

A Comparative Study of Generalized Classes of Estimators and Alternative Estimators in Regression Models Incorporating a Priori Information

Dr. Syed Qaim Akbar Rizvi

Department of statistics, Shia P.G. College, Lucknow University, Lucknow

ABSTRACT

In this chapter we discuss the generalized classes of estimator in regression model with apriori information deals with develop generalized classes of estimators quassi minimax, mock minimax with small asymptotically and large asymptotically approaches, restricted regression model, mixed regression model. In this chapter we deal the generalized classes of estimator in regression model finally more efficient estimators of the classes than the existing ones under different criteria of choosing better estimators like expected loss or concentration probability.

1. INTRODUCTION

In social, chemical, physical, biological and medical sciences, engineering, business, planning and many other fields of practical importance, we require the analysis of causes and effects involved in the mechanism of various intricate phenomena under our study. On the basis of best experience or some theoretical justification, the relationship between causes and effects are specified. This specified relationship between causes and effects together with certain other assumptions constitutes the theoretical framework in the form of providing the data generating mechanism of the phenomena under our study. The simple well known theoretical frame-work representing the complex mechanism underlying various intricate phenomena is the linear regression model which is linear in parameters in its basic nature. Linear regression model under certain assumptions contains some unknown parameters which are to be estimated on the basis of sample and apriori information.

2. THE LINEAR MODEL WITH CONSTRAINTS

Let the classical linear model be

$$y = X\beta + u \quad (1.2.1)$$

where y is $T \times 1$ vector of observations on the variable to be explained, X is full column rank matrix of observations on p explanatory variables, β is a $p \times 1$ vector of unknown regression coefficients and u is $T \times 1$ vector of disturbances which are independently and identically distributed with mean zero and the same variance σ^2 , that is, we have

$$E(u) = 0$$

and

$$E(uu') = \sigma^2 I_T. \quad (1.2.2)$$

We now consider the above linear model under three types of constraints:

Exact Constraints

On some prior information, we may have J linear equality restrictions on the $p \times 1$ parameter vector β in linear model (1.2.1) as

$$R\beta = r \quad (1.2.3)$$

where R is $J \times p$ matrix of rank J and r is a $J \times 1$ vector, both consisting of known nonstochastic elements.

Mixed Regression Stochastic Restricted Model

Theil and Goldberger (1961) developed the mixed regression model using the linear restrictions $R\beta = r$ stochastically. Thus, we have the linear restrictions on the parameter vector β as

$$r = R\beta + v \quad (1.2.4)$$

where r is a $J \times 1$ vector of $j(\leq p)$ random variables, R is a $J \times p$ matrix with known elements and v is a $J \times 1$ vector of stochastic elements having mean vector null. dispersion matrix Ψ and the elements of v are uncorrelated with those of u in (1.2.1), that is

$$\begin{aligned} E(v) &= 0 \\ E(vv') &= \Psi \end{aligned} \quad (1.2.5)$$

and

$$E(uv') = 0.$$

Noting $E(v) = 0$ in (1.2.4), we see that the restrictions (1.2.4) may be written as which shows that r contains the unbiased estimators of J linear parametric functions.

The main idea in developing mixed regression model is to combine the prior information $R\beta + v$ and the sample ($y = x\beta + u$) information and get the model

$$\begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} x \\ R \end{bmatrix} \beta + \begin{bmatrix} u \\ v \end{bmatrix} \quad (1.2.7)$$

with the assumptions in (1.2.1) and (1.2.5).

(iii) Model Under Ellipsoidal Constraints

Instead of being interested in the entire parametric space of a model, we may sometimes be interested in its subspace which may be constrained in the form of nonlinear ellipsoid given by

$$\beta'H\beta \leq k, \quad (1.2.8)$$

where H and k respectively are known positive definite symmetric matrix and a positive scalar. In ellipsoidal constraints (1.2.8), ellipsoid has its center at the origin, but if our interest lies in an ellipsoidal constraint with its ellipsoid having its center at any known p - dimensional point β_0 in the p - dimensional space, then the nonlinear ellipsoidal constraints may be written as

$$(\beta - \beta_0)'H(\beta - \beta_0) \leq k. \quad (1.2.9)$$

3. CRITERIA FOR THE PERFORMANCE OF ESTIMATORS

There are several methods of obtaining estimators in the literature but the question now arises whether some of many possible estimators are better in some sense, than the others. To solve this problem, we present some criteria of judging the performance of an estimator that will help us in deciding whether one estimator is better than another.

Closeness

Given a random sample $X_1, X_2, \dots, X_n$, an estimator of a parameter θ is a statistic, say $\ell(x_1, x_2, \dots, x_n)$, whose value is used as an estimate of θ . Any estimation procedure aims at getting the value as close to the true unknown parameter θ as possible. Ideally, one would like the value of the estimator $\ell(x_1, x_2, \dots, x_n)$

to be the unknown θ , which is not practically feasible except in some trivial cases. Thus, in the absence of being able to achieve the ultimate of always correctly estimating the unknown parameter θ , the alternative is to look for an estimator $1(X_1, X_2, \dots, X_n)$ that remains closed to the unknown parameter θ . Closeness may be defined in several ways. The distribution of statistic $T = 1(X_1, X_2, \dots, X_n)$ tells how the value of t of T are distributed and one would like to have the values of T distributed near θ , that is, one would prefer to select $1(X_1, X_2, \dots, X_n)$ so that its values may be concentrated near θ . This makes the requirement for an estimator to have its mean near or equal to θ with smaller mean square error or variance.

Rather than resorting to characteristics like mean and variance of a distribution, one may consider 'Concentration or closeness' in terms of the distribution itself. In this context, we have the following two definitions depending upon what 'concentration' might mean in terms of the distribution itself.

More Concentrated And Most Concentrated Given two estimators $T = 1(x_1, x_2, \dots, x_n)$ and $T' = \ell'(X_1, X_2, \dots, X_n)$ of θ , T' is called a more concentrated estimator of θ than T if and only if

$$P(\theta - \lambda < T' \leq \theta + \lambda) \geq P(\theta - \lambda < T \leq \theta + \lambda)$$

for all $\lambda > 0$ and each θ in the parametric space θ . If an estimator $T^* = \ell^*(X_1, X_2, \dots, X_n)$ is more concentrated than any other estimator, it is called most concentrated.

PITMAN - Closer and PITMAN - Closest

Given two estimators $T = \ell(X_1, X_2, \dots, X_n)$ and $T' = \ell'(X_1, X_2, \dots, X_n)$, T' is called a Pitman-Closer estimator of θ than T if and only if

$$P[|(T' - \theta)| < |(T - \theta)|] \geq \frac{1}{2} \text{ for each } \theta \text{ in } \theta.$$

An estimator $T^* = \ell^*(X_1, X_2, \dots, X_n)$ is called Pitman - Closest if it is Pitman - Closer than any other estimator.

We know that mean squared error is widely used criteria for choice of estimators on the justification of easier algebraic manipulations. However, it has got some limitations. Since, it depends upon squares of deviations from the true value, it is, in some situations highly affected by extreme high or low values. This results in deviation from the main idea of concentration of estimator around true unknown parameter. The other drawback of the mean squared error criterion is that it fails to say anything for the choice of estimators which do not possess second order finite moments. Alternative in such situations is to make use of some other criterion like Pitman (1937) nearness or concentration probability. We, here, go for the use of criterion of concentration probability to study the performance of some estimation procedures.

4. REVIEW OF RELEVANT LITERATURE

In the context of linear restricted regression model, Srivastava and Srivastava (1984), proposed two families of shrinkage estimators by synthesizing Steinrule estimator with the restricted regression estimator in an appropriate manner, discussed their properties with respect to bias vectors, mean squared error matrices and the risks under quadratic loss function when the disturbances

are assumed to be small and later on, Srivastava and Chandra (1991) studied these two families of estimators when disturbances are not necessarily normal. Srivastava and Srivastava (1983) considered such families of estimators for linear regression model with incomplete prior information in the form of stochastic linear constraints and studied their properties. Shukla (1993) again made a study regarding these two families of estimators under exact and stochastic linear restrictions with respect to the criterion of concentration probability of estimator around true unknown parameter.

Extending the results of Srivastava and Srivastava (1984), Srivastava and Chandra (1991) and many others, Singh (1994) proposed a generalized class of restricted regression estimators which includes some improved efficient estimators than the existing ones. Further Singh et. al. (1995) considered a generalized class of mixed regression estimators containing an improved subclass of estimators over the existing mixed or mixed type regression estimators with respect to a general quadratic loss function in linear regression model under constraints.

REFERENCES

1. Blvth. (R (1972) Some Probability Paradoxes in Choice from Among Rankem Alternatives. Journal of the American Statistical Association 67 336-373
2. Bunke, O (1975) Minimax Linear, Ridge and Shrunken Estimators for I mear Parameters. Math. Operationsforschung Statistik 6 697-701
3. Durbin, J (1953). A note on Regression when there is Extraneous Information about one of the Coefficients. Journal of the American Statistical Association 48 799-808
4. Giles. D E. A and V.K. Srivastava (1991) An Unbiased Estimator of the Concerrame Matrix of the Mixed Regression Estimator. American Statistical Association 86 441-444.
5. Guilkey, D.K. and J M. Price (1981). On Comparing Restricted Least Square Estimators. Econometrics 15 397-404.
6. Gujati, Damodar (1978). Basic Econometric, McGraw-Hill Book Compan
7. Hartung, J. (1978). Zur Verwendung von Vorinformation in der Regresson Andve (Asing Prior Information in Regression.) Research Report, Department of Applied Statistics, University of Bonn
8. Hoffmann. K. (1977). Admissibility of Linear Listanctors with repeal to Ro and Parameter Sees. Math. Operationsforschung Statistin 8 +25-488