Ann. Inst. Statist. Math.
Vol. 44, No. 2, 387–403 (1992)

# EXPANSIONS FOR STATISTICS INVOLVING THE MEAN ABSOLUTE DEVIATIONS

GUTTI JOGESH BABU* AND C. RADHAKRISHNA RAO**

*Statistics Department, 219 Pond Laboratory, Pennsylvania State University, University Park, PA 16802, U.S.A.*

(Received May 22, 1990; revised October 1, 1990)

**Abstract.** Expansion for the difference of mean absolute deviations from the sample mean and the population mean is derived. This result is used to obtain strong representations for mean absolute deviations from the sample mean and the sample median. Edgeworth expansions for some scale invariant statistics involving the mean absolute deviations are studied. These expansions are shown to be valid in spite of the presence of a lattice variable.

*Key words and phrases*: Mean absolute deviations, Edgeworth expansions, quantiles, linear model, $L_1$-norm.

## 1. Introduction

Recent research on robust statistical inference has motivated the development of statistical methodology based on the $L_1$-norm in preference to the traditional methods based on the $L_2$-norm or least squares. However, from time to time attempts have been made to use the $L_1$-norm in estimation and tests of significance, but the complexity of the distributions involved stood in the way of their use in practical applications.

For instance, if $X_1, \dots, X_n$ is an i.i.d. sample from a distribution function $F$, the alternatives

$$(1.1) \qquad \bar{X} = \text{mean}\{X_1, \dots, X_n\} \quad \text{and} \quad \tilde{X} = \text{median}\{X_1, \dots, X_n\}$$

were considered as estimators of a location parameter, and the mean absolute deviations

$$(1.2) \qquad M_1 = n^{-1} \sum |X_i - \bar{X}| \quad \text{and} \quad M_2 = n^{-1} \sum |X_i - \tilde{X}|$$

* Research supported in part by NSA Grant MDA904-90-H-1001.

** Research supported by the Air Force Office of Scientific Research under Grant AFOSR-89-0279.

were considered as estimates of a scale parameter as alternatives to the root mean square deviation

$$s = \left[n^{-1} \sum (X_i - \bar{X})^2\right]^{1/2}. \tag{1.3}$$

Godwin (1943) found the distribution of $M_1$ when $F$ is normal, but the expression for the density turned out to be very complicated. Geary (1936) considered one of the following statistics ($W_1$)

$$W_1 = \frac{M_1}{s}, \quad W_2 = \frac{M_2}{s} \tag{1.4}$$

as a test of the normality of $F$. In fact, he obtained the Edgeworth expansion of the distribution of $W_1$ when $F$ is normal. Herrey (1965) discussed the use of one of the following statistics ($t_1$)

$$t_1 = \frac{\sqrt{n}(\bar{X} - \mu)}{M_1} \quad \text{and} \quad t_2 = \frac{\sqrt{n}(\bar{X} - \mu)}{M_2} \tag{1.5}$$

to obtain a confidence interval for the unknown mean $\mu$ of the population when $F$ is normal. His analysis rests on the fact that $\bar{X} - \mu$ and $M_1$ are independent for samples from a normal distribution. Two other possibilities for this purpose are

$$t_3 = \frac{\sqrt{n}(\tilde{X} - \mu)}{M_1} \quad \text{and} \quad t_4 = \frac{\sqrt{n}(\tilde{X} - \mu)}{M_2}. \tag{1.6}$$

In this paper, we derive the asymptotic representations of the statistics

$$M_1, M_2; W_1, W_2; t_1, t_2 \tag{1.7}$$

for a general d.f. $F$, and obtain Edgeworth expansions under some moment conditions.

We also consider the Gauss-Markov linear model, define statistics of the form (1.7) depending on $L_1$ and $L_2$-norms, and study their asymptotic distributions.

We use the elementary integral representation

$$|x| - |x - a| = a \int_0^1 \operatorname{sign}(x - ya)\,dy \tag{1.8}$$

to express the statistics (1.7) as a sum of i.i.d. random variables plus quadratic term plus a remainder which is negligible in large samples. The recent work of Babu and Singh (1989) is used to obtain the Edgeworth expansions.

## 2. Asymptotics of the mean absolute deviations

Let $X_1, \dots, X_n$ be i.i.d. random variables with a common d.f. $F$ such that

$$E(X_1) = \mu, \qquad V(X_1) = \sigma^2 < \infty \tag{2.1}$$

and define the statistics (mean absolute deviations)

$$M_1 = n^{-1}\sum_1^n |X_i - \bar{X}|, \qquad M_2 = n^{-1}\sum_1^n |X_i - \tilde{X}| \tag{2.2}$$

where $\bar{X}$ and $\tilde{X}$ are the sample mean and median respectively. In this section, we obtain representations of $M_1$ and $M_2$ to derive their asymptotic distributions and for later use in Section 3 in getting the Edgeworth expansions of the distributions of certain statistics. We denote

$$M_1' = n^{-1}\sum_1^n |X_i - \mu|, \qquad M_2' = n^{-1}\sum_1^n |X_i - \nu| \tag{2.3}$$

where $\mu$ and $\nu$ are the population mean and median respectively. Let $F_n$ denote the empirical distribution function

$$F_n(x) = n^{-1}\sum_1^n I(X_i \le x). \tag{2.4}$$

We have the following theorems.

THEOREM 2.1. *Suppose that for some $c > 1$ and $0 < \beta \le 1$, $F$ satisfies*

$$|F(x) - F(\mu)| \le c|x - \mu|^\beta. \tag{2.5}$$

*Then*

$$\begin{aligned} M_1 = M_1' &+ (2F_n(\mu) - 1)(\bar{X} - \mu) + (\bar{X} - \mu)R_n \\ &+ 2(\bar{X} - \mu)\int_0^1 [F(\mu + (\bar{X} - \mu)y) - F(\mu)]dy, \end{aligned} \tag{2.6}$$

*where for some $k > 0$,*

$$P(|\bar{X} - \mu| \le n^{-1/2}\log n, |R_n| > k(\log n)^{(1+\beta)/2} n^{-(2+\beta)/4}) = O(n^{-2}). \tag{2.7}$$

PROOF. Clearly by (1.8) and Lemma A.1,

$$\begin{aligned} M_1 - M_1' &= (\bar{X} - \mu)\int_0^1 n^{-1}\sum_1^n [2I(X_i - \mu \le (\bar{X} - \mu)y) - 1]dy \\ &= (\bar{X} - \mu)\left\{(2F_n(\mu) - 1) + 2\int_0^1 [F_n(\mu + (\bar{X} - \mu)y) - F_n(\mu)]dy\right\} \\ &= (\bar{X} - \mu)\left\{(2F_n(\mu) - 1) + 2\int_0^1 [F(\mu + (\bar{X} - \mu)y) - F(\mu)]dy + R_n\right\}, \end{aligned}$$

where

$$R_n = 2\int_0^1 [F_n(\mu + (\bar{X} - \mu)y) - F_n(\mu) - F(\mu + (\bar{X} - \mu)y) + F(\mu)]dy. \tag{2.8}$$

By taking $a_n$ as in Lemma A.2 to be $n^{-1/2}\log n$, we get (2.7). This completes the proof.

THEOREM 2.2. *If $F$ is differentiable in a neighborhood of $\mu$ and the derivative $f$ at $\mu$ is positive, then with probability 1,*

$$M_1 - M_1' = (\bar{X} - \mu)(2F_n(\mu) - 1) + (\bar{X} - \mu)^2(f(\mu) + o(1)) + O(n^{-5/4}(\log n)^2). \tag{2.9}$$

PROOF. By Lemma A.4, we have with probability 1

$$n^{1/2}|\bar{X} - \mu| \le \log n \tag{2.10}$$

for all large $n$. In view of (2.10) and

$$\begin{aligned}\int_0^1 [F(\mu + (\bar{X} - \mu)y) - F(\mu)]dy &= \int_0^1 y(\bar{X} - \mu)(f(\mu) + o(1))dy\\ &= \frac{1}{2}(\bar{X} - \mu)(f(\mu) + o(1)),\end{aligned}$$

the result (2.9) follows from Theorem 2.1, by taking $\beta = 1$.

THEOREM 2.3. *Suppose that for some $c > 1$ and $0 < \beta \le 1$, $F$ satisfies the condition*

$$|F(x) - F(\mu)| \le c|x - \mu|^{\beta}. \tag{2.11}$$

*Then*

$$\sqrt{n}[M_1 - M_1' - (2F(\mu) - 1)(\bar{X} - \mu)] \xrightarrow{P} 0. \tag{2.12}$$

*Consequently*

$$\sqrt{n}(M_1 - (2F(\mu) - 1)(\bar{X} - \mu) - E|X_1 - \mu|) \xrightarrow{\mathcal{D}} N(0, \sigma_1^2), \tag{2.13}$$

*where*

$$\begin{aligned}\sigma_1^2 &= V(|X_1 - \mu| + (2F(\mu) - 1)(X_1 - \mu))\\ &= (2F(\mu) - 1)^2 V(X_1) + (4F(\mu) - 2)\operatorname{cov}(|X_1 - \mu|, X_1 - \mu) + V(|X_1 - \mu|).\end{aligned}$$

*In particular, if $F(\mu) = 1/2$, i.e., the population mean and median coincide, then*

$$\sqrt{n}(M_1 - E|X_1 - \mu|) \xrightarrow{\mathcal{D}} N(0, V(|X_1 - \mu|)). \tag{2.14}$$

PROOF. Since $\sqrt{n}|\bar{X}-\mu|$ is bounded in probability and $F_n(\mu) \xrightarrow{P} F(\mu)$, (2.12) follows from Theorem 2.1, from which (2.13) follows.

THEOREM 2.4. *Let $F$ be differentiable in a neighborhood of $\mu$ and the derivative $f$ at $\mu$ be positive. We have:*

(i) *If $F(\mu) \neq 1/2$, then*

$$\sqrt{n}[2F(\mu)-1]^{-1}(M_1-M_1') \xrightarrow{\mathcal{D}} N(0, V(X_1)).$$

(ii) *If $F(\mu) = 1/2$, then*

$$n(M_1-M_1') \xrightarrow{\mathcal{D}} f(\mu)U^2 - UV,$$

*where $(U,V)$ is a bivariate normal variable with mean zero and covariance matrix*

$$\Sigma = \begin{bmatrix} V(X_1) & E|X_1-\mu| \\ E|X_1-\mu| & 1 \end{bmatrix}.$$

PROOF. The results of Theorem 2.4 follow from the representation (2.9) of Theorem 2.2, and (2.12) of Theorem 2.3, which imply that

$$\sqrt{n}[M_1-M_1'-(\bar{X}-\mu)(2F(\mu)-1)] \xrightarrow{P} 0, \quad \text{if} \quad F(\mu)\neq 1/2$$

and

$$n[(M_1-M_1')-(\bar{X}-\mu)^2 f(\mu)-(\bar{X}-\mu)(2F_n(\mu)-1)] \xrightarrow{P} 0, \quad \text{if} \quad F(\mu)=1/2.$$

THEOREM 2.5. *Let $F$ have a unique median $\nu$ and satisfy the condition*

$$|F(x)-F(\nu)| \leq c|x-\nu|^{\beta}$$

*for some $c>1$ and $0<\beta\leq 1$. If $\sqrt{n}(\tilde{X}-\nu)$ is bounded in probability, then*

$$\sqrt{n}(M_2 - E|X_1-\nu|) \xrightarrow{\mathcal{D}} N(0,\xi^2), \tag{2.15}$$

*where*

$$\xi^2 = E(X_1-\nu)^2 - (E|X_1-\nu|)^2.$$

*Remark* 1. The conditions of the theorem hold, in particular, if $F$ has a derivative $f$ in a neighborhood of $\nu$ and $f(\nu)>0$.

PROOF. Using the result (1.8)

$$M_2 - M_2' = (\tilde{X}-\nu)\int_0^1 [2F_n(\nu+(\tilde{X}-\nu)y)-1]dy.$$

Since

$$\sup_{0<y<1}\left|F_n(\nu+(\tilde{X}-\nu)y)-\frac{1}{2}\right| \xrightarrow{P} 0,$$

we have

$$\sqrt{n}(M_2 - M_2') \xrightarrow{P} 0.$$

The result (2.15) follows since

$$\sqrt{n}(M_2' - E|X_1-\nu|) \xrightarrow{\mathcal{D}} N(0,\xi^2).$$

This completes the proof.

Let $\hat{\mu}_\alpha$ and $\mu_\alpha$ denote the $\alpha$-th sample and population quantiles respectively. Then we have the following theorem:

THEOREM 2.6. *Suppose that $F$ has a continuous derivative $f$ in a neighborhood of $\mu_\alpha$ and $f(\mu_\alpha) > 0$. Then, with probability* 1,

$$\begin{aligned}\sum_{i=1}^{n}(|X_i-\hat{\mu}_\alpha| - |X_i-\mu_\alpha|) &= -n(1-2\alpha)(\hat{\mu}_\alpha-\mu_\alpha) - n(\hat{\mu}_\alpha-\mu_\alpha)^2(f(\mu_\alpha)+o(1)) \\ &\quad + O(n^{-1/4}(\log n)^{3/2}).\end{aligned} \tag{2.16}$$

PROOF. Note that by Lemma A.3, we have with probability 1

$$\hat{\mu}_\alpha - \mu_\alpha = O(n^{-1/2}(\log n)^{1/2}).$$

Then the result (2.16) follows from Lemmas A.1 and A.2 using the same lines of proof as in Theorems 2.1 and 2.2.

From (A.6), we have

COROLLARY 2.1. *Under the conditions of Theorem* 2.6, *if $\alpha = 1/2$, we have with probability* 1 *from* (2.16),

$$n(M_2' - M_2) = n(\tilde{X}-\nu)^2(f(\nu)+o(1)) + O(n^{-1/4}(\log n)^2).$$

As a consequence, it follows that

$$\frac{n}{4}f(\nu)(M_2'-M_2) \xrightarrow{\mathcal{D}} \chi_1^2. \tag{2.17}$$

## 3. Edgeworth expansions for some scale invariant statistics

In this section we obtain the Edgeworth expansions for the distributions of the scale invariant statistics

$$W_1 = \frac{M_1}{s}, \quad W_2 = \frac{M_2}{s} \tag{3.1}$$

$$t_1 = \frac{\sqrt{n}(\bar{X}-\mu)}{M_1}, \quad t_2 = \frac{\sqrt{n}(\bar{X}-\mu)}{M_2} \tag{3.2}$$

using the asymptotic representations of $M_1$ and $M_2$ derived in Section 2. It turns out that the main terms in the asymptotic expansions of these statistics involve lattice as well as non-lattice random variables and the standard results on Edgeworth expansions of their distributions are not applicable. However, the lattice variables appear only in the quadratic term for $W_1$ and $W_2$, in which case the recent results of Babu and Singh (1989) enable the evaluation of two-term Edgeworth expansion for $W_1$ and $W_2$. By using a generalization (see Babu (1991)) of the results of Babu and Singh (1989), we develop three term Edgeworth expansions for $t_i$, $i = 1, 2$.

The following notations are used

$$\begin{aligned} &V(X_1) = E(X_1-\mu)^2 = \sigma^2, \\ &Y_i = (X_i - \mu)/\sigma, \quad \gamma_1 = E|Y_1|, \quad \bar{Y} = n^{-1}\sum Y_i, \\ &Z_i = |X_i - \nu|/\sigma, \quad \gamma_2 = E(Z_1), \quad \bar{Z} = n^{-1}\sum Z_i. \end{aligned}$$

We now state and prove the main theorems.

THEOREM 3.1. *Suppose that $F$ has a continuous density $f$ in a neighborhood of $\mu$ and $f(\mu) > 0$. For some $c > 0$ and $\beta > 0$, let $|f(\mu) - f(x)| < c|x-\mu|^\beta$ in a neighborhood of $\mu$. If $E(X_1^6) < \infty$, then we have*

$$W_1 - \gamma_1 = \bar{L} - \bar{Y}\bar{H} + \frac{1}{2}\overline{1-Y^2}\;\overline{|Y|-\gamma_1} + \frac{3}{8}\gamma_1(\overline{1-Y^2})^2 + R_{n1} \tag{3.3}$$

*where the notation $\bar{a}$ is used for the average of $a_1, \ldots, a_n$,*

$$L_i = |Y_i| - \frac{1}{2}\gamma_1(1+Y_i^2), \qquad H_i = \operatorname{sign} Y_i - \left[\sigma f(\mu) + \frac{1}{2}\gamma_1\right] Y_i,$$

*and*

$$P(|R_{n1}| > \kappa n^{-1-\epsilon}) = o(n^{-1/2})$$

*for some $\kappa > 0$ and $\epsilon > 0$.*

PROOF. The theorem follows from Theorem 2.2 and the observation that

$$\frac{\sigma}{s} - 1 = \frac{1}{2}(1-(s/\sigma)^2) + \frac{3}{8}(1-(s/\sigma)^2)^2 + O((1-(s/\sigma)^2)^3)$$

and the moderate deviation result of Michel (1976) that for some $A$

$$P(\sqrt{n}|\bar{X}-\mu| > A\log n) = o(n^{-1}). \tag{3.4}$$

(In fact Michel (1976) shows that (3.4) holds under the condition $E(X_1^4) < \infty$.)

THEOREM 3.2. *Suppose that $F$ has a continuous density $f$ in the neighborhood of $\nu$ and $f(\nu) > 0$. For some $c > 0$ and $\beta > 0$, let $|f(\nu) - f(x)| < c|x-\nu|^\beta$ in a neighborhood of $\nu$. If $E(x_1^6) < \infty$, then we have*

$$\begin{aligned} W_2 - \gamma_2 = \bar{D} &- (1/4\sigma f(\nu))(1-2F_n(\nu))^2 + \frac{1}{2}\overline{Z-\gamma_2}\;\overline{1-Y^2} \\ &+ \gamma_2\left[\frac{1}{2}\bar{Y}^2 + \frac{3}{8}(\overline{1-Y^2})^2\right] + R_{n2}, \end{aligned} \tag{3.5}$$

*where*

$$D_i = Z_i - \frac{1}{2}\gamma_2(1+Y_i^2) \quad \textit{and} \quad P(|R_{n2}| > \kappa n^{-1-\epsilon}) = o(n^{-1/2})$$

*for some $\kappa > 0$ and $\epsilon > 0$.*

Proof of this theorem is similar to that of Theorem 3.1, and so is omitted.

Note that if $\nu = \mu$, then $Z_i = |Y_i|$ and $\gamma_1 = \gamma_2$ and hence $D_i = L_i$. Further, in both the representations (3.3) and (3.5) the lattice variables appear only in the quadratic term. So the results of Babu and Singh (1989) which generalize some of the results of Bhattacharya and Ghosh (1978), apply and both $W_1$ and $W_2$ have valid two-term Edgeworth expansions. The next theorem gives these expansions.

THEOREM 3.3. *Under the conditions of Theorem 3.1, we have uniformly in $x$*

$$P(\sqrt{n}(W_1-\gamma_1) \le x) = \Phi_{\eta_1^2}(x) + \frac{1}{\sqrt{n}}p_1(x)\varphi_{\eta_1^2}(x) + o(n^{-1/2}). \tag{3.6}$$

*Under the conditions of Theorem 3.2, we have uniformly in $x$*

$$P(\sqrt{n}(W_2-\gamma_2) \le x) = \Phi_{\eta_2^2}(x) + \frac{1}{\sqrt{n}}p_2(x)\varphi_{\eta_2^2}(x) + o(n^{-1/2}). \tag{3.7}$$

*In (3.6) and (3.7),*

$$\begin{aligned} \eta_1^2 &= E(L_1^2) = 1 + \left(\frac{1}{2}\gamma_1\right)^2 [E(X_1-\mu)^4\sigma^{-4} - 1] - \gamma_1 E|X_1-\mu|^3\sigma^{-3}, \\ \eta_2^2 &= E(D_1^2) = E(X_1-\nu)^2\sigma^{-2} + \frac{1}{2}\gamma_2[E(X_1-\nu)^4\sigma^{-4} - 1] \\ &\qquad - \gamma_2 E[(X_1-\mu)^2|X_1-\nu|\sigma^{-3}], \\ p_1(x) &= \kappa_1 + \frac{1}{6}\kappa_{13}(1-(x^2/\eta_1^2)), \\ p_2(x) &= \kappa_2 + \frac{1}{6}\kappa_{23}(1-(x^2/\eta_2^2)), \end{aligned}$$

*where*

$$\begin{aligned}\kappa_1 &= \frac{1}{8}\gamma_1[5 - 3\sigma^{-4}E(X_1-\mu)^4] - \sigma f(\mu) + \frac{1}{2}\sigma^{-3}E|X_1-\mu|^3,\\ \kappa_2 &= \frac{1}{4f(\nu)} - \frac{1}{2}\gamma_2 + \frac{1}{2}E[(X_1-\mu)^2|X_1-\nu|\sigma^{-3}],\\ \eta_1^2\kappa_{13} &= E(L_1^3) - 6E(L_1Y_1)E(L_1H_1) - 3E(L_1|Y_1|E(L_1Y_1^2))\\ &\quad + \frac{9}{4}\gamma_1(E(L_1Y_1^2))^2,\\ \eta_2^2\kappa_{23} &= E(D_1^3) - \frac{3}{2f(\nu)}E[D_1\operatorname{sign}(X_1-\nu)]^2 - 3E(D_1Z_1)E(D_1Y_1^2)\\ &\quad + 3(E(D_1Y_1))^2 + \frac{9}{4}\gamma_2(E(D_1Y_1^2))^2.\end{aligned}$$

*Remark* 2. If $F$ is symmetric, then

$$\eta_1^2\kappa_{13} = \eta_2^2\kappa_{23} = E(L_1^3) - 3E(|Y_1|L_1)E(Y_1^2L_1) + \frac{9}{4}\gamma_1(E(L_1Y_1^2))^2.$$

*Remark* 3. $\gamma_1 = E(|X-\mu|/\sigma)$ is known for some parametric families. For example in Gaussian case, $\gamma_1 = \sqrt{2/\pi}$. The order of error in (3.6) does not change if the coefficients of the polynomial $p_1$ are replaced by their sample estimates.

In view of the unknown parameter $\eta_1$, (3.6) may not be of much help in practice. However, such expansions are essential, in establishing superiority of the bootstrap estimate. By some additional algebra, it is possible to show, as in Liu and Singh (1987), that the bootstrap distribution of $\sqrt{n}(W_1-\gamma_1)$ gives a better approximation to its sampling distribution than $\Phi_{\eta_{1,n}^2}(x)$, in terms of asymptotic mean squared error, where $\eta_{1,n}^2$ is an estimator of $\eta_1^2$. Similar comments apply throughout this section.

We now consider representations and three-term Edgeworth expansions for $t_1$ and $t_2$ defined in (3.2).

From Theorem 2.1 and (3.4) we have the following representations.

THEOREM 3.4. *Suppose that $F$ has a continuous density $f$ in a neighborhood of $\mu$ and $f(\mu) > 0$. For some $c > 0$ and $\beta > 0$, let $|f(\mu) - f(x)| < c|x-\mu|^\beta$ in a neighborhood of $\mu$. If $E(X_1^4) < \infty$, then*

$$t_1 = \frac{\sqrt{n}}{\gamma_1\sigma}(\overline{X-\mu})(1 - \bar{U} + \bar{V}(\overline{X-\mu})\sigma^{-1} + \bar{U}^2) + R_{n3} \tag{3.8}$$

*where*

$$\begin{aligned}\sigma\gamma_1 U_i &= |X_i-\mu| - \gamma_1 + (2F(\mu)-1)(X_i-\mu),\\ \gamma_1 V_i &= 2F(\mu) - 2I(X_i \le \mu) - f(\mu)(X_i-\mu)\end{aligned}$$

*and*

$$P(|R_{n3}| > \kappa n^{-1-\epsilon}) = o(n^{-1})$$

*for some* $\kappa > 0$ *and* $\epsilon > 0$.

THEOREM 3.5. *Under the conditions of Theorem* 3.4, *if* $\mu = \nu$, *then*

$$t_2 = \frac{\sqrt{n}}{\gamma_1 \sigma}(\bar{X} - \mu)\left(1 - \bar{U} + \bar{U}^2 + \frac{(1 - 2F_n(\nu))^2}{4\gamma_1 f(\nu)}\right) + R_{n4} \tag{3.9}$$

*where* $P(|R_{n4}| > \kappa n^{-1-\epsilon}) = o(n^{-1})$ *for some* $\kappa > 0$ *and* $\epsilon > 0$.

The results of Babu and Singh (1989) do not give a valid three-term Edgeworth expansion for $t_1$ and $t_2$. However, a refinement of it due to Babu (1991) provides the desired expansion.

THEOREM 3.6. *Under the conditions of Theorem* 3.4, *we have*

$$\begin{aligned} P(t_1 \le x) &= \Phi_{\gamma_1 - 2}(x) \\ &\quad + \frac{1}{\gamma_1 \sqrt{n}}\left(P_1(\gamma_1 x) + \frac{1}{\gamma_1 n}[P_2(\gamma_1 x) + P_{2i}(\gamma_1 x)]\right) \\ &\quad \cdot \varphi_{\gamma_1 - 2}(x) + o(n^{-1}) \end{aligned} \tag{3.10}$$

*for* $i = 1$. (3.10) *also holds for* $i = 2$ *if in addition* $\mu = \nu$.

In (3.10),

$$\begin{aligned} P_1(x) &= \kappa_1 + \kappa_3((1 - x^2)/6), \\ P_2(x) &= -\frac{\kappa_2}{2}x + \frac{\kappa_4}{24}(3x - x^3) + \frac{\kappa_3^2}{72}(10x^3 - 15x - x^5), \\ P_{21}(x) &= -x^3 E(Y_1 V_1) = x^3(\eta f(\mu) - \gamma_1)/\gamma_1, \\ P_{22}(x) &= [(\gamma_1 - (1/\gamma_1))x - \gamma_1 x^3]/(4f(\mu)), \end{aligned}$$

where

$$\begin{aligned} \kappa_1 &= E(U_1 Y_1) = [E(Y_1|Y_1|) + 2F(\mu) - 1]/\gamma_1 \\ &= [E(\sigma^{-2}(X_1 - \mu)^2 \operatorname{sign}(X_1 - \mu)) + 2F(\mu) - 1]/\gamma_1, \\ \kappa_2 &= 6\kappa_1^2 - 2E(U_1 Y_1^2) + 3E(U_1^2) \\ &= 6\kappa_1^2 - 2\gamma_1^{-2}[1 - \gamma_1^2 + (2F(\mu) - 1)^2 + (4F(\mu) - 2)E(Y_1^2 \operatorname{sign} Y_1)] \\ &\quad + 3E(U_1^2), \\ \kappa_3 &= E[\sigma^{-3}(X_1 - \mu)^3] + 6\kappa_1, \\ \kappa_4 &= E(Y_1^4) - 3 + 16\kappa_1 E(Y_1^3) - 12E(U_1 Y_1^2) + 12E(U_1^2) + 84\kappa_1^2. \end{aligned}$$

*Remark* 4. If $X_1$ is symmetrically distributed, then clearly $\kappa_1 = \kappa_3 = 0$ and hence $P_1(x) \equiv 0$. Further $U_i = \gamma_1^{-1}|Y_i| - 1$ so that

$$\begin{aligned} \kappa_2 &= 3\gamma_1^{-2} - 1 - 2\gamma_1^{-1}\sigma^{-3}E|X_1 - \mu|^3, \\ \kappa_4 &= E[(X_1 - \mu)^4\sigma^{-4}] - 3 - 12\gamma_1^{-2}(1 - \gamma_1\sigma^{-3}E|X_1 - \mu|^3). \end{aligned}$$

*Remark* 5. $\gamma_1$ here is the shape parameter which is known for some parametric families. The coefficients of the polynomials $P_2$, $P_{21}$ and $P_{22}$ can be replaced by their estimates without changing the magnitude of the error. For symmetric populations $P_1 \equiv 0$. So

$$P(t_1 \le x) = \Phi_{\gamma_1-2}(x) + \frac{1}{\hat{\gamma}_1^2 n}(\hat{P}_2(\hat{\gamma}_1 x) + \hat{P}_{2i}(\hat{\gamma}_1 x))\varphi_{\hat{\gamma}_1-2}(x) + o(n^{-1}),$$

where $\hat{}$'s signal that the quantities are estimated. In the Gaussian case $\gamma_1 = \sqrt{2/\pi}$, $\kappa_1 = \kappa_3 = 0$,

$$\kappa_2 = (3r/2) - 1 - \sqrt{2r} \quad \text{and} \quad \kappa_4 = -12 + 6\sqrt{2\pi}.$$

Some of the results on Edgeworth expansions may be derived using Bai and Rao (1991), instead of Babu and Singh (1989) and Babu (1991). However, it should be noted that, in our case, it is not trivial to verify Assumption 4 of Bai and Rao (1991).

## 4. Statistics associated with linear models

In this section, we consider the Gauss-Markov linear model, define statistics similar to those introduced in Sections 2 and 3 and discuss their asymptotic distributions. To avoid complicated notation, we consider a simple model. Let

$$y_i = x_{in}\beta + \epsilon_i, \qquad i = 1, \dots, n, \tag{4.1}$$

where $\epsilon_i, \dots, \epsilon_n$ are i.i.d. random variables, $x_{in}$ are $m$-dimensional row vectors such that $\sum x'_{in}x_{in} = I$ and the true value of $\beta$ is zero without loss of generality.

We make the following assumptions

($\mathrm{A}_1$) The common d.f. $F$ of $\epsilon_i$ has median zero, has a derivative $f$ in the neighborhood of zero and for some $c > 0$

$$f(0) > 0, \quad F(0) = \frac{1}{2} \quad \text{and} \quad |f(y) - f(0)| \le c|y|^{1/2} \tag{4.2}$$

($\mathrm{A}_2$)

$$q_n(\log n)^{1/2} \to 0 \quad \text{as} \quad n \to \infty, \tag{4.3}$$

where $q_n = \max\{\|x_{in}\| : 1 \le i \le n\}$. Clearly $nq_n^2 \ge m$ using the condition $\sum x'_{in}x_{in} = I$.

We denote the least absolute deviations (LAD) and least square (LS) estimates of $\beta$ by $\tilde{\beta}$ and $\hat{\beta}$ respectively.

Using the results of Babu (1989), the following theorem can be established.

THEOREM 4.1. *Under the assumptions* ($\mathrm{A}_1$) *and* ($\mathrm{A}_2$) *we have:*

(i) *For some* $B > 0$,

$$P(\|\tilde{\beta}\| > B(\log n)^{1/2}) = O(n^{-2}). \tag{4.4}$$

(ii)

$$2f(0)\tilde{\beta} = \sum_1^n x'_{in} \operatorname{sign} y_i + R_{n1} \tag{4.5}$$

*where for some* $c > 0$, $P(|R_{n1}| > cq_n^{1/2}(\log n)^{3/4}) = O(n^{-2})$.

(iii)

$$\sum_{i=1}^n (|y_i| - |y_i - x_{in}\tilde{\beta}|) = \frac{1}{4f(0)} \left\| \sum_1^n x_{in} \operatorname{sign} y_i \right\|^2 + R_{n2} \tag{4.6}$$

*where for some* $c > 0$, $P(|R_{n2}| > cq_n^{1/2}(\log n)^{5/4}) = O(n^{-2})$.

(iv) *Using Borel-Cantelli Lemma, it follows that the errors* $R_{n1}$ *and* $R_{n2}$ *in* (4.5) *and* (4.6) *satisfy with probability* 1

$$R_{n1} = O(q_n^{1/2}(\log n)^{3/4}) \quad \textit{and} \quad R_{n2} = O(q_n^{1/2}(\log n)^{5/4}). \tag{4.7}$$

THEOREM 4.2. *Under the assumptions* ($A_1$) *and* ($A_2$) *and the additional assumption on the LS estimation* $\hat{\beta}$

$$P(\|\hat{\beta}\| > A(\log n)^{1/2}) = o(n^{-1/2}) \quad \textit{or} \quad o(n^{-1}) \tag{4.8}$$

*we have*

$$\sum_1^n |y_i - x_{in}\hat{\beta}| = \sum_1^n |y_i| + f(0)\|\hat{\beta}\|^2 + 2\sum_1^n x_{in}[I(y_i \leq 0) - F(0)]\hat{\beta} + R_n \tag{4.9}$$

*where for some* $B > 0$

$$P(|R_n| > Bq_n^{1/2}(\log n)^{5/4}) = o(n^{-1/2}) \quad \textit{or} \quad o(n^{-1}). \tag{4.10}$$

*If the stronger condition*

$$\hat{\beta} = O((\log n)^{1/2}) \quad \textit{with probability } 1, \tag{4.11}$$

*holds for the LS estimator* $\hat{\beta}$, *then*

$$R_n = O(q_n^{1/2}(\log n)^{5/4}) \quad \textit{with probability } 1. \tag{4.12}$$

*Remark* 6. If $E(\epsilon_i) = 0$, $E(\epsilon_1^4) < \infty$ and $q_n = O(n^{-1/4}(\log n)^{-1})$ then by Lemma A.5 of the Appendix, (4.8) holds with $o(n^{-1/2})$. By similar arguments, it can be shown that (4.8) holds with $o(n^{-1})$ if in addition $E(\epsilon_1^6) < \infty$.

In summary we have the representations:

$$2f(0)\tilde{\beta} = \sum_1^n x_{in}' \operatorname{sign} y_i + R_{n1}, \tag{4.13}$$

$$\hat{\beta} = \sum_1^n x_{in}' y_i, \tag{4.14}$$

$$\begin{aligned} S_1 &= \sum_1^n |y_i - x_{in}\hat{\beta}| \\ &= \sum_1^n |y_i| + \sum_1^n x_{in}(2I(y_i \le 0) - 1 + y_i f(0))\hat{\beta} + R_{n2}, \end{aligned} \tag{4.15}$$

$$S_2^2 = \sum_1^n (y_i - x_{in}\hat{\beta})^2 = \sum_1^n y_i^2 - \|\hat{\beta}\|^2, \tag{4.16}$$

$$S_3 = \sum_1^n |y_i - x_{in}\tilde{\beta}| - \sum_1^n |y_i| - \frac{1}{4f(0)} \left\| \sum_1^n x_{in} \operatorname{sign} y_i \right\|^2 + R_{n3}, \tag{4.17}$$

$$S_4^2 = \sum_1^n (y_i - x_{in}\tilde{\beta})^2 = \sum_1^n y_i^2 - \|\hat{\beta}\|^2 + \|\hat{\beta} - \tilde{\beta}\|^2, \tag{4.18}$$

where the errors $R_{n1}$, $R_{n2}$ and $R_{n2}$ can be ignored in determining Edgeworth expansions for statistics based on (4.13)–(4.18) provided $q_n(\log n)^{5/2} = o(1)$.

Formal Edgeworth expansions can be obtained for the distributions of statistics of the form $S_i/S_j$, $i \neq j$, but to examine their validity we need to develop new tools. We content ourselves in giving the asymptotic distributions of some statistics which may be useful in statistical data analysis.

Let the error term in the linear model (4.1) have a symmetric distribution. Then:

(i) The asymptotic distribution of $(\hat{\beta}, \tilde{\beta})$ is $2m$-variate normal with zero means and variance covariance matrix

$$\begin{pmatrix} E(y_1^2) & (2f(0))^{-1}E|Y_1| \\ (2f(0))^{-1}E|Y_1| & (2f(0))^{-2} \end{pmatrix} \otimes I_m.$$

(ii) The asymptotic distribution of

$$\hat{\beta}, \quad \sqrt{n}\left(\frac{1}{n}S_1 - E|Y_1|\right) \quad \text{or} \quad \sqrt{n}\left(\frac{1}{n}S_3 - E|Y_1|\right)$$

is $(m+1)$ variate normal with zero means and variance-covariance matrix

$$\begin{pmatrix} E(Y_1^2)I_m & 0 \\ 0 & E(Y_1^2)-(E|Y_1|)^2 \end{pmatrix}.$$

(iii) The asymptotic distribution of

$$\tilde{\beta}, \quad \sqrt{n}\left(\frac{1}{n}S_1 - E|Y_1|\right) \quad \text{or} \quad \sqrt{n}\left(\frac{1}{n}S_3 - E|Y_1|\right)$$

is $(m+1)$ variate normal with zero means and variance-covariance matrix

$$\begin{pmatrix} (2f(0))^{-2}I_m & 0 \\ 0 & E(Y_1^2)-(E|Y_1|)^2 \end{pmatrix}.$$

(iv) The asymptotic distribution of

$$\frac{1}{\sqrt{n}}(S_1, S_2^2, S_3, S_4^2) - \sqrt{n}(E|y_1|, \operatorname{Var} y_1, E|y_1|, \operatorname{Var} y_1)$$

is 4-dimensional normal $(R, S, R, S)$, where $(R, S)$ is bivariate normal with zero means and variance-covariance matrix

$$\begin{pmatrix} E(y_1^2)-(E|y_1|)^2 & E(|y_1|^3)-E(|y_1|)E(y_1^2) \\ E(|y_1|^3)-E(|y_1|)E(y_1^2) & E(y_1^4)-E(y_1^2) \end{pmatrix}.$$

## Appendix

The following lemma is an immediate consequence of (1.8).

LEMMA A.1. *For any $\theta_1$ and $\theta_2$, we have*

$$\frac{1}{2}(f_n(\theta_2)-f_n(\theta_1)) = \left(\frac{1}{2}-F_n(\theta_i)\right)(\theta_2-\theta_1) + \int_{\theta_1}^{\theta_2}(F(\theta_i)-F(x))dx + R_n(\theta_i) \tag{A.1}$$

*for $i=1,2$, where for any $\zeta$*

$$\begin{aligned} f_n(\zeta) &= \frac{1}{n}\sum_1^n(|X_i|-|X_i-\zeta|) \quad \text{and} \\ R_n(\zeta) &= \int_{\theta_1}^{\theta_2}(F_n(\zeta)-F_n(x)-F(\zeta)+F(x))dx. \end{aligned} \tag{A.2}$$

*Further, if $F$ has a derivative $f$ at $\theta_1$, then*

$$\int_{\theta_1}^{\theta_2}(F(\theta_2)-F(x))dx = \frac{1}{2}(\theta_2-\theta_1)^2(f(\theta_1)+o(1)) \tag{A.3}$$

*and*

$$\text{(A.4)} \qquad \int_{\theta_1}^{\theta_2} (F(\theta_1) - F(x))dx = -\frac{1}{2}(\theta_2 - \theta_1)^2 (f(\theta_1) + o(1))$$

*as* $\theta_2 \to \theta_1$.

LEMMA A.2. *For a fixed* $\theta$, *let*

$$\text{(A.5)} \qquad |F(x) - F(\theta)| \le c|x - \theta|^\beta,$$

*for some* $c \ge 1$ *and* $0 < \beta \le 1$. *Let* $((\log n)/n) \le a_n^\beta \le 1$ *and*

$$G_n(\theta) = \max\{|g_n(x,\theta)| : |x - \theta| \le a_n\},$$

*where* $g_n(x,\theta) = F_n(x) - F_n(\theta) - F(x) + F(\theta)$. *Then*

$$\text{(A.6)} \qquad P(G_n(\theta) > Kb_n) \le 6n^{-2},$$

*where* $b_n = n^{-1/2}(\log n)^{1/2} a_n^{\beta/2}$ *and* $K = 2c + 2 + (2\beta)^{-1}$.

PROOF. Note that for any $|y| \le 1$, $e^y \le 1 + y + y^2$. So for any $x$ in $[\theta - a_n, \theta + a_n]$, $A > 0$ and $0 < \eta \le 1$, we have by (A.5) that

$$\text{(A.7)} \qquad \begin{aligned} P(g_n(x,\theta) > A) &\le e^{-nA\eta}(1 + \eta^2 |F(x) - F(\theta)|)^n \\ &\le \exp(-nA\eta + n\eta^2|F(x) - F(\theta)|) \\ &\le \exp(-nA\eta + cn\eta^2 a_n^\beta). \end{aligned}$$

By putting $\eta = b_n a_n^{-\beta}$ and $A = (c + 2 + (2\beta)^{-1})b_n$ in (A.7), we get that $0 < \eta \le 1$ and that the r.h.s. of (A.7) is not more than

$$\text{(A.8)} \qquad \exp(-(2 + (2\beta)^{-1})\log n) = n^{-(2+(2\beta)^{-1})}.$$

We now let $\nu_0 = \theta - a_n$,

$$\begin{aligned} &\nu_{i+1} = \nu_i + b_n^{1/\beta} \qquad \text{for} \qquad i = 0, 1, \ldots, r-2, \\ &\nu_r = \theta + a_n, \\ &0 \le \nu_r - \nu_{r-1} \le b_n^{1/\beta}. \end{aligned}$$

Clearly $1 \le r \le 2a_n b_n^{-1/\beta} \le 2n^{1/2\beta} \le 3n^{1/2\beta} - 1$ and

$$\text{(A.9)} \qquad G_n(\theta) \le \max_{0 \le j \le r} |g_n(\nu_j, \theta)| + \max_{1 \le j \le r} (F(\nu_j) - (F(\nu_{j-1})).$$

It follows from (A.5), (A.8) and (A.9) that

$$\begin{aligned} P(G_n(\theta) > Kb_n) &\le \sum_{j=0}^{r} P(|g_n(\nu_j, \theta)| > (K - c)b_n) \\ &\le 2(r+1)n^{-2-(2\beta)^{-1}} \le 6n^{-2}. \end{aligned}$$

This completes the proof.

*Remark* A.1. It is obvious from Lemma A.2 that with probability 1,

$$\limsup_{n\to\infty} G_n(\theta)b_n^{-1} \le K. \tag{A.10}$$

LEMMA A.3. *Suppose that $F$ has a continuous derivative $f$ in a neighborhood of $\mu_\alpha$ and $f(\mu_\alpha) > 0$. Then with probability* 1,

$$\hat{\mu}_\alpha - \mu_\alpha = O(n^{-1/2}(\log n)^{1/2}). \tag{A.11}$$

PROOF. Let $d_n = n^{-1/2}(\log n)^{1/2}$ and

$$F_n^{-1}(t) = \inf\{x : F_n(x) \ge t\}$$

for $0 < t < 1$. Since $F$ is continuous in a neighborhood of $\mu_\alpha$ it follows that for all large $n$,

$$F(F^{-1}(y)) = y \quad \text{for all} \quad y \in [\mu_\alpha - d_n, \mu_\alpha + d_n].$$

Further, for any $c > 0$, we have on $D = \{\sup_x |F_n(x) - F(x)| \le cd_n\}$,

$$\begin{aligned} F_n(F^{-1}(\alpha - 2cd_n)) &\le F(F^{-1}(\alpha - 2cd_n)) + cd_n < \alpha \\ &< F(F^{-1}(\alpha + 2cd_n)) - cd_n \le F_n(F^{-1}(\alpha + 2cd_n)). \end{aligned}$$

Hence on $D$,

$$F^{-1}(\alpha - 2cd_n) \le \hat{\mu}_\alpha \le F^{-1}(\alpha + 2cd_n).$$

Since $F^{-1}$ is differentiable in a neighborhood of $\alpha$, the lemma follows from the well known result that with probability 1,

$$\sup_x |F_n(x) - F(x)| = O(d_n).$$

LEMMA A.4. *Let $\{Z_n\}$ be a sequence of i.i.d. random variables with mean zero and $E(Z_i^2) < \infty$. Then with probability* 1, *for all large $n$,*

$$\sqrt{n}|\bar{Z}| \le \log n.$$

PROOF. Let

$$Z_i' = Z_i I(|Z_i| \le \sqrt{i}\,) \quad \text{for} \quad i = 1, 2, \dots .$$

By Markov inequality

$$P\left(2\left|\sum_{i=1}^n Z_i'\right| \ge \sqrt{n}\log n\right) \le n^{-2}\prod_{i=1}^n E[\exp(4Z_i' n^{-1/2})]. \tag{A.12}$$

But

$$E(\exp(4Z_i'n^{-1/2})) \leq 1 + 4n^{-1/2}E(|Z_i|I(|Z_i| > \sqrt{i})) + 8E(Z_1^2)n^{-1}e^2$$
$$= 1 + O((in)^{-1/2}) + O(n^{-1}).$$

So the r.h.s. of (A.12) is $O(n^{-2})$. Since

$$\sum_{i=1}^{\infty} P(Z_i \neq Z_i') \leq \sum_{i=1}^{\infty} P(|Z_i| > \sqrt{i}) = \sum_{i=1}^{\infty} P(|Z_1| > \sqrt{i})$$
$$\leq 1 + E(Z_1^2) < \infty,$$

it follows by the Borel Cantelli lemma that with probability 1, $Z_i = Z_i'$ for all large $n$ and $|\sum_{i=1}^{n} Z_i'| < \sqrt{n}(\log n)/2$ for all large $n$. These two facts together imply the result.

LEMMA A.5. *Let $E(y_1) = 0$ and $Ey_1^4 < \infty$. Let $q_n = O(n^{-1/4}(\log n)^{-1})$, then for some $A > 0$*

$$P(\|\hat{\beta}\| > A(\log n)^{1/2}) = o(n^{-1/2}).$$

The lemma can be proved using a result on exponential inequality similar to Lemma 2 of Babu (1989).

## REFERENCES


Babu, G. J. (1989). Strong representations for LAD estimators in linear models, *Probab. Theory Related Fields*, **83**, 547–558.

Babu, G. J. (1991). Edgeworth expansions for statistics which are functions of lattice and non-lattice variables, *Statist. Probab. Lett.*, **12**, 1–7.

Babu, G. J. and Singh, K. (1989). On Edgeworth expansions in mixture cases, *Ann. Statist.*, **17**, 443–447.

Bai, Z. D. and Rao, C. R. (1991). Edgeworth expansions of a function of sample means, *Ann. Statist.*, **19**, 1295–1315.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.*, **6**, 434–451.

Geary, R. C. (1936). Moments of the ratio of the mean deviation to the standard deviation for normal samples, *Biometrika*, **28**, 295–305.

Godwin, H. J. (1943). On the distribution of the estimate of mean deviation obtained from samples from a normal population, *Biometrika*, **33**, 254–256.

Herrey, E. M. J. (1965). Confidence intervals based on the mean absolute deviation of a normal sample, *J. Amer. Statist. Assoc.*, **60**, 257–269.

Liu, R. Y. and Singh, K. (1987). On partial correction by the bootstrap, *Ann. Statist.*, **15**, 1713–1718.

Michel, R. (1976). Non-uniform central limit bounds with applications to probabilities of deviations, *Ann. Probab.*, **4**, 102–106.