.5610
	(0.96)	(0.83)	(0.71)	
0.9	0.8841	0.7540	0.7181	1.1492
	(0.77)	(0.66)	(0.62)	
	$e \sim T_{(4)}$,75 Observation			
0.2	0.2355	0.1978	0.1877	0.1862
	(1.26)	(1.06)	(1.00)	
0.7	0.2549	0.2122	0.1939	0.1995
	(1.28)	(1.06)	(0.97)	
0.8	0.3384	0.2836	0.2514	0.2778
	(1.22)	(1.02)	(0.90)	
0.9	0.5515	0.4912	0.4345	0.5396
	(1.02)	(0.91)	(0.81)	
	$e \sim T_{(4)}$,100 Observation			
0.2	0.2381	0.2112	0.2042	0.2032
	(1.17)	(1.03)	(1.00)	
0.7	0.2241	0.1897	0.1758	0.1756
	(1.28)	(1.08)	(1.00)	
0.8	0.2924	0.2442	0.2204	0.2298
	(1.27)	(1.06)	(0.96)	
0.9	0.4838	0.4206	0.3688	0.4276
	(1.13)	(0.98)	(0.86)	

Table (2.12) Estimated MSEs ,RMSEs [shown between parentheses]for the different formulas of ridge parameter when P=5, $e \sim \chi^2_{(1)}$

corr	$\hat{k}_{HK}$	$\hat{k}_{N1}$	$\hat{k}_{N2}$	OLS
	$e \sim \chi^2_{(1)}$,10 Observation			
0.2	0.8237	0.8488	0.9090	1.3617
	(0.60)	(0.62)	(0.66)	
0.7	1.4850	1.2330	1.2505	3.7790
	(0.39)	(0.33)	(0.33)	
0.8	2.0162	1.4104	1.2763	5.5656
	(0.36)	(0.25)	(0.22)	
0.9	3.5275	1.6684	1.1880	10.593
	(0.33)	(0.16)	(0.11)	
	$e \sim \chi^2_{(1)}$,30 Observation			
0.2	0.3398	0.2927	0.2771	0.2743
	(1.23)	(1.06)	(1.01)	
0.7	0.4840	0.4338	0.4043	0.4419
	(1.09)	(0.98)	(0.91)	
0.8	0.5917	0.5596	0.5335	0.6229
	(0.95)	(0.89)	(0.86)	
0.9	0.8961	0.8596	0.8485	1.1612
	(0.77)	(0.74)	(0.73)	
	$e \sim \chi^2_{(1)}$,50 Observation			
0.2	0.2837	0.2361	0.2230	0.2212
	(1.28)	(1.06)	(1.00)	
0.7	0.4178	0.3556	0.3040	0.3900
	(1.07)	(0.91)	(0.78)	
0.8	0.5729	0.4920	0.4131	0.6011
	(0.95)	(0.82)	(0.69)	
0.9	0.9411	0.7846	0.7426	1.2677
	(0.74)	(0.62)	(0.59)	
	$e \sim \chi^2_{(1)}$,75 Observation			
0.2	0.2414	0.2033	0.1934	0.1920
	(1.25)	(1.05)	(1.00)	
0.7	0.2724	0.2305	0.2128	0.2209
	(1.23)	(1.04)	(0.96)	
0.8	0.3549	0.3059	0.2761	0.3071
	(1.16)	(0.99)	(0.89)	
0.9	0.5685	0.5188	0.4724	0.5866
	(0.97)	(0.88)	(0.81)	
	$e \sim \chi^2_{(1)}$,100 Observation			
0.2	0.2397	0.2117	0.2046	0.2036
	(1.17)	(1.03)	(1.00)	
0.7	0.2299	0.1939	0.1796	0.1805
	(1.27)	(1.07)	(0.99)	
0.8	0.2955	0.2459	0.2213	0.2323
	(1.27)	(1.06)	(0.95)	
0.9	0.4740	0.4111	0.3605	0.4164
	(1.14)	(0.99)	(0.87)	

Appendix B

The probability density function and the distribution function of $\hat{\mathbf{k}}_{\mathbf{HK}}$:

To derive the pdf of $\hat{k}_{HK}$ in equation (20)

$$\text{put} \quad y = \frac{\lambda}{\hat{k}_{HK}} . \qquad (1\text{-B})$$

Since the numerator of y follows central Chi-square distribution and the denominator follows a non-central Chi-square distribution then:

$$y \sim F_{(1,n-p)}(\theta, 0) ,$$

where:

y is a non-central *F* with (1 , *n*-*p*) degrees of freedom and with θ and 0 as first and second non-central parameters respectively [See Johnson and Kotz. (1970)].

$$f(y,1,(n-p),\theta,0)=e^{-\frac{\theta}{2}}\sum_{j=0}^{\infty}\left[\frac{(\theta/2)^j}{j!}\right].\left[\frac{(1/(n-p))^{j+\frac{1}{2}}y^{\frac{1}{2}+j-1}}{\left(1+\frac{1}{(n-p)}y\right)^{\frac{1+(n-p)}{2}+j}}\right]\left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right], \qquad (2\text{-B})$$

The density function of $\hat{k}_{HK}$ is obtained by replacing the variable y in equation (1-B) in to equation (2-B) as follows:

$$f(\hat{k}_{HK})=e^{-\frac{\theta}{2}}\sum_{j=0}^{\infty}\left[\frac{(\theta/2)^j}{j!}\right].\left[\frac{(\lambda/(n-p))^{j+\frac{1}{2}}\left(\frac{1}{\hat{k}_{HK}}\right)^{j+\frac{3}{2}}}{\left(1+\frac{\lambda}{(n-p)}\left(\frac{1}{\hat{k}_{HK}}\right)\right)^{\frac{1+(n-p)}{2}+j}}\right]\left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right], \qquad (3\text{-B})$$

$$\hat{k}_{HK} > 0 \quad .$$

The distribution function :

To derive the distribution function of $\hat{k}_{HK}$ in equation (22) let $F(X_i)$=$P(\hat{k}_{HK} \le X_i)$, Using equation (20), then

$$F(X_i) = P\left(\frac{\lambda}{y} \le X_i\right) = P\left(y \ge \frac{\lambda}{X_i}\right) = 1 - P\left(y \le \frac{\lambda}{X_i}\right), \quad (4\text{-B})$$

Put $y \sim F_{(1,n-p)}(\theta,0) = (n-p)G_{(1,n-p)}(\theta,0)$,

$$G_{(1,n-p)}(\theta,0) = \frac{\chi^2_{(1)}(\theta)}{\chi^2_{(n-p)}(0)} .$$

[See Johnson and Kotz (1970)].

By using the following transformation in equation (22),

let $g = \frac{y}{(n-p)}$ then $y = (n-p)g$ and $dy = (n-p)\,dg$ then

$$f(g) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{(g)^{j+\frac{1}{2}-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right], g > 0 \quad (5\text{-B})$$

then

$$P\left(y \le \frac{\lambda}{X_i}\right) = P\left((n-p)G \le \left(\frac{\lambda}{X_i}\right)\right) = P\left(G \le \left(\frac{\lambda}{X_i(n-p)}\right)\right)$$

$$= \int_0^{\left(\frac{\lambda}{X_i(n-p)}\right)} e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{(g)^{\frac{1}{2}+j-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] dg$$

then

$$f(g) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \cdot \int_0^{\left(\frac{\lambda}{X_i(n-p)}\right)} \left[\frac{(g)^{j+\frac{1}{2}-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] dg, \quad (6\text{-B})$$

$$0 \le g \le \left(\frac{\lambda}{X_i(n-p)}\right)$$

To calculate the integral make the following change of variables:

let $Z = \frac{g}{1+g}$, then $g = \frac{Z}{1-Z}$, $|dg| = \frac{dz}{(1-Z)^2}$.

after changing the border $0 \le Z \le \left(\frac{\lambda}{X_i(n-p)+\lambda}\right)$

put $r = \left(\frac{\lambda}{X_i(n-p)+\lambda}\right)$. then

$$P(z \le r) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] . \int_0^r (Z)^{j+\frac{1}{2}-1} (1-Z)^{\frac{(n-p)}{2}-1} dz$$

$$= e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \beta_r \left(\frac{1}{2}+j, \frac{(n-p)}{2}\right). \qquad \text{(7-B)}$$

transfer y to z the equation becomes

$$F(X_i) = 1 - P\,(z \le r).$$

(8-B)

Substitute equation (7-B) into (8-B), then the distribution function of $\hat{k}_{HK}$ is given as:

$$F(X_i) = 1 - e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] . \ \beta_r \left(j+\frac{1}{2}, \frac{(n-p)}{2}\right). \qquad \text{(9-B)}$$

where:

β_r is the incomplete beta $\int_0^{\frac{\lambda}{X_i(n-p)+\lambda}} (\boldsymbol{Z})^{\,j+\frac{1}{2}-1} (\boldsymbol{1}-\boldsymbol{Z})^{\frac{(n-p)}{2}-1} \ \boldsymbol{dz}$.

The probability density function and the distribution function of $\hat{k}_{N1}$

The probability density function:

To derive the pdf of $\hat{k}_{N1}$ from equation (24),

$$\text{Put} \quad y = \left[\frac{\lambda}{[\hat{\mathrm{k}}_{\mathrm{N1}}]^{p}}\right] \tag{10-B}$$

Since the numerator of y follows central Chi-square distribution and denominator follows a non-central Chi-square distribution then:

$$y \sim F_{(1,n-p)}(\theta, 0) \; .$$

where:

y is a non-central F with (1, n-p) degrees of freedom and with θ and 0 as first and second non-central parameters respectively.

Using equation (2-B) the density function of $\hat{k}_{N1}$ is obtained by replacing the variable y in to equation

(10-B) as follows:

$$f\left(\hat{k}_{N1}\right) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^{j}}{j!}\right] . \left[\frac{p\,(\lambda/(n-p))^{j+\frac{1}{2}} \left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]^{p}}\right]^{j+\frac{1}{2}} \left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]}\right]}{\left(1 + \frac{\lambda}{(n-p)}\left[\frac{1}{[\hat{\mathrm{k}}_{\mathrm{N1}}]^{p}}\right]\right)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right],$$

(11-B)

$$\hat{\mathrm{k}}_{\mathrm{N1}} > 0 \; .$$

The distribution function

To derive the distribution function of $\hat{\mathrm{k}}_{\mathrm{N1}}$ in equation (24) let $F(X_i)$=$P\left(\hat{\mathrm{k}}_{\mathrm{N1}} \leq X_i\right)$, using equation (11-B) ,then

$$F(X_i) = P\left(\left[\frac{\lambda}{y}\right]^p \leq X_i\right) = P\left(y \geq \frac{\lambda}{[X_i]^p}\right) = 1 - P\left(y \leq \frac{\lambda}{[X_i]^p}\right) \tag{12-B}$$

Put $y \sim F_{(1,n-p)}(\theta, 0) = (n-p)G_{(1,n-p)}(\theta, 0)$.

and $G_{(1,n-p)}(\theta, 0) = \frac{\chi^2_{(1)}(\theta)}{\chi^2_{(n-p)}(0)}$,

using the following transformation in equation (2-B),

let $g = \frac{y}{(n-p)}$ then $y = (n-p)g$, $dy = (n-p)\, dg$. then

$$f(g) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{(g)^{j+\frac{1}{2}-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right], g > 0 \tag{13-B}$$

Then

$$P\left(y \leq \frac{\lambda}{[X_i]^p}\right) = P\left((n-p)G \leq \left(\frac{\lambda}{[X_i]^p}\right)\right) = P\left(G \leq \left(\frac{\lambda}{[X_i]^p (n-p)}\right)\right) =$$

$$= e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \cdot \int_0^{\left(\frac{\lambda}{[X_i]^p (n-p)}\right)} \left[\frac{(g)^{j+\frac{1}{2}-1}\, dg}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right], \tag{14-B}$$

$$0 \leq g \leq \left(\frac{\lambda}{[X_i]^p (n-p)}\right)$$

To calculate the integral in equation (14-B) make the following change of variables:

let $Z = \frac{g}{1+g}$ then $g = \frac{Z}{1-Z}$, $|dg| = \frac{dz}{(1-Z)^2}$.

after changing the border $0 \leq Z \leq \left(\frac{\lambda}{[X_i]^p(n-p)+\lambda}\right)$. then

$$P\big(z(q)\big) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \cdot \int_0^q \left[\frac{\left(\frac{Z}{1-Z}\right)^{j+\frac{1}{2}-1} \frac{dz}{(1-Z)^2}}{\left(1+\frac{Z}{1-Z}\right)^{\frac{1+(n-p)}{2}+j}}\right],$$

(15-B)

where:

$$q = \left(\frac{\lambda}{[X_i]^p(n-p)+\lambda}\right),$$

then

$$P\big(z \leq (q)\big) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \cdot \int_0^q (Z)^{j+\frac{1}{2}-1}(1-Z)^{\frac{(n-p)}{2}-1} dz$$

$$= e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j, \frac{(n-p)}{2}\right)}\right] \beta_q \left(j+\frac{1}{2}, \frac{(n-p)}{2}\right).$$

(16-B)

transfer y to z the equation becomes

$$F(X_i) = 1 - P\big(z \leq (q)\big). \qquad (17\text{-B})$$

then

$$F(X_i) = 1 - e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2}+j,\frac{(n-p)}{2}\right)}\right] \cdot \beta_q\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right).$$

(18-B)

where:

β_q is the incomplete beta $\int_{\mathbf{0}}^{\frac{\lambda}{[X_i]^p(n-p)+\lambda}} (\boldsymbol{Z})^{j+\frac{1}{2}-1}(\mathbf{1}-\boldsymbol{Z})^{\frac{(n-p)}{2}-1} \boldsymbol{dz}.$

The probability density function and the distribution function of $\hat{k}_{N2}$

The probability density function

To derive the pdf of $\hat{k}_{N2}$ from equation (27)

put $y = \left[\frac{\lambda}{[\hat{k}_{N2}]^{-1}}\right] = \left[\lambda[\hat{k}_{N2}]\right]$ (19-B)

The density function of $\hat{k}_{N2}$ is obtained by replacing the variable y in equation (2-B) in to equation (19-B) as follows:

$$f(\hat{k}_{N2}) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{(\lambda/(n-p))^{j+\frac{1}{2}}\left[[\hat{k}_{N2}]\right]^{j+\frac{1}{2}-1}}{\left(1+\frac{\lambda}{(n-p)}\left[[\hat{k}_{N2}]\right]\right)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(j+\frac{1}{2},\frac{(n-p)}{2}\right)}\right]$$

(20-B)

$$\hat{k}_{N2} > 0 \quad .$$

The distribution function

To derive the distribution function of $\hat{\mathrm{k}}_{\mathrm{N2}}$ in equation (29)

let $F(X_i)$=$P\left(\hat{\mathrm{k}}_{\mathrm{N2}} \leq X_i\right)$ using equation (27)

$$F(X_i) = P\left(\left[\frac{\lambda}{y}\right]^{-1} \leq X_i\right) = P\left(y \geq \frac{\lambda}{[X_i]^{-1}}\right) = P(y \geq \lambda\,[X_i]) = 1 - P\,(y \leq \lambda\,[X_i]) \tag{21-B}$$

put $\quad y \sim F_{(1,n-p)}(\theta, 0) = (n-p)G_{(1,n-p)}(\theta, 0)$.

and $\quad G_{(1,n-p)}(\theta, 0) = \dfrac{\chi^2_{(1)}(\theta)}{\chi^2_{(n-p)}(0)}$,

using the following transformation in equation (2-B),

let$g = \frac{y}{(n-p)}$, then $y = (n-p)g$, $|dy| = (n-p)\,dg$.

Substitute y in equation (2-B) then

$$f(g) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{(g)^{j+\frac{1}{2}-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right)}\right], g > 0$$

let $P\,(y \leq \lambda\,[X_i])$ =$P\big((n-p)G \leq (\lambda\,[X_i])\big) = P\left(G \leq \left(\frac{\lambda\,[X_i]}{(n-p)}\right)\right) =$

$$= \int_0^{\left(\frac{\lambda[X_i]}{(n-p)}\right)} e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] . \left[\frac{(g)^{j+\frac{1}{2}-1}}{(1+g)^{\frac{1+(n-p)}{2}+j}}\right] \left[\frac{1}{\beta\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right)}\right] dg$$

$$= e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!} \right] \cdot \left[\frac{1}{\beta\left(\frac{1}{2} + j, \frac{(n-p)}{2}\right)} \right] \cdot \int_0^{\left(\frac{\lambda[X_i]}{(n-p)}\right)} \left[\frac{(g)^{\frac{1}{2}+j-1} \; dg}{(1+g)^{\frac{1+(n-p)}{2}+j}} \right]$$

(22-B)

$$0 \le g \le \left(\frac{\lambda[X_i]}{(n-p)} \right)$$

To calculate the integral in equation (22-B) make the following change of variables:

let $Z = \frac{g}{1+g}$,then $g = \frac{Z}{1-Z}$, $|dg| = \frac{dz}{(1-Z)^2}$.

after changing the border

$$0 \le Z \le \left(\frac{\lambda[X_i]}{(n-p)+\lambda[X_i]} \right).$$

Then equation (22-B) can be written as

$$P\big(z(w)\big) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!} \right] \cdot \left[\frac{1}{\beta\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right)} \right] \cdot \int_0^{w} \left[\frac{\left(\frac{Z}{1-Z}\right)^{j+\frac{1}{2}-1} \frac{dz}{(1-Z)^2}}{\left(1+\frac{Z}{1-Z}\right)^{\frac{1+(n-p)}{2}+j}} \right],$$

(23-B)

where: $w = \left(\frac{\lambda\,[X_i]}{(n-p)+\lambda\,[X_i]} \right)$,

then

$$P\big(z \le (w)\big) = e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!} \right] \cdot \left[\frac{1}{\beta\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right)} \right] \cdot \int_0^{w} (Z)^{j+\frac{1}{2}-1} (1-Z)^{\frac{(n-p)}{2}-1} dz$$

$$= e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(j+\frac{1}{2},\frac{(n-p)}{2}\right)}\right] \beta_w \left(j + \frac{1}{2}, \frac{(n-p)}{2}\right).$$

(24-B)

transfer y to z equation (21-B) becomes

$$F(X_i) = 1 - P\big(z \leq (w)\big) \qquad \text{(25-B)}$$

Substituting equation (24-B) into (25-B)then the distribution function of $\hat{k}_{N2}$ is given as:

$$F(X_i) = 1 - e^{-\frac{\theta}{2}} \sum_{j=0}^{\infty} \left[\frac{(\theta/2)^j}{j!}\right] \cdot \left[\frac{1}{\beta\left(j+\frac{1}{2}, \frac{(n-p)}{2}\right)}\right] \cdot \beta_w \left(j + \frac{1}{2}, \frac{(n-p)}{2}\right)$$

(26-B)

where:

β_w is the incomplete beta $\int_{\mathbf{0}}^{\frac{\lambda[X_i]}{(n-p)+\lambda[X_i]}} (\boldsymbol{Z})^{j+\frac{1}{2}-1} (\mathbf{1} - \boldsymbol{Z})^{\frac{(n-p)}{2}-1} \boldsymbol{dz}$.