Values of Asymptotic Efficiency in Linear Regression Models Concepts and Applications

Adel Ahmed Haddaw
Isra University – Jordan- Amman

ABSTRACT

The purpose of this paper is to derive asymptotic efficiency formula in which an error variable ε has some probability density function in linear regression models such as normal, extreme- value for largest values and logistic distribution. The variances of the regression coefficients are estimated by ordinary least squares estimator and maximum likelihood estimator in order to obtain variance efficiencies. Some applications of asymptotic efficiency without simulation and with simulation have been done. Asymptotic efficiency values without simulation for the regression coefficient β_1 were (1, 0.608, 0.912) for normal, extreme-value and logistic distributions respectively.

1- Introduction

The linear regression model, together with variances of regression coefficients of the least- squares estimator and maximum likelihood estimator, plays a fundamental role in derivation of asymptotic efficiency formula (Cox and Hinckley, 1968).

Consider the multiple linear regression model (Jin, Z.L et al, 2006) and (Draper and Smith, 1982) and (Lawless, 1982).

$$Y_i = \mu_i + \varepsilon_i = \underline{x_i}^{`} \underline{\beta} + \varepsilon_i , \quad i = 1, 2, \ldots n \tag{1}$$

Where $\underline{x_i}^{`} = (1, x_{i1}, x_{i2}, \ldots x_{ik})$ , $\underline{\beta}^{`} = (\beta_0 , \beta_1 , \ldots, \beta_k)$, the values $x_{i1}, x_{i2}, \ldots x_{ik}$, representing observations on k non-random explanatory variables for the i^{th} individual and ε_i random errors.

We shall assume that ε has independently and identically distributed with

$$E(\varepsilon_i) = 0 \quad , \quad v(\varepsilon_i) = \sigma^2 , \ i = 1, 2 \ldots, n \tag{2}$$

Assuming that $\underline{x}$ is of rank k+1, the OLS estimator is given by

$$\underline{\beta}^{\sim} = (\underline{x}^{/} \ \underline{x})^{-1} \ \underline{x}^{`} \ \underline{Y} \tag{3}$$

With

$$E(\underline{\beta}^{\sim}) = \underline{\beta} \quad , \quad \mathrm{COV}(\underline{\beta}^{\sim}) = \sigma^2 (\underline{x}^{/} \ \underline{x})^{-1} \tag{4}$$

Where $\underline{x}$ is the design matrix with (n (k+1)) .

The OLS estimators of $\underline{\beta}$, denoted by $\underline{\beta}^{\sim}$ and the ML estimators of $\underline{\beta}$, denoted by $\underline{\beta}^{\wedge}$. Thus asymptotic efficiency for OLSE formula is

$$E_r = V(\beta_r^{\wedge}) / V(\beta_r^{\sim}) = \{ A_{\epsilon} v(\epsilon) \} \ , \ r = 0, 1, 2, \ldots k \tag{5}$$

Where E_r are asymptotic efficiency for $(\beta_0 , \beta_1 , \ldots, \beta_k)$

2- Derivative of asymptotic efficiency formula

In this section, we will derive asymptotic efficiency formula for some probability density function in linear regression models such as (Normal, Extreme, Logistic distributions).

Statistically the general mean is very important in linear model, if we consider that the expiration without loss of generality, we will assume that

$$x_{i0} = 1 , \quad i = 1, 2, \ldots, n \quad , \sum_{i=1}^{n} x_{ir} = 0 , \ r = 1, 2, \ldots, k \tag{6}$$

Also, statistically the orthogonal is important to the general mean. In this paper we will consider that regression coefficients $(\beta_1 , \ldots, \beta_k)$ so that to derive asymptotic efficiency formulae.

If we have the probability density function (p.d.f) of errors (ε_i) , denoted by $f(\varepsilon , \underline{\lambda})$ .Where $\underline{\lambda}$ are unknown parameters, representing dispersion parameter such as scale parameter or shape parameter or location parameter of underline distribution.

If we refer to equation (1), the log likelihood function is

$$L = \sum_{i=1}^{n} \log f(Y_i - \mu i, \underline{\lambda}) = \sum_{i=1}^{n} g(Y_i - \mu i, \underline{\lambda}) \qquad (7)$$

To get the expected values for the second partial derivative of the log likelihood function as follows:

$$\partial L / \partial \beta_r = \sum_{i=1}^{n} g^{/}(Y_i - \mu i, \underline{\lambda}) \, \partial \mu i / \beta_r$$

And

$$\partial^2 L / \partial \beta_r \beta_s = \sum_{i=1}^{n} g^{//}(Y_i - \mu i, \underline{\lambda}) \, \partial \mu i / \beta_r \, \partial \mu i / \beta_s$$

$$= \sum_{i=1}^{n} g^{//}(Y_i - \mu i, \underline{\lambda}) \, x_{ir} x_{is} \qquad (8)$$

Let $\underline{\lambda}$ are fixed, then the expectation will be as follows:

$$E(-\partial^2 L / \partial \beta_r \beta_s) = (\sum_{i=1}^{n} x_{ir} x_{is}) A_{\epsilon} \qquad (9)$$

Where

$$A_{\epsilon} = E\{-g^{//}(\epsilon, \underline{\lambda})\}$$

$$= -\int_{-\infty}^{\infty} f(\epsilon, \underline{\lambda}) \, \partial^2 \operatorname{Log} f(\epsilon, \underline{\lambda}) / \partial \epsilon^2 \, d\epsilon \qquad (10)$$

Also

$$\partial^2 L / \partial \beta_r \lambda_u = \sum_{i=1}^{n} \partial / \partial \beta_r \{\partial g(Y_i - \mu i, \underline{\lambda}) / \partial \lambda_u\}$$

$$= \sum_{i=1}^{n} \partial / \partial \epsilon_i \{\partial g(\epsilon_i, \underline{\lambda}) / \partial \lambda_u\} \, \partial \epsilon_i / \partial \beta_r$$

$$= -\sum_{i=1}^{n} \{\partial^2 g(\epsilon_i, \underline{\lambda}) / \partial \epsilon_i \, \partial \lambda_u\} \, x_{ir} \qquad (11)$$

Since

$E\{\partial^2 g(\epsilon_i, \underline{\lambda}) / \partial \epsilon_i \, \partial \lambda_u\}$ is the same for each i and $\sum_{i=1}^{n} x_{ir} = 0$

We get

$$E(\partial^2 L / \partial \beta_r \lambda_u) = 0 \quad , r = 0, 1, \ldots, k \ , u = 1, 2, \ldots, v \qquad (12)$$

Using (8), we get the following

$$E(\partial^2 L / \partial \beta_r \lambda_u) = \sum_{i=1}^{n} x_{ir} E\{g^{//}(\epsilon_i, \underline{\lambda})\} \qquad (13)$$

$$= 0$$

Because $E\{g^{//}(\epsilon_i, \underline{\lambda})\}$ is a constant and $\sum_{i=1}^{n} x_{ir} = 0$.

Using (12) and (13), we can obtain the information matrix as follows:

$$\underset{(v+k+1)(v+k+1)}{I} = \begin{Bmatrix} \underline{I}_1 & \underline{0} \\ \underline{0} & \underline{I}_2 \end{Bmatrix} \qquad (14)$$

Where $\underline{I}_1$ belong to β_0 and $\underline{\lambda}$, $\underline{I}_2$ belong to $\beta_1, \ldots, \beta_k$. From (9), we obtain (r, s)th element in $\underline{I}_2$.

We consider orthogonal condition that showed in (6), so that the covariance matrix for OLSE ($\beta^{\sim}_1$, $\beta^{\sim}_2$, … $\beta^{\sim}_k$) as follows :

$$V(\underline{\beta}^{\sim}) = \{\sum_{i=1}^{n} x_{ir} x_{is}\}^{-1} V(\epsilon) \tag{15}$$

Also the inverse of the information matrix as follows:

$$\underline{I}^{-1} = \begin{Bmatrix} \underline{I}_1^{-1} & \underline{0} \\ \underline{0} & \underline{I}_2^{-1} \end{Bmatrix}$$

Whereas the covariance matrix for MLE ($\beta^{\wedge}_1$, $\beta^{\wedge}_2$, … $\beta^{\wedge}_k$) as follows:

$$V(\underline{\beta}^{\wedge}) = \underline{I}_2^{-1} = \{\sum_{i=1}^{n} x_{ir} x_{is}\}^{-1} \; A_{\epsilon}^{-1} \tag{16}$$

Thus, asymptotic efficiency formula is

$$E_r = V(\beta_r^{\wedge}) / V(\beta_r^{\sim}) = \{ A_{\epsilon} v(\epsilon) \}^{-1} \;, r = 0,1,2,...k \tag{17}$$

Where E_r are asymptotic efficiency for (β_0 , β_1 , …., β_k)

3- Application Side (Results and Discussion)

In this section, there are two applications as follows:

3-1 Application without Simulation

This section deal with some applications of asymptotic efficiency formula (17) without simulation when ε has independently and identically distributed with the following distributions:

1- **Normal Distribution** (Jin, Z.L et al, 2006) and see (Haddaw, 2014)

The p.d.f. of the errors ε being

$$f(\varepsilon) = 1 / \sqrt{2\pi}\sigma \exp(-½ \epsilon^2 / \sigma^2) \quad , \quad -\infty < \epsilon < \infty \tag{18}$$

The log likelihood function is

$$\text{Log } f(\varepsilon) = -½ \log(2\pi) - \log\sigma - (-½ \epsilon^2 / \sigma^2) ,$$

$$\partial \text{ Log } f(\varepsilon) / \partial \varepsilon = -\varepsilon / \sigma^2 \;, \partial^2 \text{ Log } f(\varepsilon) / \partial \varepsilon^2 = -1 / \sigma^2$$

Thus $A_{\epsilon} = 1/\sigma^2$, $v(\epsilon) = \sigma^2$

Thus, asymptotic efficiency formula is

$$V(\beta^{\wedge}_r) / V(\beta^{\sim}_r) = \{ A_{\epsilon} V(\epsilon) \}^{-1} = (1/\sigma^2 \; \sigma^2)^{-1} = 1 \,, r=1, 2,…k \tag{19}$$

Statistically, the OLS estimators $\underline{\beta}$~ and the ML estimators $\underline{\beta}$^ for the normal case are the same.

2- A Type 1 Extreme Value Distribution for the largest values (Haddaw and Young,1986) , the p.d.f. of the errors ε being

$$f(\varepsilon) = 1/\theta \exp\{-(\epsilon/\theta + \gamma) - \exp(-(\epsilon/\theta + \gamma)\} \quad , \quad -\infty < \epsilon < \infty \tag{20}$$

Where $\gamma = 0.57722$, $v(\epsilon) = 1/6\pi^2 \theta^2$, θ is scale parameter of the distribution.

The log likelihood function is

$$\text{Log } f(\varepsilon) = -\log\theta - \epsilon/\theta - \gamma - \exp(-(\epsilon/\theta + \gamma)$$

$$\partial \text{ Log } f(\varepsilon) / \partial \varepsilon = -1/\theta + 1/\theta \exp(-(\epsilon/\theta + \gamma)$$

$$\partial^2 \text{ Log } f(\varepsilon) / \partial \varepsilon^2 = -1/\theta^2 \exp(-(\epsilon/\theta + \gamma)$$

Then

$$A_{\epsilon} = -\int_{-\infty}^{\infty} 1/\theta \exp\{-(\epsilon/\theta+\gamma) - \exp(-(\epsilon/\theta+\gamma)\}\{-1/\theta^2 \exp(-(\epsilon/\theta+\gamma)\}\partial\varepsilon$$

$$= 1/\theta^3 \int_{-\infty}^{\infty} \exp[-2(\epsilon/\theta+\gamma) - \exp(-(\epsilon/\theta+\gamma)]\, d\varepsilon$$

Using integral by part

$u = \exp(-(\epsilon/\theta+\gamma)$

$du = -1/\theta \exp\{-(\epsilon/\theta+\gamma)\}\, d\varepsilon$, we obtain

$$A_{\epsilon} = \int_{0}^{\infty} u \exp(-u)\, du = 1/\theta^2$$

Thus, asymptotic efficiency formula is

$$V(\beta^{\wedge}_{r})/V(\beta^{\sim}_{r}) = \{A_{\epsilon} V(\epsilon)\}^{-1} = (1/\theta^2\ 1/6\pi^2\theta^2)^{-1} = 6/\pi^2 = 0.608, \quad (21)$$

r=1, 2,…k

3- Logistic Distribution (Al-Sarraf, 1986)
The p.d.f. of the errors ε being

$$f(\varepsilon) = 1/\theta\ \{\exp(-\epsilon/\theta)/(1+\exp(-\epsilon/\theta)^2\} \quad , \quad -\infty < \epsilon < \infty \quad (22)$$

$v(\epsilon) = \pi^2\theta^2/3$

The log likelihood function is

$$\text{Log } f(\varepsilon) = -\log\theta - \epsilon/\theta - 2\log(1+\exp(-\epsilon/\theta)$$

$$\partial \text{Log } f(\varepsilon)/\partial\varepsilon = -1/\theta + \{2\exp(-\epsilon/\theta)/\theta(1+\exp(-\epsilon/\theta)\}$$

$$\partial^2 \text{Log } f(\varepsilon)/\partial\varepsilon^2 = -2/\theta^2\, 2\exp(-\epsilon/\theta)/\theta(1+\exp(-\epsilon/\theta)\} + 2\exp(-\epsilon/\theta)/\theta\{1/\theta\ \{\exp(-\epsilon/\theta)/(1+\exp(-\epsilon/\theta)^2\}$$
$$= -2\exp(-\epsilon/\theta)/\theta^2(1+\exp(-\epsilon/\theta)^2$$

$$A_{\epsilon} = 2/\theta^3 \int_{-\infty}^{\infty} \exp(-\epsilon/\theta)/\ 1+\exp(-\epsilon/\theta)^2\, d\epsilon$$

$$= 2/\theta^3 \int_{-\infty}^{\infty} \exp(-2\epsilon/\theta)/\ 1+\exp(-\epsilon/\theta)^4\, d\epsilon$$

$$= 2/\theta^2 \int_{-\infty}^{\infty} \exp(-2x)/\ 1+\exp(-x)^4\, dx$$

$$= 2/\theta^2 \int_{0}^{\infty} v\, dv\ /(1+v)^4$$

$$= 2/\theta^2\{[v(1+v)\ /-3]_{0}^{\infty} + 1/3\int_{0}^{\infty}(1+v)^{-3}\, dv\}$$

$$= 2/\theta^2\,[0 - 1/6(1+v)^{-2}]_{0}^{\infty} = 1/3\,\theta^2$$

Thus, asymptotic efficiency formula is

$$V(\beta^{\wedge}_{r})/V(\beta^{\sim}_{r}) = \{A_{\epsilon} V(\epsilon)\}^{-1} = \{1/3\,\theta^2\,\pi^2\theta^2/3\}^{-1} = 9/\pi^2 = 0.912, \quad (23)$$

r=1, 2,…k

3-2 Application with Simulation

This section deal with some applications of asymptotic efficiency formula (17) with simulation when ε has normal, extreme- value or logistic distribution .

In this section we conducted simulation studies, a Monte Carlo simulation study was made for the case of a single explanatory variable to assess the performance of asymptotic efficiency of the OLS estimator relative to the ML estimator, for of the regression coefficient β_1.

For normal distribution, Y_i having the density

$$f(y_i) = 1/\sigma\sqrt{2\pi}\exp[-1/2(y_i - \beta o - \beta 1\, x_i / \sigma)^2], \quad -\infty < \mathbf{y_i} < \infty, \tag{24}$$

For extreme- value distribution, Y_i having the density

$$f(y_i) = 1/\theta\exp\{-(y_i - \beta o - \beta 1 x_i/\theta + \gamma) - \exp(-(y_i - \beta o - \beta 1\, x_i/\theta + \gamma)\}, -\infty < yi < \infty, \tag{25}$$

For logistic distribution, Y_i having the density

$$f(y_i) = 1/\theta\ \{\exp(-(y_i - \beta o - \beta 1\, x_i/\theta))/(1 + \exp(-(y_i - \beta o - \beta 1\, x_i)/\theta)^2\}, -\infty < yi < \infty, \tag{26}$$

An assessment of the Ml estimator and the OLS estimator, for β_1, the variance efficiency of the estimator for β_1 has been made by a Generalized Linear Modeling using SPSS Version 20 for the case of simple linear regression, Y_i having the density normal, extreme-value and logistic distribution.

Equally spaced values for x were taken with $x_i = i - ½(n+1)$, i=1 , … n . Equal sample sizes n=5, 10, 20 with a run-size of 3000 were used. Without loss of generality, the y- observations were generated putting $\beta_0 = \beta_1 = 0$ and θ or $\sigma=1$ in the regression model.

The values of variances of the ML estimators, OLS estimators were estimated by simulation. The variance efficiencies of the estimator for $\beta_1$are shown in Tables 1, 2, 3 .

Table 1 Variance efficiencies for estimation of $\beta_{1,\ \mathrm{Normal}}$ Error

Sample Size	$V(\beta_1^{\sim})$	$V(\beta_1^{\wedge})$	Efficiency
5	0.205	0.201	0.980
10	0.195	0.186	0.986
20	0.147	0.145	0.996

Table 2 Variance efficiencies for estimation of $\beta_{1,}$ E-Value Error

Sample Size	$V(\beta_1^{\sim})$	$V(\beta_1^{\wedge})$	Efficiency
5	3.250	2.145	0.660
10	1.600	1.008	0.630
20	0.820	0.508	0.62

Table 3 Variance efficiencies for estimation of $\beta_{1,}$ Logistic Error

Sample Size	$V(\beta_1^{\sim})$	$V(\beta_1^{\wedge})$	Efficiency
5	0.510	0.454	0.890
10	0.459	0.411	0.895
20	0.402	0.362	0.900

In Table 1, Table 2 and Table 3, it can be seen

1- The OLS efficiency for normal error is appreciably higher than the asymptotic value (one) when n=5, 10, but for higher value of n=20 its performance is much better than that of n=5, 10.

2- The OLS efficiency for extreme error is appreciably higher than the asymptotic value (0.608) when n=5, 10, but for higher value of n=20 its performance is much better than that of n=5, 10.

3- The OLS efficiency for logistic error is appreciably higher than the asymptotic value (0.912) when n=5, 10, but for higher value of n=20 its performance is much better than that of n=5, 10.

4-Conclusion

From literature review, there are some of authors such as (Lawless,1982) , (Haddaw and Young,1986), Jin, Z.L et al, 2006) and others considered in which the error variable ε has non normal distribution. Asymptotic efficiency formula of the OLS estimator relative to the ML estimator, for some probability density function in linear regression models are derived. Asymptotic efficiency values for the regression coefficient β_1 were (1, 0.608,0.912) for normal, extreme-value and logistic distributions respectively. The OLS efficiency for normal, extreme-value and logistic error with n=20 its performance is much better than that of n=5, 10.

ACKNOWLEDGEMENT

The author thanks the reviewers for their helpful comments.

REFERENCES

Al-Sarraf, Z.(1986).Some Statistical Problem Connected with Logistic Regression. Thesis, Brunel the University of West London, U. K.

Cox, D. R. and Hinckley, D.V.(1968). A Note on the Efficiency of Least- Squares Estimates.

Draper, N.R. and Smith, H.(1982).Applied Regression Analysis 2nd Ede. New York .John Wily and sons, inc.

Haddow, A. A. and Young, D.H. (1986). Moment Properties of Estimators for A Type 1 Extreme -Value Regression Model. Communications In Statistics, Theory and Methods, 15 (8).

Haddaw,A.A.(2014). Maximum Likelihood Normal Regression with Censored and Uncensored Data. Journal of Natural Sciences Research (IISTE Org.). Vol.4, No.3, 2014.U.S.A.

Lawless, J. F.(1982). Statistical Models and Methods for Lifetime Data. New York.

Jin, Z.L et al, (2006).On Least Squares Regression With Censored Data.Biomerika,93 ,1.

.