Journal of the Korean Statistical Society 45 (2016) 477–489

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Comparisons of estimators for regression coefficient in a misspecified linear model with elliptically contoured errors

Ping Peng, Guikai Hu *, Muhua Ding

School of Science, East China University of Technology, Nanchang 330013, China

ARTICLE INFO

Article history:
Received 24 January 2015
Accepted 25 January 2016
Available online 16 February 2016

AMS 2000 subject classifications:
primary 62J07
secondary 62F10

Keywords:
Comparisons of estimators
Misspecified linear model
Elliptically contoured distribution

ABSTRACT

In a misspecified linear regression model with elliptically contoured errors, the exact risks of generalized least squares (GLS), restricted least squares (RLS), preliminary test (PT), Stein-rule (SR) and positive-rule shrinkage (PRS) estimators of regression coefficient are derived. Risk superiority conditions dependent on prior constraint error and the model specification error are obtained. When the model is misspecified and the error terms obey the elliptically contoured distribution, it is shown analytically that the PRS estimator dominates uniformly the SR estimator. However, the PRS and SR estimators do not dominate uniformly the GLS estimator. Furthermore, the dominance of the RLS estimator over the GLS estimator does not hold necessarily even if the model has linear constraint.

© 2016 The Korean Statistical Society. Published by Elsevier B.V. All rights reserved.

1. Introduction

In the context of linear regression, the ordinary least squares (OLS) estimator for regression coefficient is well known as the best linear unbiased estimator. However, Stein-rule (SR) estimator proposed by James and Stein (1961) and Stein (1956) dominates OLS estimator in terms of mean squared error of prediction. From then on, some shrinkage estimators have been proposed, and their sampling properties have also been examined. Some examples are Baranchik (1970), Farebrother (1975), Ohtani (1996), Theil (1971) and so on. For the error variance, it is well known that Stein (1964) variance estimator is better than the usual estimator under mean squared error. Gelfand and Dey (1988) proved that the Stein variance estimator is in fact a preliminary test estimator after a pre-test of linear restrictions on the coefficients. In the literature, much more work have been done on the pre-test estimator of the error variance. For more details, the readers are referred to Clarke, Giles, and Wallace (1987), Ohtani (1988), Wan (1997), Wan, Zou, and Ohtani (2006), Zhu and Zhou (2011), and so on.

Their studies have two common characteristics. Firstly, the linear regression model considered was usually specified correctly. Secondly, the error term was usually assumed to obey a normal distribution. However, on the one hand, the model may be specified incorrectly due to the lack of data, ignorance or simplification in the real world. On the other hand, many economic data may be generated by the thick tail distribution. At this time, the multivariate normal distribution is not suitable. Meanwhile, the elliptically contour distribution is an extension of the multivariate normal distribution, and it has many similarities with the multivariate normal distribution. Moreover, Elliptically contoured distribution can be regarded as a proper alternative for multivariate normal distribution satisfying robustness and stableness in improved estimator. Therefore, the problem of estimation in a misspecified linear regression model with elliptically contoured distribution arises naturally.

* Corresponding author. Tel.: +86 791 83897509; fax: +86 791 83897509.
E-mail address: huguikai97@163.com (G. Hu).

http://dx.doi.org/10.1016/j.jkss.2016.01.007
1226-3192/© 2016 The Korean Statistical Society. Published by Elsevier B.V. All rights reserved.

Consider a linear model of the form

$$y = X\beta + \varepsilon, \quad \varepsilon \sim E_n(0, \sigma^2 V, \psi, \phi), \tag{1.1}$$

where $y \in \mathbb{R}^n$ is an observable random vector, X is a known constant matrix of conformable dimensions, β is an unknown parameter, ε is an error term assumed to obey an elliptically contoured distribution $E_n(0, \sigma^2 V, \psi, \phi)$ with a probability density. By Fang, Kotz, and Ng (1990), function $\psi : [0, \infty] \longrightarrow \mathbb{R}$ is called a characteristic generator, and the generated characteristic function is $\psi(\sigma^2 \tau' V \tau)$. According to Chu (1973), the probability density can be represented as a scale mixture between a normal distribution with scale parameter t^2 and a weight function $\omega(t)$ by

$$f_E(u) = \theta|\sigma^2 V|^{-\frac{1}{2}} \phi\left(\frac{1}{2\sigma^2} u' V^{-1} u\right) = \int_0^\infty f_N(u|t)\omega(t)dt. \tag{1.2}$$

Here θ is a normalizing constant, $\sigma^2 > 0$ is an unknown parameter, V is a known $n \times n$ positive definite matrix, $f_N(u|t)$ denotes the probability density function of multivariate normal distribution $N_n(0, \sigma^2 t^{-1} V)$, function ϕ is called a probability density generator. Also ϕ and ψ determine each other for each specific member of this family. It is easy to verify that ε has zero mean vector and its covariance matrix is

$$Cov(\varepsilon) = -2\sigma^2 \psi'(0) V = \sigma_\varepsilon^2 V, \quad \text{where } \sigma_\varepsilon^2 = -2\sigma^2 \psi'(0).$$

Some results related to the sampling performances of estimators in the linear model with elliptically contoured errors have been established. For example, Arashi, Saleh, and Tabatabaey (2010, 2013) discussed the estimation of parameters in parallelism model and regression model, respectively. Under a balanced loss function, Arashi (2009, 2012) analyzed the problem of estimation in multiple regression model and a seemingly unrelated regression system, respectively. However, little is known about the sampling performance of estimators in a misspecified linear model with elliptically contoured errors.

Partition X and β as

$$X = \begin{pmatrix} X_1 & \vdots & X_2 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} \beta_1 \\ \cdots \\ \beta_2 \end{pmatrix},$$

where $\beta_1 \in \mathbb{R}^{k_1}$, $\beta_2 \in \mathbb{R}^{k_2}$, $X_1 \in \mathbb{R}^{n\times k_1}$ and $X_2 \in \mathbb{R}^{n\times k_2}$ are of full column rank matrices, respectively. Now, suppose that the researcher specifies the model as

$$y = X_1\beta_1 + \gamma, \tag{1.3}$$

where $\gamma = \eta + \varepsilon$, $\eta = X_2\beta_2$. The vector $\eta \in \mathbb{R}^n$ represents the contribution of the omitted explanatory variables to the expected value of the explained variable y.

Nowadays, Hu, Yu, and Luo (2015) have studied the comparisons of variance estimators in a misspecified linear model with elliptically contoured distribution. However, such comparison among the estimators of regression coefficient has not been made so far. In this paper, we will derive the explicit formula for the risks of the generalized least squares, restricted least squares, preliminary test, Stein-rule and positive-rule shrinkage estimators of regression coefficient in model (1.3), and examine their sampling performances. In the next section we will introduce five estimators of regression coefficient and obtain the explicit formulas for the risks of proposed estimators. In Section 3, we will analyze the risk performances of five estimators by theoretical analysis. Considering the complication of the risks, we also perform the comparisons by numerical analysis in Section 4. In Section 5, we give a simulation example to test our results.

2. Estimators and its risks

In the model (1.3), we define the generalized least squares estimator of β_1 as

$$\hat{\beta}_1^{GLS} = (X_1' V^{-1} X_1)^{-1} X_1' V^{-1} y \triangleq B^{-1} X_1' V^{-1} y.$$

Similarly, the least squares estimator of σ_ε^2 is

$$S^2 = \frac{(y - X_1\hat{\beta}_1^{GLS})' V^{-1} (y - X_1\hat{\beta}_1^{GLS})}{m},$$

where $m = n - k_1$. In many practical situations, it is suspected that β_1 has the linearly independent prior constraints $H\beta_1 = 0$, where H is a known row full rank matrix of $p \times k_1$ order and p is asked to satisfy $p \geq 3$. For test of $H\beta_1 = 0$, we define the restricted least squares estimator as

$$\hat{\beta}_1^{RLS} = \hat{\beta}_1^{GLS} - B^{-1} H' (H B^{-1} H')^{-1} H \hat{\beta}_1^{GLS}.$$

Download English Version:

https://daneshyari.com/en/article/1144516

Download Persian Version:

https://daneshyari.com/article/1144516

Daneshyari.com