[6] Coutis, C. (1996), " Partial Least Squares Algorithm Yields Shrinkage Estimators", Annals of Statistics, 24: 816-824.
[7] Droge, B. and T. George (1995), " On Selecting the Smoothing Parameter of Least Squares Regression Estimates using The Minimax Regret Approach ", Statistics and Decisions 13,1-20.
[8] Giles, D. E. A. (2002), " Preliminary-test and Bayes Estimation of a Location Parameter under ' Reflect Normal ' loss ", In A. Ullah, A. Wan and A. Chaturvedi (Eds). Handbook of Applied Econometric and Statistical Inference, NewYork: Marcker Becker.
[9] Giles, J. A. and D. E. A. Giles (1993), " Pre-test Estimation and Testing in Econometrics: Recent Developments ", J. Econ. Surveys, 7, 145-197.
[10] Goodman, J. (1953), " A Simple Method for Improving Some Estimators ", Ann. Math. Statist., 24, 114-117.
[11] Gruber, M. H. J. (1998), " Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimator ", Marcel Decker, New York.
[12] Judge, G. G. And T. A. Yancey (1986), " Improved Methods of Inference in Econometrics ", North Holland, Amestrdam.
[13] Lehmann, E. L. (1985), " Theory of Point Estimation ", Wiley, NewYork.
[14] Mehta, J. S. , and Srinivasan, R. (1971). " Estimation of The Mean by Shrinkage to a Point, ", JASA, 66, 86-90.
[15] Pandy, B. N. (1979), " On Shrinkage Estimation of Normal Population Variance ", Communication in Statistics. Theory, Mathematical. A8(4),359-365.
[16] Thompson, J. R. (1968), " Some Shrinkage Techniques for Estimating The Mean ", JASA, 63, 113-122.

.............................................................

...................................................

...........................

(ii) the restricted estimator $b_r$ coincide with the estimator $t(x)=0$ (the silly estimator)

4. The usual estimator $t(x)=x$ is meaningful and more reasonable and logical when we deal with the $N(\eta,1)$ problem, whereas the unrestricted estimator $b_u$ will have less meaning and less reasonability and logicality in the linear regression problem, because it involve the choice of $b_u$, whatever the value of the nuisance parameter $\gamma$ .
5. Because of the existence of the equivalence between the estimator $b_u$ and the usual estimator $t(x)=x$ in the $N(\eta,1)$ problem, we have to try to find an alternative
   estimator for it. If $t(x)=\lambda(x)x$ is the usual estimator, then the equivalence between the $N(\eta,1)$ problem and the regression problem will ensure the existence of the same function $\lambda$ (which is the optimal shrinkage function $\lambda^*$ of (43)), which provides WALS estimator of (29), without the need for depend on the regressor variable.
6. The estimators $b$, $b_u$ and $b_r$ are unbiased as it is clear from (20),(26) and (35) respectively.
7. The estimator $b_r$ is better than the estimator $b_u$, since $V(b_r)<V(b_u)$ according to (21) and (27).
8. The estimator $b$ will be better as long as $\lambda \to 0$ as it is clear from (29).

## REFERENCES


[1] Abadir, K. M. and J. R. Magnus (2002), " Notation in Econometrics: A Proposal for A Standard ", Econometrics Journal. Vol 5.

[2] Abadir, K. M. and J. R. Magnus (2003), " Normal's Deconvolution and The Independent of Sample Mean and Variance ", Econometric Theory, 19(4), 691-.

[3] Al-Zaidi , A. F. (2005), " A Comparison Study Between A New $N(\eta,1)$ with the (Normal Bayes, Pretest, Laplace Proposed Estimator for The Mean and Burr) Estimators", M.Sc. Thesis, College of Science, Al-Mustansiriya University.

[4] Berger, J.O. (1985), " Statistical Decision Theory and Bayesian Analysis ", 2nd edition, Springer-Verlag, New York.

[5] Blyth, C.R. (1951), " On Minimax Statistical Decision Procedures and their Admissibility ", Annals of Mathematical Statistics 22,22-42.

$$b = \lambda b_u + (1-\lambda)b_r = \lambda(b_r - d\hat{\eta}) + (1-\lambda)b_r = br - d\lambda\hat{\eta}$$

**Therefore,**

$$MSE(b) = V(b) = V(b_r) - dV(\lambda(\hat{\eta})\hat{\eta})(-d)' = V(b_r) + dMSE(\lambda(\hat{\eta})\hat{\eta})d' \quad (45)$$

**We conclude from (45) that**

$MSE(b) \propto MSE(\lambda(\hat{\eta})\hat{\eta})$

**That is,**

$MSE(b)$ **is minimized iff** $MSE(\lambda(\hat{\eta})\hat{\eta})$ **is minimized** (46)

**Also,** $\lambda = 0$ **and** $\lambda = 1$ **are natural endpoint (since** $0 \le \lambda \le 1$**) because they correspond (according to the equivalence between the** $N(\eta,1)$ **problem and the regression problem) to the restricted and unrestricted estimator, respectively. That is,** $\lambda = 0$ **correspond to** $b_r$ **estimator and** $\lambda = 1$ **correspond to** $b_u$ **estimator. In other words, if we can think of** $t(x)$ **as weighted average of** $x$ **(the usual estimator of (37)) and 0 (the silly estimator of (38)), then**

$$t(x) = \lambda(x)x + (1-\lambda(x))0 \quad (47)$$

**Finally, the larger is** $|x|$**, the better is** $x$ **as an estimator of** $\eta$**.**

## 8. CONCLUSIONS

1. **Depending on (46), we conclude that finding best WALS estimator for** $\beta$ **in the linear regression problem is equivalent to finding best estimator for** $\eta$ **in the** $N(\eta,1)$ **problem.**
2. **Also, we conclude that the regression problem will be solved if and only if the** $N(\eta,1)$ **problem is solved**
3. **The Equivalence that is gotten include the correspondence for any estimator of** $\eta$ **in the** $N(\eta,1)$ **problem, requires the existence of a unique estimator for** $\beta$ **in the linear regression problem. This conclusion is clear from (47) as it is shown in the following two points:**
   **(i) the unrestricted estimator** $b_u$ **coincide with the estimator** $t(x) = x$ **(the usual estimator) for** $\eta$ **in the** $N(\eta,1)$ **problem.**

admissible (because $t(x,0)=x$ is the usual estimator of (37)). For $0<c<\infty$, the result follow from Berger (1985), P. 127-28 [4].

Now, since $\lambda_c = 1/(1+c)$ (see (40)) and the optimal c is given by $c^* = 1/\eta^2$ (see (42)), we find the optimal $\lambda$ to be $\lambda^* = \lambda_{c^*} = \eta^2/(1+\eta^2)$, see (43).

The optimal $\lambda^*$, as function of $\eta$, thus satisfies $0 \le \lambda^*(\eta) \le 1$, and $\lambda^*(\eta)$ is an even function, that is, $\lambda^*(-\eta) = \lambda^*(\eta)$, and $\lambda^*$ is an increasing function on $(0,\infty)$.

For $c \ge 0$, the estimator $t(x,c)$ will be called the normal Bayes estimator of $\eta$, because it is the Bayes estimator induced by a normal prior with mean 0 and variance 1/c, ( for details see Al-Zaidi (2005) [3]).

## 7. THE EQUIVALENTS BETWEEN THE TWO PROBLEMS

In this section, we prove that the $N(\eta,1)$ problem is equivalent to the regression problem. This will be done if we can show that finding the best estimator of $\beta$ is equivalent to finding the best estimator of $\eta$ and this will be done if we can show that the $MSE(b)$ is minimized iff (if and only if) $MSE(\lambda(\hat{\eta})\hat{\eta})$ is minimized.

Now, from (12), (13), (14) and (19), then

$$b_u = (X'X)^{-1}X'y_u - (X'X)^{-1}X'z\frac{z'M}{z'Mz}y_u = b_r - (X'X)^{-1}X'z\hat{\gamma}$$

$$= b_r - (X'X)^{-1}X'z\frac{1}{z'Mz}\hat{\eta} = b_r - d\hat{\eta} \qquad (44)$$

where

$$d = (X'X)^{-1}X'z\frac{1}{z'Mz}$$

Now, using (44) in (29), then

**as a shrinkage function. Using a square error loss function, then the risk function is**

$$R(\eta,c) = R(\eta,\lambda_c) = E[L(t(x,\lambda),\eta)] = E[t(x,\lambda)-\eta]^2 = E_\eta(\frac{x}{1+c}-\eta)^2$$

$$= E_\eta\left[\frac{(x-\eta)-\eta c}{1+c}\right]^2 = \frac{1}{(1+c)^2}\left\{E(x-\eta)^2 - 2E[(x-\eta)\eta c] + E(\eta c)^2\right\}$$

$$= \frac{1}{(1+c)^2}\left\{V(x)+c^2\eta^2\right\} = \frac{1+c^2\eta^2}{(1+c)^2}$$

**Which coincide with the result of Goodman (1953) [10], and $R(\eta,c)$ is minimized when**

$$c^* = \frac{1}{\eta^2} \tag{42}$$

**with minimum risk**

$$R(\eta,c^*) = \frac{1+c^{*2}\eta^2}{(1+c^*)^2} = \frac{1+(\frac{1}{\eta^2})^2\eta^2}{(1+\frac{1}{\eta^2})^2} = \frac{\eta^2}{1+\eta^2}$$

**Therefore, the optimal shrinkage function $\lambda$ at $c^*$ is**

$$\lambda^* = \lambda_{c^*} = \frac{1}{1+c^*} = \frac{\eta^2}{1+\eta^2} \tag{43}$$

**(iii) For $c \geq 0$, then the estimator is admissible as it is shown now. Since $R(0,\lambda) \geq 0$ with equality if and only if $\lambda = 0$, we see that $t(x,\infty)$ dominates every other estimator at $\eta = 0$, and hence is admissible. Also, $t(x,0) = x$ (from (41)), is**

$$t(x) = x \tag{37}$$

and we call it the usual estimator, and (ii) when $\lambda(x) = 0$, then

$$t(x) = 0 \tag{38}$$

and we call it the silly estimator. Let the squared error loss function be defined as

$$L(t(x,\lambda),\eta) = [t(x,\lambda) - \eta]^2 \tag{39}$$

Then, the risk function of the usual estimator of (37) using the squared error loss function of (39) is

$$R(\eta,1) = E(x-\eta)^2 = V(x) = 1$$

Hence, the usual estimator $t(x) = x$ is unbiased and has constant risk (variance) equal to 1. Blyth (1951) [5], showed that x is admissible (see also Berger (1985), P.548) [4].

Now, since x is admissible and has constant risk, it must be minimax Berger (1985), P.382 [4]. This means that the usual estimator $t(x) = x$ is unbiased, admissible, has constant risk (variance) equal to 1 and minimax.

These are strong properties in favor of x as an estimator of $\eta$. For this reason, we might want to choose an estimator different from x. Define

$$\lambda_c(x) = \frac{1}{1+c}, \quad c \neq -1 \tag{40}$$

as a shrinkage estimator for all x, where c is a constant. So that, in particular, from (40) we have the following cases

(i) For $c = 0$, then

$\lambda_0(x) = 1$

(ii) For $c = \infty$, then

$\lambda_\infty(x) = 0$

Now, consider

$$t(x,c) = t(x,\lambda_c(x)) = \frac{x}{1+c} \tag{41}$$

**Now, using (26), (30), (32) and (34) then**

$$E(b) = \lambda.0 + \beta \ = \beta \qquad (35)$$

**and**

$$V(b) = \{\lambda(A_1 - A_2) + A_2\}\sigma^2 I_n \{\lambda(A_1 - A_2) + A_2\}'$$

$$= \sigma^2 \left\{ -\lambda(X'X)^{-1} X'zz' \frac{M}{z'Mz} + (X'X)^{-1}X' \right\} \left\{ -\lambda \frac{M}{z'Mz} zz'X(X'X)^{-1} + X(X'X)^{-1} \right\}$$

$$= \sigma^2 \left\{ \lambda^2 (X'X)^{-1} X'zz' \frac{M}{z'Mz} \frac{M'}{z'Mz} zz'X(X'X)^{-1} - \lambda(X'X)^{-1} X'zz' \frac{M}{z'Mz} X(X'X)^{-1} - \lambda(X'X)^{-1} X' \frac{M}{z'Mz} zz'X(X'X)^{-1} + (X'X)^{-1} X'X(X'X)^{-1} \right\}$$

$$= \sigma^2 (X'X)^{-1} + \lambda^2 \frac{\sigma^2}{z'Mz} (X'X)^{-1} X'zz'X(X'X)^{-1} \qquad (36)$$

**When we compare (36) with (21), we note that the estimator $b$ is better than the estimator $b_u$ for all $\lambda$ except when $\lambda = 0.1$. When $\lambda = 0$ then $V(b) = V(b_r)$ which coincide with (27) and when $\lambda = 1$ then $V(b) = V(b_u)$, which coincide with (21).**

## 6. THE USUAL ESTIMATOR

**Let $x \sim N(\eta, 1)$, and let**

$$t(x, \lambda) = \lambda(x)x$$

**be an estimator of $\eta$ . (i) when $\lambda(x) = 1$, then**

$$b_r = A_2 y_r \quad , \quad E(b_r) = \beta \quad , \quad V(b_r) = \sigma^2 (X'X)^{-1}$$

**where $A_2$ is given in (25), and from (13), (20) and (21), we know that**

$$b_u = A_1 y_u \quad , \quad E(b_u) = \beta \quad ,$$

$$V(b_u) = \sigma^2 (X'X)^{-1} + \frac{\sigma^2}{z'Mz}(X'X)^{-1} X'zz'X(X'X)^{-1}$$

**where $A_1$ is given in (14). Then, from (13), (24), (2) and (3) we have**

$$b_u - b_r = A_1 y_u - A_2 y_r = (A_1 - A_2)X\beta + \gamma A_1 z + (A_1 - A_2)\varepsilon \tag{31}$$

**Now, using (20 ) and (26 ), we get**

$$E(b_u - b_r) = \beta - \beta = 0 \tag{32}$$

**and using (14), (25) and (31), we get**

$$V(b_u - b_r) = (A_1 - A_2)V(\varepsilon)(A_1 - A_2)' = \sigma^2 (A_1 - A_2)(A_1 - A_2)'$$

$$= \sigma^2 \left[ -(X'X)^{-1} X'zz' \frac{M}{z'Mz} \right] \left[ \frac{-M'}{z'Mz} zz'X(X'X)^{-1} \right]$$

$$= \frac{\sigma^2}{z'Mz}(X'X)^{-1} X'zz'X(X'X)^{-1} \tag{33}$$

**from (32) and (33), we have**

$$b_u - b_r \sim N\left(0, \frac{\sigma^2}{z'Mz}(X'X)^{-1} X'zz'X(X'X)^{-1}\right)$$

**Substituting (24) and (31) in (30) gives**

$$b = \lambda(A_1 - A_2)X\beta + \gamma\lambda A_1 z + \lambda(A_1 - A_2)\varepsilon + A_2 X\beta + A_2\varepsilon$$

$$= \{\lambda(A_1 - A_2) + A_2\}X\beta + \gamma\lambda A_1 z + \{\lambda(A_1 - A_2) + A_2\}\varepsilon \tag{34}$$

where $k(\hat{\theta})$ is a constant function in $\hat{\theta}$ and $0 \le k(\hat{\theta}) \le 1$. In the notation of this work, (28) is written as

$$b = \lambda b_u + (1-\lambda) b_r \tag{29}$$

where $\lambda = \lambda(\hat{\eta})$ is a constant function in $\hat{\eta}$ and $0 \le \lambda \le 1$. Then the shrinkage estimator considers $\lambda$ as amount of confident for the estimator $b_u$ and $(1-\lambda)$ as amount of confident for the estimator $b_r$. The estimator $b$ of (29) can be rewritten as

$$b = \lambda (b_u - b_r) + b_r \tag{30}$$

The estimator $b$ of (29) is called the WALS (weighted average least squares).

The shrinkage estimators are of two types

(1) Constant weighted shrinkage function (which do not depend upon the parameter to be estimated).

(2) Nonconstant weighted shrinkage function (which depend upon the parameter to be estimated), i.e., $\lambda = \lambda(\hat{\eta})$.

The choice of the type does not depend on any condition, it is up to the aim and desire of the researcher. In both cases the choice of the shrinkage function is done almost arbitrary, and sometimes the choise of the shrinkage function is done depending on the prior informations.

In the following section, we will study the properties of the estimator $b$ when the shrinkage function is of the first type.

## 5. THE PROPERTIES OF THE ESTIMATOR $b$

In this section, we study the properties of the estimator $b$ when it is of the first type, that is, when we have a constant weighted shrinkage function. The properties that we study are

(1) The Bias

(2) The mean squared error (MSE)

From (29), we note that the estimator $b$ depends upon the estimators $b_u$ and $b_r$. To study the properties of the estimator $b$, we have to know the properties of the estimators $b_u$, $b_r$ and $b_u - b_r$. From (24), (26) and (27) we know that and

$$V(b_r) = (X'X)^{-1} X'V(y_r) X(X'X)^{-1} = \sigma^2 (X'X)^{-1} \tag{27}$$

hence, $b_r$ is unbiased estimator of $\beta$ and

$$b_r \sim N(\beta, \sigma^2 (X'X)^{-1})$$

## 4. THE SHRINKAGE ESTIMATOR $\bar{b}$

Although that the estimated regression parameters by the OLS method are unbiased, but still researchers may face a problem because in these estimators some difficulties of certain types could appear. For example, the researcher could make a mistake not on purpose, like adopting a small sample, the result of this mistake could imply that the resulting estimator be away from the true value of the parameter that we want to estimate or these estimators could have big value and consequently be away from the true value, too.

The classical methods for estimations with the OLS method included assume that the parameter which we want to estimate is unknown and we do not know any think about it. We do the estimation process by drawing a sample out of the population to estimate the unknown parameter. Thus, the classical methods do not depend on prior works or on the researchers knowledge. This problem is well known and it is not new. It was discussed by many researchers and they have found many solutions for it. One of these solutions is using shrinkage estimator (see [4], [6], [11], [14], [15], [16] ). The shrinkage estimator is one of the methods which takes in account the use of the prior informations about the vector of unknown parameters which we want to estimate. These informations could be as initial values $\theta_0$ and estimated values $\hat{\theta}$ , computed from a small random sample by any of the classical methods. Then the final estimator would be a shrinkage estimator which is a linear combination of $\theta_0$ and $\hat{\theta}$. The shrinkage estimator will take the form of the weighted shrinkage function and is denoted by $\tilde{\theta}$ and is written as

$$\tilde{\theta} = k(\hat{\theta})\hat{\theta} + (1 - k(\hat{\theta})\theta_0 \tag{28}$$

$$+(X'X)^{-1}X'z\frac{z'M}{(z'Mz)}\frac{Mz}{(z'Mz)}z'X(X'X)^{-1}\Bigg]$$

$$=\sigma^2(X'X)^{-1}+\frac{\sigma^2}{z'Mz}(X'X)^{-1}X'zz'X(X'X)^{-1} \quad (21)$$

From (13), we note that $b_u$ is a linear function in $y_u$ and $y_u$ is distributed normally according to (16), then from (20), (21) and (16), the distribution of $b_u$ is

$$b_u \sim N(\beta, V(b_u)).$$

The estimator $\hat{\gamma}$ of (12) and the estimator $b_u$ of (13) are ordinary least squares (OLS) estimators for the parameters $\gamma$ and $\beta$ in the unrestricted model of (2).

**3.3 The distribution of $b_r$:** Now, letting $\gamma = 0$ in (2), then the model becomes

$$y_r = X\beta + \varepsilon \quad , \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (22)$$

and

$$y_r \sim N(X\beta, \sigma^2 I_n) \quad (23)$$

the model of (22) is restricted model, and the OLS estimator of $\beta$ is

$$b_r = A_2 y_r \quad (24)$$

where

$$A_2 = (X(X)^{-1}X' \quad (25)$$

Using (22), (23), (24) and (25), then the mean and variance of $b_r$ are

$$E(b_r) = (X'X)^{-1}X'E(y_r) = (X'X)^{-1}X'X\beta = \beta \quad (26)$$

Let

$$\eta = \frac{\gamma}{\sigma / \sqrt{z'Mz}}$$

then

$$\hat{\eta} = \frac{\hat{\gamma}}{\sigma / \sqrt{z'Mz}} \quad \text{and} \quad \hat{\eta} \sim N(\eta, 1) \tag{19}$$

where $\eta$ represents the non-central parameter associated with t statistic for testing that $\gamma=0$.

In this paper, we assumed that $\sigma^2$ is known which is mentioned earlier, hence the value of the parameter $\eta$ will depend on the value of the parameter $\gamma$, i.e. $\eta \propto \gamma$.

**3.2 The distribution of $b_u$:** From (13), (14), (15) and (11), we have

$$E(b_u) = A_1 E(y_u)$$

$$= \left[ (X'X)^{-1} X' - (X'X)^{-1} X'z \frac{z'M}{z'Mz} \right] (X\beta + \gamma\, z)$$

$$= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\gamma\, z - \gamma (X'X)^{-1} X'z \frac{z'Mz}{z'Mz} = \beta \tag{20}$$

that is, $b_u$ is unbiased estimator of $\beta$, and

$$V(b_u) = A_1 V(y_u) A_1'$$

$$= \left[ (X'X)^{-1} X' - (X'X)^{-1} X'z \frac{z'M}{z'Mz} \right] \sigma^2 I_n \left[ (X'X)^{-1} X' - (X'X)^{-1} X'z \frac{z'M}{z'Mz} \right]'$$

$$= \sigma^2 \left[ (X'X)^{-1} - (X'X)^{-1} X' \frac{Mz}{z'Mz} z'X(X'X)^{-1} - (X'X)^{-1} X'z \frac{z'M}{z'Mz} X(X'X)^{-1} \right.$$

$$b_u = A_1 y_u \tag{13}$$

where

$$A_1 = (X'X)^{-1}X'\left(\mathrm{I} - \mathrm{z}\frac{\mathrm{z'M}}{\mathrm{z'Mz}}\right) = (X'X)^{-1}X' - (X'X)^{-1}X'z\frac{z'M}{z'Mz} \tag{14}$$

and $b_u$ is also a linear function in the unrestricted variable $y_u$.

## 3. DISTRIBUTION OF $\hat{\gamma}$ , $b_u$ و $b_r$

To find the distribution of $\hat{\gamma}$, $b_u$ and $b_r$ we need to know the distribution of $y_u$. From (2), we know that

$$E(y_u) = X\beta + \gamma\ z \quad , \qquad V(y_u) = \sigma^2 I_n \tag{15}$$

then

$$y_u \sim N(X\beta + \gamma\ z,\ \sigma^2 I_n) \tag{16}$$

**3.1. The distribution of $\hat{\gamma}$:** From (11), (12) and (15), we have

$$E(\hat{\gamma}) = \frac{z'M}{z'Mz}E(y_u) = \frac{z'M}{z'Mz}(X\beta + \gamma\,\mathrm{z}) = \gamma \tag{17}$$

that is, $\hat{\gamma}$ is unbiased estimator of $\gamma$, and

$$V(\hat{\gamma}) = \frac{z'M}{z'Mz}V(y_u)\frac{M'z}{z'Mz} = \frac{z'M}{z'Mz}\sigma^2 I_n \frac{M'z}{z'Mz} = \sigma^2\frac{z'Mz}{(z'Mz)^2} = \frac{\sigma^2}{z'Mz} \tag{18}$$

Since $\hat{\gamma}$ is a linear function in $y_u$, see (12) then from (16), (17) and (18), we have

$$\hat{\gamma} \sim N(\gamma\ ,\frac{\sigma^2}{z'Mz})$$

$$X'y_u - \hat{\gamma}X'z - X'Xb_u = 0 \tag{6b}$$

**Multiplying (6a) by $(z'z)^{-1}$ , then**

$$\hat{\gamma} = (z'y_u - b_u'X'z)(z'z)^{-1} \tag{7}$$

**and premultiplying (6b) by $(X'X)^{-1}$ and using (7) , we get**

$$b_u = (X'X)^{-1}X'y_u - (X'X)^{-1}X'z\hat{\gamma} \tag{8}$$

**substituting (8) in (7) gives**

$$\hat{\gamma} = \left[ z'y_u - \left\{ (X'X)^{-1}(X'y_u - X'z\hat{\gamma}) \right\}' X'z \right](z'z)^{-1}$$
$$= z'y_u(z'z)^{-1} - y_u'X(X'X)^{-1}X'z(z'z)^{-1} + \hat{\gamma}\, z'X(X'X)^{-1}X'z(z'z)^{-1}$$
$$= z'y_u(z'z)^{-1} - y_u'z(z'z)^{-1} + y_u'Mz(z'z)^{-1} + \hat{\gamma} - \hat{\gamma}z'Mz(z'z)^{-1} \tag{9}$$

**where**

$$M = I - X(X'X)^{-1}X' \tag{10}$$

**and from (10), we note that**

$$M = M', M^2 = M, MX = 0, X'M = 0 \tag{11}$$

**that is, matrix $M$ is symmetric and idempotent. Now, since $z'y_u = y_u'z$ , then from (9), $\hat{\gamma}$ is**

$$\hat{\gamma} = \frac{y_u'Mz}{z'Mz} = \frac{z'M}{z'Mz}y_u \tag{12}$$

**where $z'My_u = y_u'Mz$ and then $\hat{\gamma}$ is a linear function in the unrestricted variable $y_u$ . Thus, substituting (12) in (8), we get**

**without affecting the main results. The existence of the term $\gamma z$ in the model (1) make the model unrestricted. For this, we write model (1) as**

$$y_u = X\beta + \gamma z + \varepsilon \ , \quad \varepsilon \sim N(0, \sigma^2 I_n) \tag{2}$$

**the estimation of the parameters vector $\beta$ will be written as $b_u$, where u stands for unrestricted model. In the case, when (γ=0 ), i.e. the term $\gamma z$ is not in the model, then model (1) will be restricted model and is written as**

$$y_r = X\beta + \varepsilon \quad , \quad \varepsilon \sim N(0, \sigma^2 I_n) \tag{3}$$

**the estimation of the parameters vector $\beta$ will be written as $b_r$, where r stands for restricted model. From (2), we note that the error sum of square is**

$$Q = \varepsilon'\varepsilon = (y_u - X\beta - \gamma z)'(y_u - X\beta - \gamma z)$$

$$= y_u' y_u - y_u' X\beta - \gamma y_u' z - \beta' X' y_u + \beta' X' X\beta + \gamma \beta' X'z - \gamma z' y_u +$$

$$\gamma z' X\beta + \gamma^2 z'z \tag{4}$$

**where $y_u' X\beta = \beta' X' y_u$ and $z' X\beta = \beta' X' z$. Then $b_u = \hat{\beta}$ and $\hat{\gamma}$ can be gotten from differentiating (4) with respect to $\beta$ and $\gamma$ respectively, as follows**

$$\frac{\partial Q}{\partial \gamma} = 0$$
$$\frac{\partial Q}{\partial \beta} = 0 , \tag{5}$$

**Then, (5) will yield the following normal equations**

$$z y_u - b_u' X' z - \hat{\gamma}\, z'z = 0 \tag{6a}$$

difficulties and it concentrates on a chosen weighted shrinkage function for this situation. For this, many research works have been done for shrinkage function ( see [6], [11], [14], [15], [16] ) to gain advantage for finding good estimators out of a very small sample which could be a single observation which is the aim of this study. In other words, as long as the prior information become less or the random sample become smaller, the probability of getting estimator away from its true value for the unknown parameter become larger. For this reason, the need for the shrinkage method increased, especially for the optimal shrinkage estimator which yield optimal estimator.

One of the common important statistical problems in Econometric is the regression problem (i.e. the problem of estimating the parameter vector $\beta$ in the multiple linear regression model) ( see [1], [2], [3], [8], [9], [12] ) which has the form

$$y = X\beta + \gamma\, z + \varepsilon \;, \quad \varepsilon \sim N(0, \sigma^2 I_n) \tag{1}$$

where $y$ is an $n \times 1$ vector of response variables, $X$ is an $n \times p$ matrix of regressor (explanatory) variables, z is an $n \times 1$ vector of additional explanatory variables, $\beta$ is an $p \times 1$ vector of unknown parameters, $\gamma$ is a nuisance parameter and $\varepsilon$ is the $n \times 1$
vector of random errors, and the problem of estimating the mean $\eta$ of the univariate normal distribution which is called as the $N(\eta,1)$ problem for single observation . In this paper, we consider the shrinkage functions for finding an estimator for the parameter without the need to depend completely on the explanatory variables, and considering estimators for the $N(\eta,1)$ problem depends upon single observation, we try to prove that the two problems are equivalent and try to prove that solving any problem will lead for solving the other one.

## 2. ESTIMATING THE REGRESSION MODEL'S PARAMETERS

Assume that the matrix $(X : z)$ has Full Column Rank and the variance $\sigma^2$ is known. We have to note that the last assumption about $\sigma^2$ to be known is unrealistic assumption, but it simplifies the analysis

## ABSTRACT

In this paper, we study the estimation of the unknown mean parameter $\eta$ of a univariate normal distribution when the variance is known ($\sigma^2 = 1$) and the sample is
of size one. The main estimation methods have been discussed along with their properties and the properties of the estimators. The focus is concentrated on the shrinkage method for finding the estimators. We wish to prove that there is a relationship between the shrinkage estimator $b$ (the estimator of the parameter vector $\beta$ in the linear regression model) and $\eta$ of $N(\eta,1)$, where for each estimator $t(x) = \lambda(x)x$ for $N(\eta,1)$ problem, there will be a corresponding WALS (weighted average least squares) estimator in the regression problem of the form $b = \lambda(\hat{\eta})b_u + (1-\lambda(\hat{\eta}))b_r$.


Key words and phrases: mean squared error criterion, model selection, regression anlysis, shrinkage estimator, univariate normal mean.

## 1. INRODUCTION

There are many methods for estimation (Classical methods, Bayes methods, Shrinkage method) ( see [4], [7], [13] ). For each method, there are some problems and there are solutions for these problems. The idea of shrinkage depends on combining the classical and the Bayes methods, which takes $\theta_0$ as an initial value and takes $\hat{\theta}$ as an estimator from a small sample. The final estimator will be a linear combination of $\theta_0$ and $\hat{\theta}$ depending on a weighted shrinkage function which is given the symbol $\lambda$, which represents a confident value for $\hat{\theta}$ and $(1-\lambda)$ as a confident value for $\theta_0$ and $0 \le \lambda \le 1$. Thus, the shrinkage estimator has the form $\tilde{\theta} = \lambda\hat{\theta} + (1-\lambda)\theta_0$.

The shrinkage estimators are very efficient when they are compared with the classical estimators. Definitely, there will be no prior informations about the parameters, which we want to estimate and there will be no information for the random sample which we depend on. Hence, we note that the estimator value will be away from its true value. This is one of the negative points against the classical and Bayes methods. The shrinkage method has important advantage because it reduces the

# EQUIVALENCE OF THE $N(\eta,1)$ PROBLEM AND THE REGRESSION PROBLEM WITH ONE NUISANCE PARAMETER

# تكافؤ مسألة $N(\eta,1)$ و مسألة الانحدار مع معلمة مقلقة واحدة

سامي داود كبارة *
عامر فاضل الزيدي * *

**المستخلص**

في هذا البحث تم دراسة تقدير معلمة المعدل غير المعلومة $\eta$ في التوزيع الطبيعي الاحادي عندما يكون التباين معلوم $(\sigma^2 = 1)$ و العينة ذات حجم واحد . تم مناقشة طرق التقدير الرئيسة مع خواصها و خواص مقدراتها . تم التركيز على طريقة التقلص لايجاد المقدرات. كما تم أثبات وجود علاقة بين مقدر التقلص $b$ (مقدر متجه المعلمات $\beta$ في نموذج الانحدار الخطي) و $\eta$ من التوزيع $N(\eta,1)$ , بحيث لكل مقدر $t(x) = \lambda(x)x$ لمسألة $N(\eta,1)$ , يوجد مقدر متوسط أصفر المربعات الموزن (WALS) مناظر في مسألة الانحدار من الصيغة $b = \lambda(\hat{\eta})b_u + (1-\lambda(\hat{\eta}))b_r$.

* استاذ مساعد / الجامعة المستنصرية / كلية العلوم/ قسم الرياضيات

* * مدرس مساعد / الجامعة المستنصرية / كلية العلوم/ قسم الرياضيات

مقبول النشر بتاريخ 2006/2/28