Journal of Statistical Research *ISSN 0256 - 422 X*
2013, Vol. 47, No. 1, pp. 95-105

L-ESTIMATION OF THE PARAMETERS IN A LINEAR MODEL BASED ON A FEW SELECTED REGRESSION QUANTILES

A.K.MD. EHSANES SELEH
Carleton University, Ottawa, Canada
Email: saleh3422@rogers.com

ABSTRACT

This paper considers the L-estimation of regression and scale parameter of the linear model $\mathbf{Y} = \beta_0 1_o + \mathbf{X}\boldsymbol{\beta} + \sigma\mathbf{e}$, where β_0 is the intercept parameters and σ is the scale in the model, based on $k(\leq n)$ *optimum regression quantiles* as defined by Koenker and Bassett (1978). In addition, the paper contains the trimmed estimation problem with continuous weight functions, the estimation of conditional regression function and the related optimum regression quantiles.

1 Introduction

Consider the model

$$\mathbf{Y}_{n\times 1} = \beta_0 1_{n\times 1} + \mathbf{X}_{n\times p}\boldsymbol{\beta}_{p\times 1} + \sigma\mathbf{e}_{n\times 1},$$

where $E(\mathbf{e}_{n\times 1}) = 0$ and $E(\mathbf{e}_{n\times 1}\mathbf{e}'_{1\times n}) = \sigma^2\mathbf{I}_n$ and the components of $\mathbf{e}_{n\times 1}$ are independent and distributed with the cdf F_0 which is known. Our basic problem is the estimation of the parameter $(\boldsymbol{\beta}'; \sigma) = (\beta_0, \ldots, \beta_p; \sigma)$ based on the observation vector $\mathbf{Y}$ and the design matrix $\mathbf{D}_n = (\mathbf{1}_n|\mathbf{X})$. For the location submodel i.e. when $\mathbf{X} = 0$, Ogawa (1951) developed the procedure of estimation the location and the scale parameters jointly as well as singly, based on a few selected sample quantiles. Subsequently, many authors like Sarhan and Greenberg (1962), Saleh (1992), Saleh (1981), Saleh and Ali (1966), and Harter (1963), among many others followed the procedure for various specific distributions and obtained a few *optimum sample quantiles* for the estimation of location and scale parameters of F_0.

Recently, Koenker and Bassett (1978) introduced the concept of regression quantiles as an extension of the sample quantiles to the linear model. This concept seems to provide a reasonable basis not only the construction of robust L-estimators of regression parameters but also develops robust test of the linear hypothesis. Koenker and Bassett (1978) also suggested the trimmed least squares estimator as an extension of the trimmed mean to the linear model. This idea has later been studied by Ruppert and Caroll (1980). They also derived the Bahadur type representation of regression quantiles up to order $\mathbf{O}_p(n^{-1/2})$ which is also extended to order $\mathbf{O}_p(n^{-3/4})$ by Jureckova (1984).

The trimmed least-squares estimation with continuous weight function have also been pursued by Koenker and Portnoy (1987) and extended by Guttenbrunner and Jureckova (1992).

The object of this paper is to propose estimation of $(\boldsymbol{\beta}, \sigma)$ based on a few *selected (optimum) regression quantiles* extending the idea of Theorem 4.3 of Koenker and Bassett (1978) from the location scale model to the linear model.

2 Estimation of $(\boldsymbol{\beta}', \sigma)$ based on a few selected regression quantiles

Let $Y_1, \ldots, Y_n$ be independent observations Y_i is distributed according to $F_0[(y-\beta_0-\sum_{i=1}^{p} x_{ij}\beta_i)/\sigma]$, where F_0 is known and absolutely continuous having pdf f_0 which is non-negative and σ is the unknown scale parameter while $\mathbf{D}_n = (\mathbf{1}_n|\mathbf{X})$ is a known design matrix satisfying the following conditions

$$(i)\ \lim_{n\to} \bar{x}_{nj} = \bar{x}_j, \quad (ii) \lim n^{-1}\mathbf{D_n'}\mathbf{D_n} = \mathbf{C},\ \text{a } (p+1)\times(p+1) \text{ matrix.}$$

Let $Q_0(\lambda)$ denote the *quantile-function* of the distribution F_0 corresponding to the spacing λ $(0 < \lambda < 1)$ and $q_0(\lambda) = f_0(Q_0(\lambda))$ be the *density-quantile function*. Assume that n is large and $k(\leq n)$ is a given integer. Consider the k fixed spacings $\lambda_1, \ldots, \lambda_k$ satisfying the relation $0 < \lambda_1 < \cdots < \lambda_k < 1$ and consider the k-regression quantiles $\hat{\beta}_j(\lambda_1), \ldots, \hat{\beta}_j(\lambda_k)$ $(j = 0, \ldots, p)$ which are the solution of the minimization problem

$$\sum_{j=1}^{n} \rho_{\lambda_i}(\mathbf{Y}_j - \beta_0 - \mathbf{x}_j' t) = min$$

where $\rho_\lambda(\mathbf{x}) = \mathbf{x}\{\lambda\mathbf{I}(x > 0) - (1-\lambda)\mathbf{I}(x \leq 0)\}$ and $\mathbf{I}(A)$ is the indicator function of the set A. The regression quantile minimization problem is equivalent to the linear program

$$\left.\begin{array}{ll} \text{minimize} & [\lambda\mathbf{1}_n' r^+ + (1-\lambda)\mathbf{1}_n' r] \\ \text{subject to} & \mathbf{Y} = \beta_0\mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + r^+ + r^-, \ \ \{(\beta_0, \boldsymbol{\beta}), r^+, r^-\} \in \mathbf{R}^{p+1} \times \mathbf{R}^{2n} \end{array}\right\} \quad (2.1)$$

Let us denote the vector of solutions for the minimization problem as

$$\hat{\boldsymbol{\beta}}_j(\lambda) = (\hat{\beta}_j(\lambda_1), \ldots, \hat{\beta}_j(\lambda_k))', \ \ j = 0, 1, \ldots, p,$$

also, let $\mathbf{Q}_0(\lambda) = (Q_0(\lambda_1), \ldots, Q_0(\lambda_k))'$ and $\mathbf{1}_k = (1, \ldots, 1)'$, a k-tuple.

Then, using Theorem 4.2 of Koenker and Bassett (1978) we obtain that the $k(p+1)$-dimensional random variable

$$\sqrt{n}\{[\hat{\beta}_0(\lambda) - \beta_0\mathbf{1}_k - \sigma\mathbf{Q}_0(\lambda)]', [\hat{\beta}_1(\lambda) - \beta_1\mathbf{1}_k]', \ldots, [\hat{\beta}_p(\lambda) - \beta_p\mathbf{1}_k]'\} = \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})',$$

say, where $(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ is $k \times (p+1)$-dimensional matrix, converges in distribution to an $k(p+1)$-dimension normal distribution with mean $\mathbf{0}$ and covariance matrix $\sigma^2(\mathbf{C}^{-1} \otimes \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ is a $k \times k$ matrix defined by

$$\boldsymbol{\Omega} = \left[\frac{\lambda_i \wedge \lambda_j - \lambda_i \lambda_j}{q_0(\lambda_i) q_0(\lambda_j)} \right], \tag{2.2}$$

which is the asymptotic covariance-matrix of k ordinary sample quantiles from the distribution F_0. This theorem parallels Mosteller (1946). Thus, following generalized least squares principle one minimizes the quadratic from $n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})'(\mathbf{C} \otimes \boldsymbol{\Omega}^{-1})(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ to obtain the normal equations for the asymptotically best linear estimators (ABLUE) of $(\boldsymbol{\beta}', \sigma)$.

$$\begin{bmatrix} K_1 & \bar{\text{x}}_{01}K_1 & \cdots & \bar{\text{x}}_{0p}K_1 & K_3 \\ \bar{\text{x}}_{01}K_1 & (s_{11}^2 + \bar{\text{x}}_{01}^2)K_1 & \cdots & (s_{1p} + \bar{\text{x}}_{01}\bar{\text{x}}_{0p})K_1 & \bar{\text{x}}_{01}K_3 \\ \bar{\text{x}}_{02}K_1 & (s_{12} + \bar{\text{x}}_{01}\bar{\text{x}}_{02})K_1 & \cdots & (s_{2p} + \bar{\text{x}}_{02}\bar{\text{x}}_{0p})K_1 & \bar{\text{x}}_{02}K_3 \\ \vdots & \vdots & \ddots & \vdots & \\ \bar{\text{x}}_{0p}K_1 & (s_{1p} + \bar{\text{x}}_{01}\bar{\text{x}}_{0p})K_1 & \cdots & (s_{pp}^2 + \bar{\text{x}}_{0p}^2)K_1 & \bar{\text{x}}_{0p}K_3 \\ K_3 & \bar{\text{x}}_{01}K_3 & \cdots & \bar{\text{x}}_{0p}K_3 & K_2 \end{bmatrix} \begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \beta_2^* \\ \vdots \\ \beta_p^* \\ \sigma \end{bmatrix} = \begin{bmatrix} V_0 \\ V_1 \\ V_2 \\ \vdots \\ V_p \\ V_{p+1} \end{bmatrix}$$

where

$$\begin{aligned} V_0 &= Z_0 + \bar{x}_{01}Z_1 + \cdots + \bar{x}_{0p}Z_p,\ V_j = \bar{x}_{0j}V_0 + s_{j1}Z_1 + \cdots + s_{jp}Z_p, (j = 1, \ldots, p) \\ V_{p+1} &= Z_0^* + \bar{x}_{01}Z_1^* + \cdots + \bar{x}_{0p}Z_p^*,\ Z_j = \mathbf{1}'\boldsymbol{\Omega}^{-1}\hat{\boldsymbol{\beta}}_j(\lambda), \\ Z_j^* &= \mathbf{Q}'(\lambda)\boldsymbol{\Omega}^{-1}\hat{\boldsymbol{\beta}}_j(\lambda),\ j = 0, 1, \ldots, p, K_1 = \mathbf{1}_k'\boldsymbol{\Omega}^{-1}\mathbf{1}_k, \\ K_2 &= \mathbf{Q}'(\lambda)\boldsymbol{\Omega}^{-1}\mathbf{Q}(\lambda), K_3 = \mathbf{1}_k'\boldsymbol{\Omega}^{-1}\mathbf{Q}(\lambda), \Delta = K_1, K_2 - K_3^2 \\ ns_{ij} &= \sum_{h=1}^{n}(x_{ih} - \bar{x}_i)(x_{jh} - \bar{x}_j),\ i, j = 0, 1, \ldots, p. \end{aligned}$$

Then, the asymptotic distribution of

$$\{\sqrt{n}(\beta_0^* - \beta_0), \sqrt{n}(\beta_1^* - \beta_1), \ldots, \sqrt{n}(\beta_p^* - \beta_p), \sqrt{n}(\sigma^* - \sigma)\} \text{ is } N_{p+2}(\mathbf{O}, \sigma^2\mathbf{K}^{-1})$$

where

$$\begin{bmatrix} K_1 & \cdots & \bar{x}_0K_1 & \cdots & \bar{x}_{01}K_1 & \cdots & K_3 \\ \bar{x}_{01}K_1 & \cdots & (s_{11} + \bar{x}_{01}^2)K_1 & \cdots & (s_{1p} + \bar{x}_{01}\bar{x}_p)K_1 & \cdots & \bar{x}_{01}K_3 \\ \vdots & \vdots & \ddots & \vdots & & & \\ \bar{x}_{0p}K_1 & \cdots & (s_{p1} + \bar{x}_{01}\bar{x}_{0p})K_1 & \cdots & (s_{pp} + \bar{x}_{0p}^2)K_1 & \cdots & \bar{x}_{0p}K_3 \\ K_3 & \cdots & \bar{x}_{01}K_3 & \cdots & \bar{x}_{0p}K_3 & \cdots & K_2 \end{bmatrix}$$

It may be verified that

$$\mid \mathbf{K} \mid = \mid \mathbf{C} \mid K_1^p\Delta, \quad \text{where } \Delta = K_1K_2 - K_3^2. \tag{2.3}$$

Also, it may also be shown that the estimates are asymptotically unbiased.

Further, we note that the expressions corresponding to K_1, K_2 and K_3 are given by

$$\left.\begin{aligned} K_1 &= \sum_{i=1}^{k+1}[q_0(\lambda_i)-q_0(\lambda_{i-1})]^2/(\lambda_i-\lambda_{i-1}) \\ K_2 &= \sum_{i=1}^{k+1}[q_0(\lambda_i)Q_0(\lambda_i)-q_0(\lambda_{i-1})Q_0(\lambda_i)]^2/(\lambda_i-\lambda_{i-1}), \quad \text{and} \\ K_3 &= \sum_{i=1}^{k+1}[q_0(\lambda_i)-q_0(\lambda_{i-1})][q_0(\lambda_i)Q_0(\lambda_i)-q_0(\lambda_{i-1})Q_0(\lambda_{i-1})]/(\lambda_i-\lambda_{i-1}) \end{aligned}\right\} \tag{2.4}$$

with $\lambda_0 = 0$ and $\lambda_{k+1} = 1$. If one chooses symmetric spacings then K_3 reduces to 0 and the ABLUE are given by $\boldsymbol{\beta}^* = K_1^{-1}C^{-1}\mathbf{V}_*$ and $\sigma^* = K_2^{-1}\mathbf{V}_{p+1}$, where $\mathbf{V}'_* = (V_1,\ldots,V_p)$.

It may be shown that the asymptotic relative efficiency (ARE) of the AVLUE based on the regression quantile, with spacings $(\lambda_1,\ldots,\lambda_k)$ relative to the usual least squares estimation (LSE) is given by ARE (ABLUE: LSE) $= \frac{K_1^p\Delta}{I_{11}^p(I_{11}I_{22}-I_{12}^2)}$, where $\mathbf{I} = (I_{ij})$ is the information matrix of the location-scale family with cdf F_0.

Thus, in order to obtain the optimum spacing vector $(\lambda_1^0,\ldots,\lambda_k^0)$ we maximize $K_1^p\Delta$ with respect to $(\lambda_1^0,\ldots,\lambda_k^0)$ subjects to $0 < \lambda_1 < \ldots < \lambda_k < 1$. Thus, we solve the system of equations

$$p\Delta K_1^{p-1}\frac{\delta K_1}{\delta\lambda_i} + K_1^p\frac{\delta\Delta}{\delta\lambda_i} = 0,\ i = 1,\ldots,k.$$

Therefore, the ABLUE of $(\beta_0,\ldots,\beta_p,\sigma)$ are obtained first by computing $\hat{\beta}_j(\lambda_i^0),\ j = 0,1,\ldots,p$ and $i = 1,2,\ldots,k$ and $K_1, K_2, K_2, Z_1, Z_2, Z_1^*$ and Z_2^* using the optimum spacings $(\lambda_1^0,\ldots,\lambda_k^0)$ then using the normal equations (2.9). Then, the ARE is given by $\frac{K_{10}^p\Delta_0}{I_{11}^p(I_{11}I_{22}-I_{12}^2)}$, where K_{10} and Δ_0 are the maximum values of K_1 and Δ is achieved. We shall consider an example in section 4 using exponential errors.

3 Estimation of conditional quantile function

Consider the conditional quantile functions $Q(\xi) = l_0\beta_0 + l_1\beta_1 + \ldots + l_p\beta_p + \sigma Q_0(\xi),\ 0 < \xi < 1$. The estimate of $Q(\xi)$ is obtained by substituting ABLUE of $(\beta_0,\ldots,\beta_p;\sigma)$ which is

$$Q^*(\xi) = l_0\beta_0^* + l_1^*\beta_1 + \ldots + l_p^*\beta_p + \sigma^*Q_0(\xi).$$

The asymptotic variance of $Q^*(\xi)$ is given by

$$\sigma^2 l'K^{-1}l,\ l = (l_0,\ldots,l_p,Q(\xi))'.$$

Here, the vector l is known. Similarly, the asymptotic variance of the LSE of the parameter is $\sigma^2 l'\mathbf{I}^{*-1}l$, where

$$I^* = \begin{bmatrix} I_{11} & \cdots & \bar{x}_0 I_{11} & \cdots & \bar{x}_{01} I_{11} & \cdots & I_{12} \\ \bar{x}_{01} I_{11} & \cdots & (s_{11}+\bar{x}_{01}^2) I_{11} & \cdots & (s_{1p}+\bar{x}_{01}\bar{x}_{0p}) I_{11} & \cdots & \bar{x}_{01} I_{12} \\ \bar{x}_{02} I_{11} & \cdots & (s_{12}+\bar{x}_{01}\bar{x}_{02}) I_{11} & \cdots & (s_{2p}+\bar{x}_{02}\bar{x}_{0p}) I_{11} & \cdots & \bar{x}_{02} I_{12} \\ \vdots & \vdots & \ddots & \vdots & & & \\ \bar{x}_{0p} I_{11} & \cdots & (s_{p1}+\bar{x}_{01}\bar{x}_{0p}) I_{11} & \cdots & (s_{pp}+\bar{x}_{0p}^2) I_{11}\bar{x}_{0p} & \cdots & I_{12} \\ I_{12} & \cdots & \bar{x}_{01} I_{12} & \cdots & \bar{x}_{0p} I_{12} & \cdots & I_{22} \end{bmatrix}$$

Thus, the ARE of $Q^*(\xi)$ relative to LSE $Q^*(\xi)$ is given by

$$\text{ARE}(Q^*(\xi) : Q^{**}(\xi)) = \frac{l'\mathbf{I}^{*-1}l}{l'\mathbf{K}^{-1}l}.$$

Since, $\xi\epsilon(0,1)$ can assume infinitely many values tabular values of ARE becomes, prohibitive except for chosen values of ξ. However, it is well-known that

$$\text{Ch}_{min}(\mathbf{I}^{*-1}\mathbf{K}) \leq \frac{l'\mathbf{I}^{*-1}l}{l'\mathbf{K}^{-1}l} \leq \text{Ch}_{max}(\mathbf{I}^{*-1}\mathbf{K}),$$

where $\text{Ch}_{min}(A)$ and $\text{Ch}_{max}(A)$ and the minimum and maximum characteristic roots.

This means that the maximum ARE is $\text{Ch}_{max}(\mathbf{I}^{*-1}\mathbf{K})$. Thus, one can maximize $\text{Ch}_{max}(\mathbf{I}^{*-1}\mathbf{K})$ or $\text{tr}(\mathbf{I}^{*-1}\mathbf{K})$ with respect to $(\lambda_1, \ldots, \lambda_k)$ to obtain $(\lambda_1^*, \ldots, \lambda_k^*)$ to be the optimum values. These spacings will be used to obtain the appropriate regression quantiles to obtain the optimum estimator of $Q(\xi)$.

Special case of interest is the vector $(1, 0, \ldots, 0, Q(\xi))$ which defines the quantile-function, $Q(\xi)$ and the ABLUE of $Q(\xi)$. Then,

$$l'\mathbf{I}^{*-1}l = \mathbf{I}^{*(11)} + 2Q_0(\xi)\mathbf{I}^{*(p+2,1)} + Q_0^2(\xi)\mathbf{I}^{*(p+2,p+2)} \quad \text{and}$$
$$l'\mathbf{K}^{-1}l = \mathbf{K}^{(11)} + 2Q_0(\xi)\mathbf{K}^{(p+2,1)} + Q_0^2(\xi)\mathbf{K}^{(p+2,p+2)}.$$

Hence, we maximize $\text{tr}(\mathbf{M})$ or $\text{Ch}_{max}(\mathbf{M})$, where

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}^{*(11)} & \mathbf{I}^{*(1,p+2)} \\ \mathbf{I}^{*(p+2,1)} & \mathbf{I}^{*(p+2,p+2)} \end{bmatrix} \begin{bmatrix} \mathbf{K}^{(11)} & \mathbf{K}^{(1,p+2)} \\ \mathbf{K}^{(p+2,1)} & \mathbf{K}^{(p+2,p+2)} \end{bmatrix},$$

where $\mathbf{I}^{*(-1)} = (\mathbf{I}^{*(i,j)})$ and $\mathbf{K}^{(-1)} = (\mathbf{K}^{(i,j)})$. Thus, one maximizes $\text{tr}(\mathbf{M})$ or $\text{Ch}_{max}(\mathbf{M})$ to obtain *optimum* few *regression quantiles*.

4 Trimmed estimates of regression parameters

In this section, we discuss the trimmed L-estimators of the regression and scale parameters with continuous weight functions as in the location-scale case in Bennet (1952) extended by Chernoff, Gastwirth and Johns (1967).

Let us consider the spacings $\lambda_{r+j} = (r+j)/(n+1),\ j = 1, \ldots, n-2r$ and let the regression quantiles be

$$\hat{\boldsymbol{\beta}}_j(\lambda) = \left(\hat{\beta}_j\Big(\frac{r+1}{n+1}\Big), \ldots, \hat{\beta}_j\Big(\frac{n-2r}{n+1}\Big)\right)',\ j = 0, 1, \ldots, p$$

for the corresponding spacings. Now, using Bennett's (1952) approximation, we obtain

$$\frac{1}{n}\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1} := \int_{\lambda_r}^{1-\lambda_r} \Big\{\frac{q_0'(u)}{q_0(u)}\Big\}^2 du + \frac{q_0^2(\lambda_r)}{\lambda_r} + \frac{q_0^2(\lambda_{n-r})}{1-\lambda_{n-r}} = \mathbf{I}_{11}^0$$

$$\frac{1}{n}Q_0'(\lambda)\boldsymbol{\Omega}^{-1}Q_0(\lambda) := \int_{\lambda_r}^{1-\lambda_r} \Big\{1 + Q_0(u)\frac{q_0'(u)}{q_0(u)}\Big\}^2 du + \frac{Q_0^2(\lambda_r)q_0^2(\lambda_r)}{\lambda_r} + \frac{q_0^2(\lambda_{n-r})}{1-\lambda_{n-r}} = \mathbf{I}_{22}^0$$

$$\frac{1}{n}\mathbf{1}'\boldsymbol{\Omega}^{-1}Q_0(\lambda) := \int_{\lambda_r}^{1-\lambda_r} \frac{q_0'(u)}{q_0(u)}\Big\{1 + Q_0(u)\frac{q_0'(u)}{q_0(u)}\Big\} du + \frac{Q_0^2(\lambda_r)q_0^2(\lambda_r)}{\lambda_r} + \frac{Q_0(\lambda_{n-r})q_0^2(\lambda_{n-r})}{1-\lambda_{n-r}} = \mathbf{I}_{12}^0$$

Further, for $i = r+2, \ldots, n-r-1$, we get

$$\frac{1}{n}\mathbf{1}'\boldsymbol{\Omega}^{-1} := \frac{q_0(\lambda_i)[\{q_0(\lambda_{i+1}) - q_0(\lambda_i)\} - \{q_0(\lambda_i) - q_0(\lambda_{i-1})\}]}{\lambda_{i+1} - \lambda_i}$$
$$:= -q_0(\lambda_i)\frac{d^2 q_0(\lambda_i)}{d\lambda_i^2} d\lambda_i - \frac{1}{n}\varphi_1(\lambda_i)(\text{say})$$

with $i = [n\lambda_i] + 1$. In particular if $i = r+1$, we get

$$\frac{1}{n}\mathbf{1}'\boldsymbol{\Omega}^{-1} := \frac{1}{n}\varphi_1(\lambda_{r+1}) + \frac{q_0^2(\lambda_{r+1})}{\lambda_{r+1}} - q_0'(\lambda_{r+1})$$

Similarly for $i = n-r$, we get

$$\frac{1}{n}\mathbf{1}'(\lambda)\boldsymbol{\Omega}^{-1} := \frac{1}{n}\varphi_1(\lambda_{n-r}) - \frac{q_0^2(\lambda_{n-r})}{\lambda_{n-r}} - q_0'(\lambda_{n-r})$$

Again, for $i = r+2, \ldots, n-r-1$, we have

$$\frac{1}{n}Q_0'(\lambda)\boldsymbol{\Omega}^{-1} := \frac{1}{n}\varphi_2(\lambda_i) = q_0(\lambda_i)\frac{d^2 Q_0(\lambda_i)q_0(\lambda_i)}{d\lambda_i^2}$$

and for $i = r+1$ and for $i = n-r$, we get

$$\frac{1}{n}Q_0'(\lambda)\boldsymbol{\Omega}^{-1} := \frac{1}{n}\varphi_2(\lambda_{r+1}) + \frac{Q_0(\lambda_{r+1})q_0(\lambda_{r+1})}{\lambda_{r+1}} - \{q_0(\lambda_{r+1}) + Q_0(\lambda_{r+1})q_0'(\lambda_{r+1})\} \text{ and}$$
$$\frac{1}{n}Q_0'(\lambda)\boldsymbol{\Omega}^{-1} := \frac{1}{n}\varphi_2(\lambda_{n-r}) - \frac{Q_0(\lambda_{n-r})q_0(\lambda_{n-r})}{\lambda_{n-r}} - \{q_0(\lambda_{n-r}) + Q_0(\lambda_{n-r})q_0'(\lambda_{n-r})\}$$

respectively. From the above calculation, the functions $\phi_1(u)$ and $\phi_2(u)$ are of the form

$$\varphi_1(u) = -q_0(u)\frac{d^2 q_0(u)}{du^2} = -\frac{d}{du}\{\frac{q_0'(u)}{q_0(u)}\}$$
$$\varphi_2(u) = -q_0(u)\frac{d^2\{uq_0(u)\}}{du^2} = -\{\frac{q_0'(u)}{q_0(u)} + Q_0(u)\frac{d}{du}[\frac{q_0'(u)}{q_0(u)}]\}$$

Thus, to obtain trimmed ABLUE of $(\beta_0, \ldots, \beta_p; \sigma)$ we solve the normal equations

$$\begin{bmatrix} I_{11}^0 & \bar{x}_{01}I_{11}^0 & \cdots & \bar{x}_{0p}I_{11}^0 & I_{12}^0 \\ \bar{x}_{01}I_{11}^0 & (s_{11}^2 + \bar{x}_{01}^2)I_{11}^0 & \cdots & (s_{1p} + \bar{x}_{01}\bar{x}_{0p})I_{11}^0 & \bar{x}_{01}I_{12}^0 \\ \vdots & \vdots & \ddots & \vdots & \\ \bar{x}_{0p}I_{11}^0 & (s_{p1} + \bar{x}_{0p}\bar{x}_{01}) & \cdots & (s_{pp}^2 + \bar{x}_{0p}^2)I_{11}^0 & \bar{x}_{0p}I_{12}^0 \\ I_{12}^0 & \bar{x}_{01}I_{12}^0 & \cdots & \bar{x}_{0p}I_{12}^0 & I_{22}^0 \end{bmatrix} \begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \vdots \\ \beta_p^* \\ \sigma^* \end{bmatrix} = \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_p \\ V_{p+1} \end{bmatrix},$$

where

$$V_0 = Z_0 + \bar{x}_{01}Z_1 + \cdots + \bar{x}_{0p}Z_p, \quad V_j = \bar{x}_{0j}V_0 + s_{j1}Z_1 + \cdots + s_{jp}Z_p$$
$$V_{p+1} = Z_0^* + \bar{x}_{01}Z_1^* + \ldots + \bar{x}_{0p}Z_p^*$$
$$Z_j = \frac{1}{n}\sum_{i=r+1}^{n-r}\varphi_1\Big(\frac{i}{n+1}\Big)\beta_j\Big(\frac{i}{n+1}\Big) \text{ and } Z_j^* = \frac{1}{n}\sum_{i=r+1}^{n-r}\varphi_2\Big(\frac{i}{n+1}\Big)\hat{\beta}_j\Big(\frac{i}{n+1}\Big)$$

with $j = 0, 1, \ldots, p$.

Thus, the trimmed ABLUE of the regression as well as the scale parameter is given by

$$\begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \vdots \\ \beta_p^* \\ \sigma^* \end{bmatrix} = \begin{bmatrix} I_{11}^0 & \bar{x}_{01}I_{11}^0 & \cdots & \bar{x}_{0p}I_{11}^0 & I_{12}^0 \\ \bar{x}_{01}I_{11}^0 & (s_{11}^2 + \bar{x}_{01}^2)I_{11}^0 & \cdots & (s_{1p} + \bar{x}_{01}\bar{x}_{0p})I_{11}^0 & \bar{x}_{01}I_{12}^0 \\ \vdots & \vdots & \ddots & \vdots & \\ \bar{x}_{0p}I_{11}^0 & (s_{p1} + \bar{x}_{0p}\bar{x}_{01}) & \cdots & \bar{x}_{0p-1}I_{11}^0 & \bar{x}_{0p}I_{12}^0 \\ I_{12}^0 & \bar{x}_{01}I_{12}^0 & \cdots & \bar{x}_{0p}I_{12}^0 & I_{22}^0 \end{bmatrix}^{-1} \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_p \\ V_{p+1} \end{bmatrix}$$

The covariance matrix of trimmed ABLUE of the parameters is given by

$$\sigma^2 \begin{bmatrix} I_{11}^0 & \bar{x}_{01}I_{11}^0 & \cdots & \bar{x}_{0p}I_{11}^0 & I_{12}^0 \\ \bar{x}_{01}I_{11}^0 & (s_{11} + \bar{x}_{01}^2)I_{12}^0 & \cdots & (s_{1p} + \bar{x}_{01}\bar{x}_{0p})I_{11}^0 & \bar{x}_{01}I_{12}^0 \\ \vdots & \vdots & \ddots & \vdots & \\ \bar{x}_{0p}I_{11}^0 & (s_{p1} + \bar{x}_{0p}\bar{x}_{01})I_{11}^0 & \cdots & (s_{pp} + \bar{x}_{0p}^2)I_{11}^0 & \bar{x}_{0p}I_{12}^0 \\ I_{12}^0 & \bar{x}_{01}I_{12}^0 & \cdots & \bar{x}_{0p}I_{12}^0 & I_{22}^0 \end{bmatrix}^{-1}$$

It may be noted that these trimmed estimators are similar to Bennetts (1952) estimator for the location and scale parameters for the location and scale model of distribution. Computation of estimators will be highly computer-intensive which is not a problem now a days. In a separate paper, the author will demonstrate the application of these formulas derived here.

5 Simple linear model and estimation using optimum regression quantiles

Consider the simple linear model will the known error distribution F_0,

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{x}_i + \mathbf{e}_i,\ i = 1, \dots, n.$$

Let the spacing vector be $\lambda_1, \dots, \lambda_k$ and for $j = (0, 1)$, the associated regression quantiles be $\hat{\boldsymbol{\beta}}_j(\lambda) = (\hat{\beta}_j(\lambda_1), \hat{\beta}_j(\lambda_2), \dots, \hat{\beta}_j(\lambda_k))'$. Then, under the assumed condition the normal equations are given by the following expression by letting $\bar{x} = \bar{x}_0$ and $c^2 = \lim \frac{1}{n} \sum x_i^2$ we have

$$\begin{bmatrix} K_1 & \bar{x}_0 K_1 & K_3 \\ \bar{x}_0 K_1 & (s_{11} + \bar{x}_0^2 K_1) & \bar{x}_0 K_3 \\ K_3 & \bar{x}_0 K_3 & K_2 \end{bmatrix} \begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \sigma^* \end{bmatrix} = \begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix}$$

using $p = 1$. The solution for the estimators becomes

$$\begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \sigma^* \end{bmatrix} = \begin{bmatrix} \frac{K_2}{\Delta} + \frac{\bar{x}_0^2}{s_{11} K_1} & -\frac{\bar{x}}{s_{11} K_1} & -\frac{K_3}{\Delta} \\ -\frac{\bar{x}_0}{s_{11} K_1} & \frac{1}{s_{11} K_1} & 0 \\ -\frac{K_3}{\Delta} & 0 & \frac{K_1}{\Delta} \end{bmatrix} \begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix},$$

$$V_0 = Z_0 + \bar{x}_0 Z_1, V_1 = \bar{x}_0 Y_0 + s_{11} Z_1, V_2 = Z_0^* + \bar{x}_0 Z_1^*$$

The covariance matrix of the estimators is given by

$$Cov(\beta_1^*, \beta_2^*, \sigma^*) = \frac{\sigma^2}{n} \begin{bmatrix} \frac{K_2}{\Delta} + \frac{\bar{x}_0^2}{s_{11} K_1} & -\frac{\bar{x}_0}{s_{11} K_1} & -\frac{K_3}{\Delta} \\ -\frac{\bar{x}_0}{s_{11} I_1} & \frac{1}{s_{11} K_1} & 0 \\ -\frac{K_3}{\Delta} & 0 & \frac{K_1}{\Delta} \end{bmatrix}$$

The covariance matrix of the LSE is given by

$$\frac{\sigma^2}{n} \begin{bmatrix} \frac{\mathbf{I}_{22}}{|\mathbf{I}|} + \frac{\bar{x}_0^2}{s_{11} \mathbf{I}_{11}} & -\frac{\bar{x}_0}{s_{11} \mathbf{I}_{11}} & -\frac{\mathbf{I}_{12}}{|\mathbf{I}|} \\ -\frac{\bar{x}_0}{s_{11} \mathbf{I}_{11}} & \frac{1}{s_{11} \mathbf{I}_{11}} & 0 \\ -\frac{\mathbf{I}_{12}}{|\mathbf{I}|} & 0 & \frac{\mathbf{I}_{11}}{|\mathbf{I}|} \end{bmatrix}$$

with $\mathbf{I} = (I_{ij})$ is the information for the location-scale distribution F_0. Then ARE of the ABLUE relative to LSE has the same expression as (2.16) with p=1. Thus, we maximize $K_1\Delta$ with respect to $\lambda_1, \ldots, \lambda_k$. If $F_0(x) = (1+e^{-x})^{-1}$, i.e. logistic distribution, it is then known that the optimum spacings vector is given by $((k+1)^{-1}, 2(k+1)^{-1}, \ldots, k(k+1)^{-1})$. Thus, for $k = 3$, we get the popular Gastwirths (1966) trimean type robust estimators in the case of this linear model.

Let $F_0(x) = 1 - e^{-x}$. In this case errors should be subtracted by 1. Then, one can easily verify that (see, Saleh and Ali, 1966; Saleh, 1981;).

$$K_1 = 1/e^{u_1} - 1,\ K_2 = (e^{u_1} - 1)\{u_1^2 + L(e^{u_1} - 1)\},\ L = \sum_{i=2}^{k} \frac{(u_i - u_{i-1})^2}{e^{u_i} - e^{u_{i-1}}},$$

$$K_3 = u_1(e^{u_1} - 1)^{-1} \text{and } \Delta = L(e^{u_1} - 1)^{-1}, u_i = \log(1-\lambda_i)^{-1},\ i = 1, 2, \ldots, k$$

The ARE is given by $K_1\Delta$.

Thus $K_1\Delta = (e^{u_1} - 1)^{-2} e^{-u_1} Q_{k-1}$, where $Q_{k-1} = \sum_{i=1}^{k-1} (t_i - t_{i-1})^2/(e^{t_i} - e^{t_{i-1}})$ as in Saleh and Ali (1966). Thus, maximizing $K_1\Delta$ $w.r.t.$ $(\lambda_1, \ldots, \lambda_k)$ one gets the optimum solution. Thus, we get

$$\lambda_1^* = \left(n + \frac{1}{2}\right)^{-1} \text{ and } \lambda_{j+1}^* = \frac{2 + (2n-1)\lambda_j^0}{2n+1},\ j = 1, \ldots, k-1,$$

where $\lambda_j^0 (j = 1, \ldots, k-1)$ are the optimum spacings for the scale-parameter alone which are available in Sarhan and Greenberg (1962). For example for $k = 5$, we have $\lambda_1^0 = .3931$, $\lambda_2^0 = .6670$, $\lambda_3^0 = .8434$, $\lambda_4^0 = .9434$ and $\lambda_5^0 = .9885$. Thus, one can use these spacings to obtain the spacings (5.1) for the six optimum regression quantiles for the estimation of β_0, β_1 and σ.

As for the conditional quantile-function for a given $\xi \in (0, 1)$

$$\begin{aligned} y(\xi) &= \beta_0 + \beta_1 x_0 + \sigma Q_0(\xi), \qquad & Q_0(\xi) &= \ln(1-\xi)^{-1} \\ &= l'(\beta_0, \beta_1, \sigma)', & &= l(1, x_0, Q_0(\xi))' \end{aligned}$$

we use the estimator $y^*(\xi) = \beta_0^* + \beta_1^* x_0 + \sigma^* Q_0(\xi)$. To obtain optimum spacings one has to maximize the maximum characteristic root of the matrix $\mathbf{M}$ given by

$$\mathbf{M} = \begin{bmatrix} \frac{1}{n-1} + \frac{\bar{x}_0^2}{ns_{11}} & -\frac{\bar{x}_0}{ns_{11}} & -\frac{1}{n-1} \\ -\frac{\bar{x}_0}{ns_{11}} & \frac{1}{ns_{11}} & 0 \\ -\frac{1}{n-1} & 0 & \frac{n}{n-1} \end{bmatrix} \begin{bmatrix} K_1 & \bar{x}_0 K_1 & K_3 \\ \bar{x}_0 K_1 & (s_{11} + \bar{x}_0^2) K_1 & \bar{x}_0 K_3 \\ K_3 & \bar{x}_0 K_3 & K_2 \end{bmatrix}$$

or tr($\mathbf{M}$) $w.r.t.$ $(\lambda_1, \ldots, \lambda_k)$. Here, the first matrix is the covariance matrix of the maximum likelihood estimators.

6 Marginal estimation of the intercept, scale and the quantile-function of the distribution

In the linear model $\mathbf{Y} = \beta_0 \mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \sigma\mathbf{e},\ (\beta_0, \sigma)$ represents the location (intercept) and the scale parameter. Further, the parameters $Q(\xi) = \beta_0 + \sigma Q_0(\xi),\ (0 < \xi < 1)$ will be called the quantile-function of the distribution F_0. If we are particularly interested in the estimation of β_0, σ and $Q(\xi)$, we may consider the regression-quantiles $\hat{\beta}_0(\lambda) = (\hat{\beta}_0(\lambda_1), \ldots, \hat{\beta}_0(\lambda_k))'$ corresponding to the spacing vector $\lambda = (\lambda_1, \ldots, \lambda_k)'$. Thus, from Section 2, we find $\sqrt{n}[\hat{\beta}_0(\lambda) - \beta_0 \mathbf{1}_k - \sigma Q_0(\lambda)] - N_k(0, \sigma^2 \mathbf{C}^{11} \mathbf{\Omega})$, where $\mathbf{C}^{-1} = (\mathrm{C}^{ij})$ and Ω is defined in (2.2). Thus, minimizing

$$n[\hat{\beta}_0(\lambda) - \beta_0 \mathbf{1}_k - \sigma Q_0(\lambda)]\mathbf{\Omega}^{-1}[\hat{\beta}_0(\lambda) - \beta_0 \mathbf{1}_k - \sigma Q_0(\lambda)]$$

one obtains the ABLUE of (β_0, σ) as

$$\beta_0^* = \frac{K_2 Z_0 - K_3 Z_0^*}{\Delta} \text{ and } \sigma^* = \frac{K_1 Z_0^* - K_3 Z_0}{\Delta},$$

where Δ and K_1, K_2, K_3 have the same definition as in (2.3-2.4). The variance-covariance matrix for (β_0^*, σ^*) is given by

$$\mathrm{Cov}(\beta_0^*, \sigma^*) = \frac{\sigma^2 \mathbf{C}^{11}}{n\Delta} \begin{bmatrix} K_2 & -K_3 \\ -K_3 & K_1 \end{bmatrix}$$

Also, the ABLUE of $Q(\xi)$ is given by

$$Q^*(\xi) = \beta_0^* + \sigma^* Q_0(\xi), \ 0 < \xi < 1$$

with asymptotic variance

$$\mathrm{Var}(Q^*(\xi)) = \frac{\sigma^2 \mathbf{C}^{11}}{n\Delta}[K_2 + K_1 Q_0^2(\xi) - 2K_3 Q_0(\xi)].$$

The corresponding LSE has the asymptotic variance given by

$$\mathrm{Var}(Q^{**}(\xi)) = \frac{\sigma^2 \mathbf{C}^{11}}{n \mid \mathbf{I} \mid}\{I_{22} + I_{11} Q_0^2(\xi) - 2I_{12} Q_0(\xi)\}.$$

The ARE of $Q^*(\xi)$ relative to $Q^{**}(\xi)$ is then given by

$$\mathrm{ARE}(Q^*(\xi) : Q^{**}(\xi)) = \frac{\Delta\{I_{22} + I_{11} Q_0^2(\xi) - 2I_{12} Q_0(\xi)\}}{\mid \mathbf{I} \mid [K_2 + K_1 Q_0^2(\xi) - 2K_3 Q_0(\xi)]}.$$

The numerator and denominator may be written as $l'\mathbf{I}^{-1}l$ and $l'\mathbf{K}^{-1}l$, where

$$\mathbf{I} = \begin{bmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{bmatrix} \text{ and } \mathbf{K} = \begin{bmatrix} K_1 & K_3 \\ K_3 & K_2 \end{bmatrix}$$

and $l' = (1, Q_0(\xi))$. Thus, ARE is the expression

$$\frac{l'\mathbf{I}^{-1}l}{l'\mathbf{K}^{-1}l} \text{ and } \mathrm{Ch}_{min}(\mathbf{I}^{-1}\mathbf{K}) \leq \frac{l'\mathbf{I}^{-1}l}{l'\mathbf{K}^{-1}l} \leq \mathrm{Ch}_{max}(\mathbf{I}^{-1}\mathbf{K}).$$

Thus, to obtain the optimum spacings for (β_0^*, σ^*) and $Q^*(\xi)$ the maximum Δ and $\mathrm{Ch}_{max}(\mathbf{I}^{-1}\mathbf{K})$ or $\mathrm{tr}(\mathbf{I}^{-1}\mathbf{K})$ respectively $w.r.t.$ $(\lambda_1, \ldots, \lambda_k)'$. The problem have been discussed in many publications listed in Saleh (1992) and available.

References

[1] Bennett, C.A. (1952). Asymptotic properties of ideal linear estimators. Ph.D. Thesis, University of Michigan.

[2] Chernoff, H., Gastwirth, J.L. and Johns, M.V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38**, *52-72.*

[3] Harter, H.L. (1961). Estimating the parameters of the negative exponential populations using one or two order statistics. *Ann. Math. Statist.* **32**, *1078-1084.*

[4] Gastwirth, J.L. (1966). Robust Procedures. *JASA,* **65**, *946-973.*

[5] Guttenbrunner, C. and Jureckova, J. (1992). Regression rank scores and regression quantiles, *Ann. Statist.* **20**, *305-329.*

[6] Jureckova, J. (1984). Regression quantiles and trimmed least squares estimator under a general design. *Kybernetika,* **20**, *345-357.*

[7] Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica,* **46**, *33-50.*

[8] Koenker, R. and Portnoy, S. (1987). L-estimation for linear models. *JASA,* **82**, *851-857.*

[9] Ogawa, J. (1951). Contributions to theory of systematic statistics. *Osaka Math. Jour.,* **3**, *175-213.*

[10] Ruppert, D. and Carroll, R.J. (1980). Trimmed least-squares estimation in the linear model, *JASA,* **75**, *828-838.*

[11] Saleh, A.K.Md.E. (1992). Contributions to order statistics, Department of Math/Statistics. Carleton University, Ottawa.

[12] Saleh, A.K.Md.E. (1981). Estimating quantiles in exponential distribution, Statistics and Related Topics (Edited by M. Csorgo, D.A. Dawson, J.N.K. Rao and A.K.Md. E. Saleh). North-Holland Publishing Company. Amsterdam.

[13] Saleh, A.K.Md.E. and Ali, M.M. (1966). Asymptotic optimum quantiles for the estimation of the parameters of the negative exponential distribution. *Ann. Math. Statist.* **37**, *143-151.*

[14] Sarhan, A.E. and Greenberg, B.G. (1962). Contributions to Order Statistics. Johns Wiley & Sons, New York.