1988

# Granularity and efficiency

Chong Sun Hong
*Iowa State University*

Order Number 8825922

# Granularity and efficiency

Hong, Chong Sun, Ph.D.
Iowa State University, 1988

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

**Granularity and efficiency**

by

Chong Sun Hong

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1988

## TABLE OF CONTENTS

# 1. INTRODUCTION

## 1.1. Background and Motivation

When rounding to a grid is considered, three possible ways of computing the sample estimates ( e.g., the sample mean, the sample median, and least-squares estimate, etc.) of population parameters might be considered.

1. Compute $X_i$ exactly and round the sample estimates computed on the basis of the exact $X_i$'s, denoted by $(\hat{\theta})_r$.
2. Round $X_i$'s and compute the sample estimate exactly, on the basis of the rounded $X_i$'s, denoted by $(\hat{\theta}_r)$.
3. Round $X_i$'s and, in addition, round the sample estimate computed on the basis of the rounded $X_i$'s, denoted by $(\hat{\theta}_r)_r$.

The analysis of relative efficiency asymptotics in the above first and third cases involves large deviations in a natural way, reminiscent of Bahadur's [2,3] and Bahadur and Rao's [4] asymptotic comparison of tests. The "outside rounding" in $(\hat{\theta})_r$ and $(\hat{\theta}_r)_r$ has the effect of converting comparative variance ratios to comparative large-deviation rates.

The "inside rounding" in $(\hat{\theta}_r)$ and $(\hat{\theta}_r)_r$ was first considered in Sheppard's correction (see Kendall and Stuart [31]). Also, "inside rounding" has been considered in the context of order statistics by David and Mishriky [20]. Kempthorne [30], Giesbrecht and Kempthorne [23], and Lambert [34,35] considered the asymptotic maximum-likelihood theory for "inside rounding" from the point of view of the asymptotic maximum-likelihood theory for parametrized multinomial distributions.

We are primarily concerned with the first rounding case, with $\hat{\theta}$ taken to be the sample mean, the sample median, and least-squares estimates. The rounded sample mean $\overline{X}_\epsilon$ and the rounded sample median $M_\epsilon$, respectively, are computed by rounding the sample mean $\overline{X}$ and the sample median $M$ to the nearest point of a grid $\{2\epsilon i\ ;\ \ i = 0, \pm 1, \pm 2, \ldots\}$, $\epsilon$ assumed greater than zero, with upward rounding when $\overline{X}$ or $M$ falls half-way between two grid-points. The rounded sample median $M_\epsilon$ is also computable as the median of rounded sample values. Rounded least-squares estimates $b_{j,\epsilon_j}$ in regression models are computed in similar fashion.

Hammersley [26] considered, among other matters, the asymptotic relative efficiency (ARE) of the above rounded sample median $M_\epsilon$ with respect to the above rounded sample mean $\overline{X}_\epsilon$, as estimates of a Normal population mean $\mu$ restricted to a uniform grid of mesh size $2\epsilon$. Here, via the theory of large deviations (Bahadur [2,3]; Bahadur and Rao [4]), we extend Hammersley's work, to a certain class of "TEIFR" distributions in order to examine the Normal multivariate case and the Normal $p$-variate regression model.

Asymptotic comparison of $\overline{X}_\epsilon$ and $M_\epsilon$ would be especially appropriate if one or the other could be shown to be fully efficient in some sense; say, for example, if $\overline{X}_\epsilon$ (respectively, $M_\epsilon$) could be shown in some sense to be most precise among unbiased grid-valued estimators of a grid-valued Normal (respectively, Laplace) population center $\mu$. Such results, however, seem not to be available. Of course, in those cases where the distribution of $\overline{X}$ or of $M$ is symmetric about $\mu$ and decreases monotonically on either side of $\mu$, the rounded versions $\overline{X}_\epsilon$ or $M_\epsilon$ maximize the likelihood of $\overline{X}$ or $M$ over the grid. In addition, Hammersley [26] pointed out that $\overline{X}_\epsilon$ is the maximum likelihood estimate for $\mu$ under Normality (as is $M_\epsilon$ under the Laplace distribution). Hammersley [26] did attempt a Cramér-Rao analysis for the Normal case, hoping thereby to obtain a benchmark for $\overline{X}_\epsilon$, but left unresolved the question of obtaining a sharp lower bound for the variance of any unbiased estimate of the grid-restricted Normal mean. Somewhat more conclusive evidence was obtained by Ghosh and Meeden [22], who showed that $\overline{X}_\epsilon$ is admissible in the Normal case among all grid-valued estimators, when estimating a grid-valued $\mu$, for loss functions $L(\mu, d) = W(\mu - d)$ such that $W(t)$ is non-negative symmetric and increasing in $|t|$, with $W(2j\epsilon) \leq K(2j\epsilon)^u$, $\quad j = 1, 2, \ldots$, for some $u$, $K > 0$.

Whereas the sample mean is an average of *iid* random variables $\{X_i \; ; \; i = 1, 2, \ldots, n\}$, a least-squares estimators $b_j$ in the $p$-variate Normal regression model is a linear function of non-*iid* random variables $\{Y_i \; ; \; i = 1, 2, \ldots, n\}$, though, in fact, when the corresponding population parameters $\beta_j$ is subtracted from $b_j$, $b_j - \beta_j$ reduces to a linear function of *iid* random variables. Here we consider the asymptotic variances and covariances of the rounded least-squares estimators (RLSE) $b_{j,\epsilon_j}$, where $b_{j,\epsilon_j}$ is obtained by rounding, to the grid of mesh size $2\epsilon_j$, the least-squares estimator of $\beta_j$ for $j = 1, 2, \ldots, p$. The asymptotic variances and covariances of these rounded regression estimates are analyzable in terms of univariate and bivariate large deviations, in a manner analogous to the analysis of the rounded sample mean $\overline{X}_\epsilon$.

We may define joint asymptotic efficiency (JAE) as the determinant of the asymptotic variance-covariance matrix. The JAE of the RLSEs in the $p$-variate regression model is a function of the sum of the large-deviation rates for the asymptotic variances of the individual RLSEs, since the asymptotic covariances are negligible. Multivariate large deviations are also used to study rounded sample means and rounded sample medians in the multivariate setting, under a certain bivariate analogue of the TEIFR condition. As in the Normal regression case, multivariate JAEs of rounded sample means and rounded sample medians involve summations of univariate large-deviation rates. This is illustrated for the Normal and Laplace distributions.

A possible, albeit *non-linear*, context for RLSEs is provided by the work of Sankoorikal, Danofsky, David, Hendrickson and Tollefson [39]. That investigation had the aim of devising methodology for locating the vibrating rod among a two-dimensional grid of rods in a nuclear reactor core; vibration signal magnitudes at certain detector sites within the core were modeled as dependent regression variables, with expectations expressed as (non-linear) functions of the respective known detector locations and the unknown 2 dimensional coordinates of the vibrating rod. Thus the rod location was a grid-valued two-dimensional regression parameter appearing in several non-linear regression equations (one each for the several detector sites), and the proposed estimate of the location of the vibrating rod is obtained by rounding (to the nearest point of the two-dimensional grid of rods) the vector regression coefficient estimated on the basis of the several observed detector signal magnitudes.

## 1.2. Findings

We consider the rounded sample mean $\overline{X}_\epsilon$ and the rounded sample median $M_\epsilon$ as estimates of the grid-valued location parameter $\mu$ under a certain class of "two-sided extended IFR (TEIFR)" distribution. The role of the TEIFR assumption is to insure that the tails of the distribution of $X_i - \mu$ fall off quickly enough to make the comparison of asymptotic probabilities of large (beyond $\epsilon$) deviations of the location-normalized sample median $M - \mu$ and sample mean $\overline{X} - \mu$ relevant to the comparison of their asymptotic mean-square errors (MSEs). A finding is that a TEIFR distribution possesses a moment generating function, as well as tail probabilities that satisfy a condition necessary for the large-deviation treatment of the rounded sample median $M_\epsilon$. Thus the asymptotic mean-square errors of the rounded sample median $M_\epsilon$ and the rounded sample mean $\overline{X}_\epsilon$ are formulated via the theory of large-deviation (Bahadur and Rao [4]; Bahadur [2,3]). We find the following:

Suppose a random sample $\{X_i \ ; \quad i = 1, 2, \ldots, n\}$ is drawn from a TEIFR distribution whose location parameter is $\mu$ and whose variance is $\sigma^2$. Then the asymptotic mean-square error $MSE_n(M_\epsilon)$ of the sample median rounded to the grid $(2k\epsilon \ ; \quad k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\lim_{n \to \infty} n^{-1} \ln MSE_n(M_\epsilon) = \ln[2(R_\epsilon(1 - R_\epsilon))^{1/2}],$$

where

$$R_\epsilon = \max\{\Pr(X_i - \mu \geq \epsilon), \quad \Pr(X_i - \mu \leq -\epsilon)\}.$$

And the asymptotic mean-square error $MSE_m(\overline{X}_\epsilon)$ of the sample mean, rounded to the grid $(2k\epsilon\ ;\quad k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\lim_{m\to\infty} m^{-1} \ln MSE_m(\overline{X}_\epsilon) = -\frac{\epsilon^2}{2\sigma^2}(1 + \delta(\epsilon)),$$

when $\epsilon$ is sufficiently small, where $m$ denotes sample sizes, and the function $\delta(\cdot)$ satisfies

$$\lim_{\epsilon\to 0} \delta(\epsilon) = 0.$$

The asymptotic relative efficiency $e_\epsilon$ of $M_\epsilon$ with respect to $\overline{X}_\epsilon$ is found, when $\epsilon$ is small, to equal:

$$e_\epsilon = -2\sigma^2\epsilon^{-2}(1 + \delta(\epsilon))^{-1} \ln[2(R_\epsilon(1 - R_\epsilon))^{1/2}].$$

The asymptotic relative efficiency $e_\epsilon$ is computed as the limiting ratio of "equivalent sample sizes" $n$ and $m$.

We find $e_\epsilon$ to be surprisingly sensitive to distribution shape, as well as to the grid mesh size and $e_\epsilon$ is right-continuous in $\epsilon$ at $\epsilon = 0$. In other words,if the population distribution possesses a density $f$ in a neighborhood of $\mu$, then

$$\lim_{\epsilon\to 0} e_\epsilon = 4f^2(\mu)\sigma^2,$$

the RHS being the usual asymptotic relative efficiency $e_0$ of $M$ vs. $\overline{X}$ for populations possessing finite variance and a density in a neighborhood of the population median. We make the point in this connection that "continuity at zero" does not hold for the form of asymptotic efficiency dealt with by Hammersley [26]. A finding related to the continuity of $e_\epsilon$ at $\epsilon = 0$ is that, within the TEIFR class, the "asymptotic effective variance" in the sense of Bahadur [2] of the sample median $M$ equals its asymptotic variance as usually defined.

We expand the analysis of $e_\epsilon$ to $\epsilon$'s of arbitrary size, in the special Laplace and Normal cases. We point to a certain equivocal behavior of asymptotic relative efficiency away from $\epsilon = 0$ in the case of the Laplace distribution. Which of $M_\epsilon$ and $\overline{X}_\epsilon$ is more efficient turns out to depend on the value of $\epsilon$. No such equivocal behavior occurs in the Normal case, where the increasing superiority of $\overline{X}_\epsilon$ over $M_\epsilon$ with increasing $\epsilon$ is conveniently quantified via a certain quartic lower bound for a symmetric version of Mill's ratio $R(\epsilon)$:

$$
\begin{aligned}
R(\epsilon)R(-\epsilon) &\equiv \frac{\Phi(\epsilon)\Phi(-\epsilon)}{\phi(\epsilon)\phi(-\epsilon)} \\
&\geq \frac{\pi}{2}\exp[(\frac{\pi-2}{\pi})\epsilon^2 + (\frac{1}{2\pi})\epsilon^4].
\end{aligned}
$$

These last findings derive from comparing the rates of decrease of the probabilities of large (beyond $\epsilon$) deviations, respectively of $|M - \mu|$ and $|\overline{X} - \mu|$. Hence they pertain as well to the comparison of the asymptotic error rates of certain tests based respectively on $M$ and $\overline{X}$, since these tests' asymptotic error rates essentially are themselves large deviation rates. Thus the findings have testing counterparts.

We also consider the asymptotic variances and covariances of the rounded least-squares estimators (RLSEs) $b_{j,\epsilon_j}$ in Normal $p$-variate regression models, in terms of the large deviations of $\Pr(|b_j - \beta_j| \geq \epsilon_j)$ and $\Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \geq \epsilon_j)$, respectively, where $b_i$ is the least-squares estimator of the grid-valued parameter $\beta_i$ for $i = 1, 2, \ldots, p$, rounded to the grid. The asymptotic variances and covariances of these rounded regression estimates are analyzable in terms of their large deviations in a manner analogous to the analysis of the rounded sample mean $\overline{X}_\epsilon$.

To implement the regression analysis, some lemmas extend to the $b_j$'s the treatment of large deviations of sample means in Bahadur and Rao [4] and Bahadur [2,3]. These lemmas yield large-deviation rates for regression parameter vectors $\underline{b}$:

$$n^{-1} \ln \Pr(\underline{b} - \underline{\beta} \geq \underline{\epsilon}) = n^{-1}(-\frac{1}{2}\underline{\epsilon}' \underline{V} \underline{\epsilon}) + o_n(1),$$

where $\underline{V}$ is the variance-covariance matrix of $\underline{b}$.

The asymptotic variances and covariances of the RLSEs are formulated as follows:

The asymptotic variance $Var_n(b_{j,\epsilon_j})$ of the RLSE $b_{j,\epsilon_j}$, obtained by rounding, to the grid $(2k\epsilon_j\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\lim_{n\to\infty} n^{-1} \ln Var_n(b_{j,\epsilon_j}) = \lim_{n\to\infty} -\frac{\epsilon_j^2}{2n\sigma_{n,j}^2} = -\frac{\epsilon_j^2 Q_{j,j}}{2\sigma^2},$$

where $\underline{Q}$ is a positive definite matrix such that $\lim_{n\to\infty} n^{-1}\underline{X}'\underline{X} = \underline{Q}$ and $Q_{j,j}$ is the $(j,j)$ element of $\underline{Q}$.

The asymptotic covariance $Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j})$ of the RLSEs $b_{i,\epsilon_i}$ and $b_{j,\epsilon_j}$, rounded the grids $(2k\epsilon_i,\ 2k\epsilon_j\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$ satisfies

$$\begin{aligned}\lim_{n\to\infty} n^{-1} \ln sgn(corr_{i,j}) Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) &= \lim_{n\to\infty} -\frac{1}{2n}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon} \\ &= -\frac{1}{2\sigma^2}\underline{\epsilon}'\underline{Q}\underline{\epsilon},\end{aligned}$$

where $corr_{i,j}$ is the correlation between $b_i$ and $b_j$, and

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

The analysis is illustrated with the simple regression model.

The joint asymptotic efficiency (JAE) may be defined as the determinant of the asymptotic variance-covariance matrix, say $\underline{V}^a$, leading to the JAE of the RLSEs in the form

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a| = \lim_{n\to\infty} \sum_{i=1}^{n} n^{-1} \ln Var_n(b_{i,\epsilon_i})$$

$$= -\sum_{i=1}^{n} \frac{\epsilon_i^2 Q_{i,i}}{2\sigma^2}.$$

Considered next are not-necessarily-Normal multivarite sample means and medians, treated by bivariate large derivations similar to those introduced for the regression analysis. For $j = 1, 2, \ldots, p$, let $\{X_{ij}\}$ be a random sample from a certain $p$-variate distribution $F$. Assume that $E(X_{ij}) = \mu_j$. Each $\overline{X}_j$ and $M_j$ is rounded to the nearest point of the uniform grid $2\epsilon_j$, denoted by $\overline{X}_{j,\epsilon_j}$ and $M_{j,\epsilon_j}$, respectively. Then the asymptotic variances and asymptotic covariances of $\underline{\overline{X}}_{\epsilon} = (\overline{X}_{1,\epsilon_1}, \overline{X}_{2,\epsilon_2}, \ldots, \overline{X}_{p,\epsilon_p})$ and $\underline{M}_{\epsilon} = (M_{1,\epsilon_1}, M_{2,\epsilon_2}, \ldots, M_{p,\epsilon_p})$ are considered. In order to examine the asymptotic covariances, the bivariate distributions are assumed to be such that both $\Pr(X_1 - \mu_1 \geq \epsilon_1,\ X_2 - \mu_2 \geq \epsilon_2) \equiv P_{\epsilon_1,\epsilon_2}$ and $\Pr(X_1 - \mu_1 \geq \epsilon_1,\ X_2 - \mu_2 \leq \epsilon_2) \equiv Q_{\epsilon_1,\epsilon_2}$ are appropriately quadrant-symmetric, and bivariate log-concave in positive $(\epsilon_1, \epsilon_2)$, a property that we can call the "Two-sided Extended Bivariate Increasing Failure Rate (TEBIFR)".

Bahadur and Rao's treatment of the univariate behaviors of large deviations can be extended to the multivariate case for sample means and to the bivariate case for sample medians:

$$n^{-1} \ln \Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}) = \ln \rho_{\underline{\bar{X}}}(\underline{\epsilon}) + o_n(1),$$

where $\rho_{\underline{\bar{X}}}(\underline{\epsilon})$ is defined as $\min_{\underline{t}}[\exp(-\sum_{j=1}^{p} \epsilon_j t_j)\phi_{\underline{X}-\underline{\mu}}(\underline{t})]$, and

$$n^{-1} \ln \Pr(M_1 - \mu_1 \geq \epsilon_1, \;\; M_2 - \mu_2 \geq \epsilon_2) = \ln \rho_{\underline{M}}(\epsilon_1, \epsilon_2) + o_n(1),$$

where $\rho_{\underline{M}}(\epsilon_1, \epsilon_2)$ is defined as $\min_{(t_1,t_2)} [\exp(-\frac{t_1}{2} - \frac{t_2}{2})\phi_{Y_{\epsilon_1}, Y_{\epsilon_2}}(t_1, t_2)]$.

The asymptotic variances of the rounded sample mean and the rounded sample median have been already considered. The preceding relations make possible the computations of the asymptotic covariances of the estimators:

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Cov_n(\overline{X}_{i,\epsilon_i}, \; \overline{X}_{j,\epsilon_j}) = \ln \rho_{\underline{\bar{X}}}(\epsilon_i, \epsilon_j),$$

and

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Cov_n(M_{i,\epsilon_i}, \; M_{j,\epsilon_j}) = \ln \rho_{\underline{M}}(\epsilon_i, \epsilon_j).$$

We thus obtain the matrices of the asymptotic variances and covariances of rounded sample means and rounded sample medians, leading to the possibility of a comparison of $\underline{\overline{X}}_{\underline{\epsilon}}$ and $\underline{M}_{\underline{\epsilon}}$ in terms of the joint asymptotic efficiencies (JAEs). It is found that the JAEs of $\underline{\overline{X}}_{\underline{\epsilon}}$ and $\underline{M}_{\underline{\epsilon}}$ are the summations of univariate large-deviation rates:

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{\bar{X}})| = \sum_{j=1}^{p} \rho_{\underline{\bar{X}}}(\epsilon_j)$$

and

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{M})| = \sum_{j=1}^{p} \rho_{\underline{M}}(\epsilon_j).$$

These computations are illustrated using the Normal and Laplace distributions.

## 1.3. Review of Large Deviations

Before discussing our research it will be beneficial to recall certain standard results concerning large deviations.

Billingsley [14] described large deviations to estimate $\Pr(Y \geq \alpha)$, where $Y$ is a discrete random variable assuming values $y_j$ with probabilities $p_j$, and $\alpha$ is positive.

Assume that

$$E(Y) < 0, \tag{1.1}$$

$$\Pr(Y > 0) > 0. \tag{1.2}$$

Let $\phi(t) = \sum_j p_j e^{ty_j}$ be the moment generating function of $Y$. Then $\phi'(0) < 0$ by (1.1), and $\phi(t) \to \infty$ as $t \to \infty$ by (1.2). Since $\phi(t)$ is convex, it has its minimum $\rho$ at a positive argument $\tau$:

$$\inf_t \phi(t) = \phi(\tau) = \rho,$$

where $0 < \rho < 1$, and $\tau > 0$.

For all positive $t$, $\Pr(Y \geq 0) = \Pr(e^{tY} \geq 1) \leq \phi(t)$ by Markov's inequality, and hence

$$\Pr(Y \geq 0) \leq \rho.$$

Billingsley [14] also provided a lower bound for $\Pr(Y \geq 0)$.

Those bounds motivate our understanding to the asymptotic behavior of $\Pr(Y \geq 0)$. In this connection, Billingsley [14] gave the following theorem by Chernoff:

Chernoff's Theorem

*Let $X_1, X_2, \ldots$ be independent, identically distributed simple random variables satisfying $E(X_n) < 0$ and $\Pr(X_n > 0) > 0$, let $\phi(t)$ be their common moment generating function, and put $\rho = \inf_t \phi(t)$. Then*

$$\lim_{n \to \infty} n^{-1} \ln \Pr(X_1 + \cdots + X_n \geq 0) = \ln \rho. \qquad \bullet$$

Bahadur and Rao [4] extended Chernoff's Theorem, which dealt with simple random variables, to continuous random variables: Let $a$ be a constant, $-\infty < a < \infty$, and for each $n = 1, 2, \ldots$ let

$$p_n = \Pr(\frac{X_1 + \cdots + X_n}{n} \geq a).$$

The distribution of $X_1$ is assumed to be such that $p_n > 0$ for each $n$, and that $p_n \to 0$ as $n \to \infty$. The objective of their paper was to obtain an estimate of $p_n$, say $q_n$, which is precise in the sense that, as $n \to \infty$,

$$\frac{q_n}{p_n} = 1 + o(1).$$

The essentials of their approach are as follows:

Let $t$ be a real variable, and let $\phi(t)$ denote the moment generating function of $X_1$, i.e., $\phi(t) = E(e^{tX_1})$. Define

$$\psi(t) = e^{-at}\phi(t).$$

Let $T$ denote the set of all values $t$ for which $\phi(t) < \infty$. Suppose that $\Pr(X_1 = a) \neq 1$, that $T$ is a non-degenerate interval, and that there exists a positive $\tau$ in the interior of $T$ such that $\psi(\tau) = \inf_t \psi(t) = \rho$.

Let $Y_1 = X_1 - a$, and let $F$ be the (left-continuous) distribution function of $Y_1$, $F(y) = \Pr(Y_1 < y)$. Let $G$ be defined as $G(z) = \int_{-\infty<y<z} \rho^{-1} e^{\tau y} dF(y)$ (the so-called exponential centering of $F$). Since $E(e^{\tau Y_1}) = \psi(\tau) = \rho$, $G$ is clearly a probability distribution function. Let $Z_1$ be a random variable distributed according to $G$. Bahadur and Rao [4] show that the moment generating function of $Z_1$ exists in a neighborhood of the origin, and that

$$E(Z_1) = 0, \qquad 0 < Var(Z_1) < \infty.$$

Further, denoting $Var(Z_1)$ by $\sigma^2$, they obtained

$$\sigma^2 = \frac{\phi''(\tau)}{\phi(\tau)} - a^2.$$

Again, with $\alpha$ defined by

$$\alpha = \sigma\tau, \qquad (0 < \alpha < \infty),$$

and $Z_1, Z_2, \ldots$ an *iid* sequence distributed according to $G$, let

$$U_n = \frac{Z_1 + \cdots + Z_n}{\sqrt{n}\sigma}$$

and

$$H_n(x) = \Pr(U_n < x), \qquad (-\infty < x < \infty).$$

Then Bahadur and Rao [4] gave the lemma:

Lemma 2 of Bahadur and Rao [4]

$$p_n = \rho^n I_n,$$

*where*

$$I_n = \sqrt{n}\alpha \int_0^\infty e^{-\sqrt{n}\alpha x}[H_n(x) - H_n(0)]dx.$$

This lemma leads to the conclusion of Chernoff's Theorem for continuously distributed $X_1$ as follows:

Since $0 \leq H_n(x) - H_n(0) \leq 1$ for every $n$ and $x \geq 0$, we have $I_n \leq 1$ and hence $p_n \leq \rho^n$ for every $n$, by Lemma 2 above. Let $\epsilon$ be a positive constant. Then

$$\begin{aligned} I_n &\geq \sqrt{n}\alpha \int_\epsilon^\infty \exp(-\sqrt{n}\alpha x)[H_n(x) - H_n(0)]dx \\ &\geq [H_n(\epsilon) - H_n(0)]\sqrt{n}\alpha \int_\epsilon^\infty \exp(-\sqrt{n}\alpha x)dx \\ &= [H_n(\epsilon) - H_n(0)]\exp(-\sqrt{n}\alpha\epsilon). \end{aligned}$$

Hence $\liminf_{n\to\infty}\{\frac{1}{\sqrt{n}}\ln I_n\} \geq -\alpha\epsilon$. Since $I_n \leq 1$ for every $n$, and since $\epsilon$ is arbitrary, it follows that $\frac{1}{\sqrt{n}}\ln I_n = o(1)$. Hence

$$n^{-1}\ln p_n = \ln\rho + o(1).$$

Bahadur provided the following additional background material: for each $n = 1, 2, \ldots$, let $T_n = T_n(x_1, x_2, \ldots, x_n)$ be a statistic such that the sequence $\{T_n\}$ is a consistent estimate of real location parameter $\theta$. Bahadur [2] defined the asymptotic effective standard deviation of $T_n$ as the solution $\tau$ of the equation $\Pr(|T_n - \theta| \geq \epsilon|\theta) = \Pr(|N| \geq \epsilon/\tau)$ when $n$ is large and $\epsilon$ is a small positive number, where $N$ denotes a standard Normal variable.

Let $S_n = \sum_{i=1}^{n} X_i$. Under the assumptions in Bahadur and Rao [4], the asymptotic effective variance of the sample mean of $n$ independent identically distributed random variables exists and equals $n^{-1}$ times the variance of each random variable $X_i$, as usually defined. A result of Bahadur [2] (with $N$ denoting a standard Normal random variable) is provided by

<u>Lemma 2.4 of Bahadur [2]</u>

*If* $\lambda_n(\epsilon)$ *is defined by*

$$\Pr(|S_n/n| \geq \epsilon) = \Pr(|N| \geq \epsilon/\lambda_n(\epsilon))$$

*for* $\epsilon > 0$ *and* $n = 1, 2, \ldots$, *then*

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \{n\lambda_n^2(\epsilon)\} = E(X^2).$$

*Equivalently,*

$$\Pr(|S_n/n| \geq \epsilon) = \exp(-\frac{n\epsilon^2}{2E(X^2)}[1 + \delta_n(\epsilon)]),$$

*where*

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \delta_n(\epsilon) = 0. \qquad \bullet$$

The preceding results provide the basis for this dissertation.

## 1.4. Overview

Under a certain class of distributions called TEIFR, Chapter 2 provides the large-deviation derivation of the asymptotic relative efficiency $e_\epsilon$ of the rounded sample median $M_\epsilon$ vs. the rounded sample mean $\overline{X}_\epsilon$, in a sense related to the concept (Hodges and Lehmann [27], Bahadur [2,3]) of limiting ratio of equivalent sample sizes.

Chapter 3 examines the behavior of $e_\epsilon$ for $\epsilon$ near zero and compares $e_\epsilon$ with the asymptotic relative efficiency used by Hammersley [26]. It also expands the analysis of Chapter 2 to $\epsilon$'s of arbitrary size, for the special Normal and Laplace cases, with interesting conclusions regarding grid size, and compares the asymptotic error rates of certain related tests based on $\overline{X}$ and $M$.

Chapter 4 considers the asymptotic variances and covariances of the rounded least-squares estimators (RLSEs) $b_{j,\epsilon_j}$ in the Normal $p$-variate regression model under certain assumptions on the explanatory vector $\underline{X}$. The discussion is illustrated by the simple regression model, and the joint asymptotic efficiency (JAE) of the RLSEs is described.

Chapter 5 extends the ideas of granularity and rounded asymptotic efficiency of sample means and medians to the multivariate case.

Chapter 6 considers alternative forms of rounding.

# 2. UNIVARIATE GRANULARITY AND RELATIVE EFFICIENCY

## 2.1. Introduction

Consider the rounded sample mean $\overline{X}_\epsilon$ and the rounded sample median $M_\epsilon$ as estimates of the grid-valued location parameter $\mu$ under a certain class of TEIFR distributions. The asymptotic mean-square errors (MSEs) of the rounded sample median $M_\epsilon$ and the rounded sample mean $\overline{X}_\epsilon$ are formulated via the theory of large-deviation (Bahadur and Rao [4]; Bahadur [2,3]).

.Suppose a random sample $\{X_i\ ;\ \ i = 1, 2, \ldots, n\}$ is drawn from a TEIFR distribution with location parameter $\mu$ and variance $\sigma^2$. Then the asymptotic mean-square error $MSE_n(M_\epsilon)$ of the sample median, rounded to the grid $(2k\epsilon\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\lim_{n \to \infty} n^{-1} \ln MSE_n(M_\epsilon) = \ln[2(R_\epsilon(1 - R_\epsilon))^{1/2}],$$

where

$$R_\epsilon = \max\{\Pr(X_i - \mu \geq \epsilon),\ \ \Pr(X_i - \mu \leq -\epsilon)\}.$$

And the asymptotic mean-square error $MSE_m(\overline{X}_\epsilon)$ of the sample mean, rounded to the grid $(2k\epsilon\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\lim_{m \to \infty} m^{-1} \ln MSE_m(\overline{X}_\epsilon) = -\frac{\epsilon^2}{2\sigma^2}(1 + \delta(\epsilon)),$$

when $\epsilon$ is sufficiently small, where $m$ denotes sample sizes, and the function $\delta(\cdot)$ satisfies

$$\lim_{\epsilon \to 0} \delta(\epsilon) = 0.$$

In terms of limiting variances or mean-square errors, the asymptotic relative efficiency $e_\epsilon$ of $M_\epsilon$ with respect to $\overline{X}_\epsilon$, when $\epsilon$ is small, is

$$e_\epsilon = -2\sigma^2 \epsilon^{-2} (1 + \delta(\epsilon))^{-1} \ln[2(R_\epsilon(1 - R_\epsilon))^{1/2}].$$

The asymptotic relative efficiency $e_\epsilon$ is computed as the limiting ratio of "equivalent sample sizes" $n$ and $m$; this computation, if it is to be based on variances rather than mean-square errors, calls for the above additional assumption of symmetry of our TEIFR distribution, under which $P_\epsilon = Q_\epsilon = R_\epsilon$.

## 2.2. Assumptions

The underlying distribution is assumed to be such that $E(X) = Med(X) = \mu$, and that both $\Pr(X - \mu \geq c)$ and $\Pr(X - \mu \leq -c)$ are log-concave in $c$ for $c \geq 0$, a property that we call "Two-sided Extended Increasing Failure Rate (TEIFR)" ("Two-sided" because both sides of zero are being considered and "extended" because, while the log-concavity does guarantee continuity, it does not guarantee absolute continuity).

Define $P_c \equiv \Pr(X - \mu \geq c)$, $Q_c \equiv \Pr(X - \mu \leq -c)$ and let $\gamma \equiv \sup\{c : P_c > 0\}$ and $\gamma' \equiv \sup\{c : Q_c > 0\}$.

Condition C1:

(1) $P_c$ is log-concave for $c$ on $(0, \gamma)$ and $Q_c$ is log-concave for c on $(0, \gamma')$.

(2) $\epsilon$ is such that $R_\epsilon \equiv \max(P_\epsilon, Q_\epsilon) > 0$.

(3a) If $\epsilon$ is such that $P_\epsilon > 0$, then $P_\epsilon > P_{3\epsilon}$.

(3b) If $\epsilon$ is such that $Q_\epsilon > 0$, then $Q_\epsilon > Q_{3\epsilon}$.

We refer to Condition C1(1), as specifying that the distribution of $X_i - \mu$ be TEIFR, because both tails of the distribution of $X_i - \mu$ are being considered and, while Condition C1(1) guarantees continuity of $P_c$ on $(0, \gamma)$ and $Q_c$ on $(0, \gamma')$, respectively, absolute continuity is not guaranteed ( i.e., $X_i - \mu$ does not necessarily have a density). Condition C1(2) insures that at least one of two tails tails of the distributions of $X_i - \mu$ protrudes beyond $\epsilon$ (respectively, -$\epsilon$) and guarantees a finite and non-zero variance $\sigma^2$. Later on, this Condition C1(2) is a requirement for being able to describe the behavior of $\overline{X}_\epsilon$ in terms of large deviations.

Finally, the rationale for requiring Condition C1(3) is as follows: Suppose that $P_\epsilon > 0$. Then since the median of $(X_i - \mu)$ is zero so that $P_\epsilon \leq 1/2$, Condition C1(3a) insures that $P_\epsilon(1 - P_\epsilon) > P_{3\epsilon}(1 - P_{3\epsilon})$, a condition required in the large-deviation treatment of $M_\epsilon$, as explained following equation (2.12).

Condition C1 is of course meant to be added to the earlier-stated blanket condition that the population mean and median coincide.

## 2.3. Asymptotic MSEs of the Rounded Sample Median $M_\epsilon$

For the IFR-type cases to be treated, both $M_\epsilon$ and $\overline{X}_\epsilon$ will equal $\mu$ with probability tending to 1, so that neither will possess a non-degenerate asymptotic distribution. We have therefore little choice but to follow Hammersley [26] in computing the asymptotic relative efficiencies of one to the other, in terms of asymptotic variances or mean-square errors, rather than in terms of variances, or of effective variances (Bahadur [2]), of their limiting distributions.

We begin with a tail computation for $M_\epsilon$, supposing first that $n$ is an odd number: $n = 2v + 1$.

For any positive $c$, in particular any positive integer multiple of $\epsilon$, let $Y_{c,i}$ be the indicator of the event $X_i - \mu \geq c$. Then

$$\Pr(M - \mu \geq c) = \Pr(\sum_{i=1}^{2v+1} Y_{c,i} \geq v + 1) \tag{2.1}$$

$$= \Pr(\sum_{i=1}^{2v+1} Y_{c,i} \geq v + 1/2)$$

$$= \Pr(\overline{Y}_c \geq 1/2),$$

where the penultimate equality is due to the fact that $\sum Y_{c,i}$ is an integer. Performing now the standard first-order large deviation computation (Bahadur and Rao [4]) for $\Pr(\overline{Y}_c \geq 1/2)$, based on the moment generating function $\phi(t) = 1 - P_c + P_c \exp(t)$ of $Y_{c,i}$, where $P_c = \Pr(Y_{c,i} = 1) = \Pr(X_i - \mu \geq c)$, we find that $\exp(-t/2)\phi(t)$ is minimized for $t = \ln((1-P_c)/P_c) \equiv \tau$, and that $\exp(-\tau/2)\phi(\tau) = 2[P_c(1-P_c)]^{1/2}$; hence

$$\Pr(M - \mu \geq c) = \Pr(\overline{Y}_c \geq 1/2) \leq 2^n [P_c(1 - P_c)]^{n/2} \tag{2.2}$$

for all n, and, in addition, for any $\delta > 0$,

$$\Pr(M - \mu \geq c) = \Pr(\overline{Y}_c \geq 1/2) \geq (1 - \delta)^n (2^n [P_c(1 - P_c)]^{n/2}) \tag{2.3}$$

for $n$ large enough.

Probabilities of the form $\Pr(M - \mu \leq -c)$ are dealt with in a similar way. Define indicators $Z_{c,i}$ of the events $X_i - \mu \leq -c$, and find analogously to (2.1), that

$$\Pr(M - \mu \leq -c) = \Pr(\overline{Z}_c \geq 1/2),$$

and, with $Q_c$ set equal to $\Pr(X_i - \mu \leq -c)$, obtain the analogues of (2.2) and (2.3):

$$\Pr(M - \mu \leq -c) = \Pr(\overline{Z}_c \geq 1/2) \leq 2^n [Q_c(1 - Q_c)]^{n/2} \tag{2.4}$$

for all n, and, in addition, for any $\delta > 0$,

$$\Pr(M - \mu \leq -c) = \Pr(\overline{Z}_c \geq 1/2) \geq (1 - \delta)^n (2^n [Q_c(1 - Q_c)]^{n/2}) \tag{2.5}$$

for $n$ large enough.

We are now ready to approximate the large-sample MSE of $M_\epsilon$ when $n$ is odd; to that end define:

$$\begin{aligned}\pi_{k,\epsilon}(n), k: 1,2,\ldots, &= k^2[\Pr(M-\mu \geq (2k-1)\epsilon) \\ &\quad - \Pr(M-\mu \geq (2k+1)\epsilon)]\end{aligned}$$

and

$$\begin{aligned}\pi'_{k,\epsilon}(n), k: 1,2,\ldots, &= k^2[\Pr(M-\mu \leq -(2k-1)\epsilon) \\ &\quad - \Pr(M-\mu \leq -(2k+1)\epsilon)].\end{aligned}$$

Then

$$(2\epsilon)^{-2}MSE_n(M_\epsilon) = (2\epsilon)^{-2}E[(M_\epsilon-\mu)^2] \tag{2.6}$$

$$= \{\Pr(M-\mu \geq \epsilon) \tag{2.7}$$

$$- \Pr(M-\mu \geq 3\epsilon) \tag{2.8}$$

$$+ \sum_{k=2}^{\infty} \pi_{k,\epsilon}(n)\} \tag{2.9}$$

$$+\{\Pr(M-\mu \leq -\epsilon) \tag{2.10}$$

$$- \Pr(M-\mu \leq -3\epsilon) \tag{2.11}$$

$$+ \sum_{k=2}^{\infty} \pi'_{k,\epsilon}(n)\}. \tag{2.12}$$

As mentioned in Section 2.2, Condition C1(3a) insures that

$$P_\epsilon(1 - P_\epsilon) > P_{3\epsilon}(1 - P_{3\epsilon})$$

which implies via (2.2) and (2.3) that (2.7) dominates (2.8).

In addition,

**Lemma 2.1** $\Pr(M - \mu \geq \epsilon)$ *dominates* $\sum_{k=2}^{\infty} \pi_{k,\epsilon}(n)$ *under Condition C1 when* $P_\epsilon > 0$.

Proof:

Assuming without loss of generality that

$$P_\epsilon > 0,$$

Condition C1(3a) implies that

$$P_\epsilon > P_{3\epsilon}. \tag{2.13}$$

Also without loss of generality, and to avoid the minor modifications required otherwise, assume that $\gamma$ in Condition C1(1) equals $+\infty$, so that

$$P_c > 0 \tag{2.14}$$

for $c \geq 0$.

The demonstration that $\Pr(M - \mu \geq \epsilon)$ dominates $\sum_{k=2}^{\infty} \pi_{k,\epsilon}(n)$ derives from the fact that it is only the lower bound (2.3) holds for large enough n, rather than the upper bound (2.2), since the lower bound is used only once, in bounding

$\Pr(M - \mu \geq \epsilon)$ from below, whereas the upper bound has to be applied to each of the terms of the infinite series $\sum_{k=2}^{\infty} \pi_{k,\epsilon}(n)$.

At any rate, recalling (2.3), (2.13) and the fact that $P_\epsilon \leq 1/2$ because $\mu$ is the median of the distribution of $X_i$, choose a $\delta$ such that

$$(P_\epsilon(1 - P_\epsilon))^{1/2}(1 - \delta) > (P_{3\epsilon}(1 - P_{3\epsilon}))^{1/2} \tag{2.15}$$

and

$$\Pr(M - \mu \geq \epsilon) \geq (1 - \delta)^n (2^n [P_\epsilon(1 - P_\epsilon)]^{n/2}) \tag{2.16}$$

for large enough $n$.

On the other hand, in view of (2.2), we have for all $n$ that

$$\begin{aligned} 0 &\leq \sum_{k=2}^{\infty} \pi_{k,\epsilon}(n) \\ &\leq \sum_{k=2}^{\infty} k^2 \Pr(M - \mu \geq (2k - 1)\epsilon) \\ &\leq 2^n \sum_{k=2}^{\infty} k^2 [P_{(2k-1)\epsilon}(1 - P_{(2k-1)\epsilon})]^{n/2}. \end{aligned} \tag{2.17}$$

Recall now that $P_\epsilon \leq 1/2$ so that $P_{3\epsilon} < 1/2$ in view of (2.13). Since $P_{3\epsilon} < 1/2$ there is a real number $a$ that is large enough to insure that $\lambda_a(p) \equiv (p(1 - p))^a / p$ is increasing on $(0,\ P_{3\epsilon}]$, so that, for $P_1 \leq P_2$ in $(0,\ P_{3\epsilon}]$,

$$\frac{\lambda_a(P_1)}{\lambda_a(P_2)} = [\frac{(P_1(1 - P_1))^a}{(P_2(1 - P_2))^a}][\frac{P_2}{P_1}] \leq 1,$$

or

$$\frac{(P_1(1 - P_1))^a}{(P_2(1 - P_2))^a} \leq \frac{P_1}{P_2}. \tag{2.18}$$

Therefore, examining the ratio of successive terms of (2.17), we find, for $k \geq 2$, that

$$\frac{(k+1)^2[P_{(2k+1)\epsilon}(1-P_{(2k+1)\epsilon})]^{n/2}}{k^2[P_{(2k-1)\epsilon}(1-P_{(2k-1)\epsilon})]^{n/2}}$$

$$= [\frac{k+1}{k}]^2[\frac{(P_{(2k+1)\epsilon}(1-P_{(2k+1)\epsilon}))^a}{(P_{(2k-1)\epsilon}(1-P_{(2k-1)\epsilon}))^a}]^{n/(2a)}$$

$$\leq [\frac{k+1}{k}]^2[\frac{P_{(2k+1)\epsilon}}{P_{(2k-1)\epsilon}}]^{n/(2a)}$$

$$\leq [\frac{k+1}{k}]^2[\frac{P_{3\epsilon}}{P_\epsilon}]^{n/(2a)}$$

$$\leq (\frac{3}{2})^2[\frac{P_{3\epsilon}}{P_\epsilon}]^{n/(2a)}, \qquad (2.19)$$

where the first inequality is due to (2.18) and the second inequality is due to the log-concavity of $P_c$ required by Condition C1(1). Now choose any $\rho < 1$ and let $n_o$ be such that (2.16) holds for $n \geq n_o$, and (2.19) is no greater than $\rho$. Then the RHS of (2.17) is no greater than

$$\frac{2^{n+2}[P_{3\epsilon}(1-P_{3\epsilon})]^{n/2}}{(1-\rho)}$$

for $n \geq n_o$, which, recalling (2.15) and (2.16), does indeed verify that, as $n$ increases,

$$\frac{\Pr(M-\mu \geq \epsilon)}{\sum_{k=2}^{\infty} \pi_{k,\epsilon}(n)} \longrightarrow 0.$$

This completes the proof of Lemma 2.1, for $n$ odd.

Next consider the case of even $n$, say $n = 2v$, and illustrate the argument with the first three terms of (2.6), the second three terms being handled in precisely the same way. With $Y_{c,i}$ defined as in (2.1), the analogue of (2.1) is

$$\begin{aligned}
\Pr(\overline{Y}_c \geq \frac{1}{2} + \frac{1}{n}) &= \Pr(\sum_{i=1}^{2v} Y_{c,i} \geq v + 1) \\
&\leq \Pr(M - \mu \geq c) \\
&\leq \Pr(\sum_{i=1}^{2v} Y_{c,i} \geq v) \qquad (2.20) \\
&= \Pr(\overline{Y}_c \geq \frac{1}{2}).
\end{aligned}$$

But, according to Lemma 2.2 of Bahadur [2], $n^{-1} \ln \Pr(\overline{Y}_c \geq 1/2 + o_n(1))$ and $n^{-1} \ln \Pr(\overline{Y}_c \geq 1/2)$ have the same limit, so that individual terms of form $\Pr(M - \mu \geq k\epsilon)$ may be handled precisely as when $n$ is odd. In addition, the infinite series treated by Lemma 2.1 is covered as well, since, fortunately, the inequality (2.20), which furnishes the upper bound to terms of form $\Pr(M - \mu \geq k\epsilon)$, is precisely the expression underlying the argument in Lemma 2.1.

**Theorem 2.2** *Suppose a random sample* $\{X_i \; ; \quad i = 1, 2, \ldots, n\}$ *is drawn from the TEIFR class distribution whose location parameter is* $\mu$. *Then the asymptotic mean-square error* $MSE_n(M_\epsilon)$ *of the sample median rounded to the grid* $(2k\epsilon \; ; \;\; k =$

$0, \pm 1, \pm 2, \ldots)$ *satisfies*

$$\lim_{n \to \infty} n^{-1} \ln MSE_n(M_\epsilon) = \ln[2(R_\epsilon(1 - R_\epsilon))^{1/2}] \equiv -\theta_M(\epsilon), \qquad (2.21)$$

*where*

$$R_\epsilon \equiv \max\{ \ P_\epsilon = \Pr(X_i - \mu \geq \epsilon), \quad Q_\epsilon = \Pr(X_i - \mu \leq -\epsilon) \ \}.$$

Proof:

A development analogous to the proof of Lemma 2.1, pertaining to the lower tail, this time featuring Condition C1(3b), leads to the conclusion that, if $Q_\epsilon > 0$, then (2.10) dominates both (2.11) and (2.12).

Hence, combining the upper tail and lower tail arguments,

$$\begin{aligned} (2\epsilon)^{-2} MSE_n(M_\epsilon) \quad &= \quad \Pr(M - \mu \geq \epsilon)(1 + o_n(1)) \\ &\quad + \Pr(M - \mu \leq -\epsilon)(1 + o_n(1)), \end{aligned} \qquad (2.22)$$

where one of the two addends of the RHS may equal zero. This last relation leads to the result (2.21) of the Theorem 2.2, using (2.2), (2.3), (2.4) and (2.5). This completes the proof of Theorem 2.2.

## 2.4. Asymptotic MSEs of the Rounded Sample Mean $\overline{X}_\epsilon$

We turn to the rounded sample mean $\overline{X}_\epsilon$, and use the symbol $m$, rather than $n$, to denote sample size.

We begin our discussion of $\overline{X}_\epsilon$ with the case of $\epsilon$ small, partly because, when $\epsilon$ is small, Condition C1, which served us in the case of $M_\epsilon$, is in addition adequate to validate our analysis of $\overline{X}_\epsilon$. We rely on Lemma 2.4 of Bahadur [2], formulated for the case of $\epsilon$ small, which happens to be couched directly in terms of the large-deviation behavior of $|\overline{X}-\mu|$ ( Bahadur in effect studies both tails of the distribution of $\overline{X} - \mu$ and then combines them, much as we did in going from (2.22) to (2.21)). Hence we write the analogue of (2.6) as the sum of three (rather than six) terms:

$$(2\epsilon)^{-2}MSE_m(\overline{X}_\epsilon) = (2\epsilon)^{-2}E[(\overline{X}_\epsilon - \mu)^2] \tag{2.23}$$

$$= \{\Pr(|\overline{X} - \mu| \geq \epsilon) \tag{2.24}$$

$$- \Pr(|\overline{X} - \mu| \geq 3\epsilon) \tag{2.25}$$

$$+ \sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m)\}, \tag{2.26}$$

where

$$\pi^*_{k,\epsilon}(m) = k^2[\Pr(|\overline{X} - \mu| \geq (2k-1)\epsilon) - \Pr(|\overline{X} - \mu| \geq (2k+1)\epsilon)].$$

**Lemma 2.3** *If the distribution $F$ of $X_i - \mu$ is TEIFR, then $F$ possesses a moment generating function.*

Proof:

Since F is TEIFR, there exist constants $a$, $b$, $c$ and $d$, with $b$, $d > 0$, such that

$$\ln(1 - F(y)) \leq a - by \tag{2.27}$$

$$\ln F(x) \leq c - dx, \tag{2.28}$$

for $y \in [0, \ \gamma)$ and $x \in [0, \ \gamma')$.

Now first consider $\int_0^\gamma \exp(tx)dF(x)$. That integral certainly is finite for $t \leq 0$. Furthermore, regarding positive $t$, write

$$\begin{aligned}
\int_0^\gamma \exp(tx)dF(x) &= \int_0^\gamma [1 + \int_0^x t\exp(ty)dy]dF(x) \\
&= F(\gamma) - F(0) + t\int_0^\gamma [\int_y^\gamma dF(x)]\exp(ty)dy \\
&\leq 1 - F(0) + t\int_0^\gamma [1 - F(y)]\exp(ty)dy \\
&\leq 1 - F(0) + t\exp(a)\int_0^\infty \exp[(t-b)y]dy \\
&< \infty
\end{aligned}$$

for $0 < t < b$, where the weak inequality is due to (2.27). Thus, all told,

$$\int_0^\gamma \exp(tx)dF(x) < \infty \tag{2.29}$$

for $t < b$.

Next consider $\int_0^{\gamma'} \exp(-tu)dG(u)$, where $G(u) = 1 - F(-u)$. That integral certainly is finite for $t \geq 0$; regarding negative $t$, write

$$\int_0^{\gamma'} \exp(-tu)dG(u) = \int_0^{\gamma'} [1 - \int_0^u t\exp(-ty)dy]dG(u)$$

$$
\begin{aligned}
&= G(\gamma') - G(0) - t\int_0^{\gamma'} [\int_y^{\gamma'} dG(u)] \exp(-ty)dy \\
&\leq 1 - G(0) - t\int_0^{\gamma'} [1 - G(y)] \exp(-ty)dy \\
&= F(0) - t\int_0^{\gamma'} F(-y) \exp(-ty)dy \\
&\leq F(0) - t\exp(c)\int_0^{\gamma'} \exp[-(d+t)y]dy \\
&< \infty
\end{aligned}
$$

for $t > -d$, where the weak inequality is due to (2.28). Thus, all told,

$$\int_{-\gamma'}^{0} \exp(tx)dF(x) < \infty \tag{2.30}$$

for $t > -d$.

Finally, combining (2.29) and (2.30), we find that

$$\int_{\gamma'}^{\gamma} \exp(tx)dF(x) < \infty$$

for $-d < t < b$, which establishes that $F$ possesses a moment generating function. This completes the proof.

Bahadur's conditions for his Lemma 2.4 are that ($\alpha$) the distribution of $X_i - \mu$ possess a moment generating function, and ($\beta$) the variance $\sigma^2$ of $X_i - \mu$ (finite in view of ($\alpha$)) exceeds zero. Lemma 2.3 shown above tells us ($\alpha$) is guaranteed by Condition C1(1) and ($\beta$) clearly is guaranteed by Condition C1(2). Hence with regard to (2.24), we may avail ourselves of Bahadur's Lemma 2.4 under Condition C1 (especially his relation (2.11)) when $\epsilon$ is sufficiently small, and write

$$\Pr(|\overline{X} - \mu| \geq \epsilon) = \exp[-\frac{m\epsilon^2}{2\sigma^2}(1 + \delta_m(\epsilon))],$$

where

$$\lim_{m \to \infty} \delta_m(\epsilon) = \delta(\epsilon)$$

with the function $\delta(\cdot)$ satisfying

$$\lim_{\eta \to 0} \delta(\eta) = 0. \tag{2.31}$$

**Lemma 2.4** $\Pr(|\overline{X} - \mu| \geq \epsilon)$ *dominates both* $\Pr(|\overline{X} - \mu| \geq 3\epsilon)$ *and* $\sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m)$.

Proof:

We first compute two lower bounds ((2.32) and (2.33) below) for $\Pr(|\overline{X} - \mu| \geq \epsilon)$; these bounds, and the insuring derivation, are valid, when $\epsilon$ is small enough, for any distribution satisfying Condition C1, and are valid as well for arbitrary $\epsilon$ in the case of the Normal and Laplace distributions. The bounds are computed by successively computing $t_{1,\epsilon} > 0$, $t_{2,\epsilon} < 0$, $a_\epsilon$, $b_\epsilon$ and $c_\epsilon$ as in Bahadur [2], the latter three being given by

$$a_\epsilon = \exp(-\epsilon t_{1,\epsilon})\phi(t_{1,\epsilon}) = \inf_{t>0} \exp(-\epsilon t)\phi(t),$$

$$b_\epsilon = \exp(\epsilon t_{2,\epsilon})\phi(t_{2,\epsilon}) = \inf_{t<0} \exp(\epsilon t)\phi(t)$$

and

$$c_\epsilon = \max(a_\epsilon, b_\epsilon).$$

Substituting now Bahadur's relation (2.9) for his relation (2.13) (or using Theorem 3.1 in Bahadur [3]), we then obtain

$$\lim_{m\to\infty} m^{-1} \ln \Pr(|\overline{X} - \mu| \geq \epsilon) = \ln c_\epsilon,$$

from which we conclude that

$$\Pr(|\overline{X} - \mu| \geq \epsilon) \geq \exp(-m\epsilon t_{1,\epsilon})\phi^m(t_{1,\epsilon}) \exp(1 + o_m(1)) \tag{2.32}$$

and

$$\Pr(|\overline{X} - \mu| \geq \epsilon) \geq \exp(m\epsilon t_{2,\epsilon})\phi^m(t_{2,\epsilon}) \exp(1 + o_m(1)). \tag{2.33}$$

We are now ready to compare $\sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m)$ to $\Pr(|\overline{X} - \mu| \geq \epsilon)$. To this end, let $I_K(z)$ be the indicator function (in z) of the event $z \geq K$. Pursuing the familiar first steps of the large deviation argument (viz., the proof of Theorem 3.1 in Bahadur [3]), we start with the fact that

$$\exp(z - K)t \geq I_k(z)$$

for any $t \geq 0$, and, upon taking expectations with respect to the distribution of $\overline{X} - \mu$, find for any multiple $h\epsilon$ of $\epsilon$, setting $K = h\epsilon$, that

$$\exp(-mh\epsilon t)\phi^m(t) \geq \Pr(\overline{X} - \mu \geq h\epsilon) \tag{2.34}$$

for any $t \geq 0$; in addition, using the fact that the moment generating function of $-(X_i - \mu)$ is $\phi(-t)$, we also find that

$$\exp(-mh\epsilon t)\phi^m(-t) \geq \Pr(\overline{X} - \mu \leq -h\epsilon) \tag{2.35}$$

for any $t \geq 0$. But

$$\begin{aligned} 0 &\leq \sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m) \\ &\leq \sum_{k=2}^{\infty} k^2 \Pr(|\overline{X} - \mu| \geq (2k-1)\epsilon) \\ &= \{\sum_{k=2}^{\infty} k^2 \Pr(\overline{X} - \mu \geq (2k-1)\epsilon) \\ &\quad + \sum_{k=2}^{\infty} k^2 \Pr(\overline{X} - \mu \leq -(2k-1)\epsilon)\}, \end{aligned} \tag{2.36}$$

so that, setting $t = t_{1,\epsilon} > 0$ in (2.34) and $t = -t_{2,\epsilon} > 0$ in (2.35),

$$\begin{aligned} 0 &\leq \sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m) \\ &\leq \{\sum_{k=2}^{\infty} k^2 \exp[-(2k-1)\epsilon m t_{1,\epsilon}]\phi^m(t_{1,\epsilon}) \\ &\quad + \sum_{k=2}^{\infty} k^2 \exp[(2k-1)\epsilon m t_{2,\epsilon}]\phi^m(t_{2,\epsilon})\}, \end{aligned}$$

and, appealing now to (2.32) and (2.33),

$$0 \leq \frac{\sum_{k=2}^{\infty} \pi^*_{k,\epsilon}(m)}{\Pr(|\overline{X} - \mu| \geq \epsilon)}$$

$$\leq \{\sum_{k=2}^{\infty} k^2 \exp[-(2k-2)\epsilon m t_{1,\epsilon}] + \sum_{k=2}^{\infty} k^2 \exp[(2k-2)\epsilon m t_{2,\epsilon}]\} \{\exp(1+o_m(1))\}. \tag{2.37}$$

It remains to show that the first square-bracketed term tends to zero with $m$. To this end set $\lambda = \min(\epsilon t_{1,\epsilon}, -\epsilon t_{2,\epsilon})$, and bound this term by

$$\begin{aligned} 2\sum_{k=2}^{\infty} k^2 \exp[-(2k-2)\lambda m] &= 2\sum_{k=1}^{\infty} (k+1)^2 \exp[-2k\lambda m] \\ &\leq 2\sum_{k=1}^{\infty} 2^{k+1} \exp[-2k\lambda m] \\ &= 4\sum_{k=1}^{\infty} \exp[k(\ln 2 - 2\lambda m)]. \end{aligned}$$

Now restrict attention to $m$ such that $\ln 2 - 2\lambda m < 0$; for such $m$, the last expression equals

$$\frac{4\exp(\ln 2 - 2\lambda m)}{[1-\exp(\ln 2 - 2\lambda m)]},$$

which clearly tends to zero with $m$.

Comparing $\Pr(|\overline{X} - \mu| \geq 3\epsilon)$ to $\Pr(|\overline{X} - \mu| \geq \epsilon)$ is of course less involved, and requires just the inequalities (2.32) and (2.33), along with the inequalities (2.34) and (2.35) with $h = 3$. So the proof of Lemma 2.4 is completed.

**Theorem 2.5** *Suppose a random sample* $\{X_i \; ; \quad i = 1, 2, \ldots, m\}$ *is drawn from* *the TEIFR class distribution whose location parameter is $\mu$. Then the asymptotic mean-square error, $MSEm(\overline{X}_\epsilon)$, of the sample mean rounded to the grid $(2k\epsilon\ ;\ k = 0, \pm 1, \pm 2, \ldots)$ satisfies*

$$\lim_{m \to \infty} m^{-1} \ln MSEm(\overline{X}_\epsilon) = -\frac{\epsilon^2}{2\sigma^2}(1 + \delta(\epsilon)) \equiv -\theta_{\overline{X}}(\epsilon) \tag{2.38}$$

*under Condition C1 when $\epsilon$ is sufficiently small.*

Proof:

Bahadur's [2] Lemma 2.4 allows the assertion, for $\epsilon$ sufficiently small, that, under Condition C1,

$$\lim_{m \to \infty} m^{-1} \ln \Pr(|\overline{X} - \mu| \geq \epsilon) = -\frac{\epsilon^2}{2\sigma^2}(1 + \delta(\epsilon)) \equiv -\theta_{\overline{X}}(\epsilon), \tag{2.39}$$

with the function $\delta(\cdot)$ satisfying (2.31).

It is shown in Lemma 2.4 above (where, again, the argument involves combining tails) that, in addition, $\Pr(\overline{X} - \mu| \geq \epsilon)$ dominates both (2.25) and (2.26) under Condition C1 when $\epsilon$ is sufficiently small, so that, in view of (2.23), relation (2.39) in fact allows concluding that

$$\lim_{m \to \infty} m^{-1} \ln MSEm(\overline{X}_\epsilon) = -\theta_{\overline{X}}(\epsilon) \tag{2.40}$$

under Condition C1 when $\epsilon$ is sufficiently small. So the proof of Theorem 2.5 is completed.

## 2.5. Asymptotic Relative Efficiency

Let us compute the asymptotic relative efficiency (ARE) $e_\epsilon$ of $M_\epsilon$ and $\overline{X}_\epsilon$, as the limiting ratio of "equivalent sample sizes." This computation, if it is based on variances rather than mean-square errors, calls for the additional assumption of symmetry of our TEIFR distribution, under which $P_\epsilon = Q_\epsilon = R_\epsilon$. What is required here is the number $e_\epsilon$ such that

$$\lim_{n\to\infty} \frac{Var_n(M_\epsilon)}{Var_m(\overline{X}_\epsilon)} = 1 \tag{2.41}$$

for

$$m = [e_\epsilon n]. \tag{2.42}$$

It is natural (because of the continuity aspect discussed in the next chapter) to replace (2.41) with

$$\lim_{n\to\infty} \ln MSE_n(M_\epsilon)/\ln MSE_m(\overline{X}_\epsilon) = 1. \tag{2.43}$$

Therefore, we assert the following:

**Theorem 2.6** *Suppose a random sample $\{X_i\ ;\ \ i = 1, 2, \ldots, m\}$ is drawn from the TEIFR class distribution whose location parameter is $\mu$. Under Condition C1 and when $\epsilon$ is sufficiently small, the asymptotic relative efficiency is:*

$$e_\epsilon = \frac{\theta_M(\epsilon)}{\theta_{\overline{X}}(\epsilon)} = -2\sigma^2\epsilon^{-2}(1+\delta(\epsilon))^{-1}\ln[2(R_\epsilon(1-R_\epsilon))^{1/2}]. \tag{2.44}$$

Proof:

Under (2.42), and in view of relation (2.21) and (2.38), relation (2.43) becomes

$$\begin{aligned}
1 &= \lim_{n\to\infty} \frac{\ln MSE_n(M_\epsilon)}{\ln MSE_m(\overline{X}_\epsilon)} \\
&= \lim_{n\to\infty} \left(\frac{n}{m}\right)\left[\frac{n^{-1}\ln MSE_n(M_\epsilon)}{m^{-1}\ln MSE_m(\overline{X}_\epsilon)}\right] \\
&= (e_\epsilon^{-1})\frac{\theta_M(\epsilon)}{\theta_{\overline{X}}(\epsilon)}.
\end{aligned}$$

This completes the proof of Theorem 2.6.

# 3. FINE GRIDS AND COARSE GRIDS

## 3.1. Introduction

In the previous chapter we found the asymptotic relative efficiency (ARE) $e_\epsilon$ of the rounded sample median $M_\epsilon$ and the rounded sample mean $\overline{X}_\epsilon$ as the limiting ratio of "equivalent sample sizes" under the TEIFR distribution. Now, we find $e_\epsilon$ to be surprisingly sensitive to distribution shape, as well as to the grid mesh size, and $e_\epsilon$ is right-continuous in $\epsilon$ at $\epsilon = 0$. In other words, if the population distribution possesses a density $f$ in a neighborhood of $\mu$, then

$$\lim_{\epsilon \to 0} e_\epsilon = 4f^2(\mu)\sigma^2,$$

the RHS being the usual asymptotic relative efficiency $e_0$ of $M$ vs. $\overline{X}$ for populations possessing finite variance and a density in a neighborhood of the population median. The asymptotic relative efficiency of $M_\epsilon$ vs. $\overline{X}_\epsilon$ may therefore be said to be "continuous at zero" in the sense that $e_\epsilon$ does converge to $e_0$. This continuity at zero has the further interpretation that the "asymptotic effective variance" of $M$ in the sense of Bahadur [2] equals its usual asymptotic variance $(4f^2(\mu)n)^{-1}$, for the types of population distributions in question. In this connection Section 3.2 dominates that "continuity at zero" does not hold for the form of asymptotic efficiency dealt with by Hammersley [26]. Furthermore, within the TEIFR class and given continuity of $e_\epsilon$ at $\epsilon = 0$, the "asymptotic effective variance" in the sense of Bahadur [2] of the sample median $M$ equals its asymptotic variance.

This chapter expands the analysis of $e_\epsilon$ to $\epsilon$'s of arbitrary size, in the special Laplace and Normal cases. We point to a certain equivocal behavior of asymptotic relative efficiency away from $\epsilon = 0$ in the case of Laplace distribution. Which of $M_\epsilon$ and $\overline{X}_\epsilon$ is more efficient turns out to depend on the value of $\epsilon$. No such equivocal behavior occurs in the Normal case, where the increasing superiority of $\overline{X}_\epsilon$ over $M_\epsilon$ with increasing $\epsilon$ is conveniently quantified via a certain quartic lower bound for a symmetric version of Mill's ratio $R(\epsilon)$ (see Kendall and Stuart [31]):

$$\begin{aligned} R(\epsilon)R(-\epsilon) &\equiv \frac{\Phi(\epsilon)\Phi(-\epsilon)}{\phi(\epsilon)\phi(-\epsilon)} \\ &\geq \frac{\pi}{2}\exp[(\frac{\pi-2}{\pi})\epsilon^2 + (\frac{1}{2\pi})\epsilon^4], \end{aligned}$$

This last resuly is derived by comparing the rates of decrease of the probabilities of large (beyond $\epsilon$) deviations, respectively of $|M - \mu|$ and $|\overline{X} - \mu|$. Hence they pertain as well to the comparison of the asymptotic error rates of certain tests based respectively on $M$ and $\overline{X}$, since these tests' asymptotic error rates essentially are themselves large deviation rates. Thus the findings of this chapter have testing counterparts which are discussed briefly in Section 3.4.

## 3.2. Fine Grids

This section uses the analysis of Chapter 2 to make the following three points:

1. $e_\epsilon$ is in a sense continuous to the right at $\epsilon = 0$.
2. That right-continuity has the interpretation that in the sense of Bahadur [2] the "asymptotic effective variance" of $M$ equals its asymptotic variance.
3. When asymptotic relative efficiency is viewed in the sense of Hammersley [26], right-continuity at $\epsilon = 0$ is no longer obtained.

Regarding the right-continuity of $e_\epsilon$ at 0, let us add the assumption that $X_i - \mu$ possesses a differentiable density $f$ in a neighborhood of 0, and define $e_0$ as the usual asymptotic relative efficiency of $M$ vs. $\overline{X}$, i.e., as limiting ratio of equivalent sample sizes (David [19]),

$$e_0 = 4\sigma^2 f^2(\mu). \tag{3.1}$$

Recall that $e_\epsilon$, as defined by relation (2.44), involves $R_\epsilon$, the maximum of $P_\epsilon$ and $Q_\epsilon$. If now $e_\epsilon^{(P)}$ and $e_\epsilon^{(Q)}$ are respectively defined as the analogues of (2.44) with $R_\epsilon$ respectively replaced by $P_\epsilon$ and $Q_\epsilon$, then L'Hôpital's Rule, plus relation (2.24), yields the conclusion that both $e_\epsilon^{(P)}$ and $e_\epsilon^{(Q)}$ tend to $e_0$ as $\epsilon$ tend to zero. Hence

$$\begin{aligned} \lim_{\epsilon \to 0} e_\epsilon &= \lim_{\epsilon \to 0} \max(\, e_\epsilon^{(P)},\ e_\epsilon^{(Q)}\,) \\ &= \max(\, \lim_{\epsilon \to 0} e_\epsilon^{(P)},\ \lim_{\epsilon \to 0} e_\epsilon^{(Q)}\,) \\ &= e_0, \end{aligned} \tag{3.2}$$

a relation that we describe by saying that $e_\epsilon$ is continuous to the right at $\epsilon = 0$.

Regarding the "asymptotic effective variance" of $M$, recall first that, since (2.7) and (2.10) dominate the other terms of (2.6), relation (2.21) has the equivalent interpretation that

$$\lim_{n\to\infty} n^{-1} \ln \Pr(|M-\mu| \geq \epsilon) = \ln[2(R_\epsilon(1-R_\epsilon))^{1/2}]. \tag{3.3}$$

We now invoke relations (1.4) and (1.7) of Bahadur [2], which say that, if there is a function $V_n(\mu)$ such that

$$\lim_{\epsilon\to 0} \lim_{n\to\infty} \frac{-\epsilon^2}{2V_n(\mu)\ln \Pr(|M-\mu| \geq \epsilon)} = 1, \tag{3.4}$$

then $V_n(\mu)$ is the "asymptotic effective variance" of $M$. But

$$V_n^*(\mu) = (4f^2(\mu)n)^{-1},$$

the usual asymptotic variance of $M$, does satisfy (3.4), and also is therefore the asymptotic effective variance of $M$. This can be shown as follows:

In view of relations (3.3) and (2.44),

$$\begin{aligned}
&\lim_{n\to\infty}\left[\frac{-\epsilon^2}{2V_n^*(\mu)\ln \Pr(|M-\mu| \geq \epsilon)}\right] \\
&= \frac{-2\epsilon^2 f^2(\mu)}{\ln[2(R_\epsilon(1-R_\epsilon))^{1/2}]} \\
&= \frac{4\sigma^2 f^2(\mu)}{e_\epsilon(1+\delta(\epsilon))},
\end{aligned}$$

and, in view of relations (2.31), (3.1) and (3.2), this last quantity tends to unity as $\epsilon$ tends to zero, so that $V_n^*(\mu)$ does satisfy relation (3.4), as claimed.

Finally, regarding Hammersley's alternative definition of asymptotic relative efficiency, Hammersley [26] computes the ARE of $M_\epsilon$ vs. $\overline{X}_\epsilon$ not in terms of the limiting ratio of equivalent sample sizes, but rather as

$$\lim_{n\to\infty} \frac{Var_n(\overline{X}_\epsilon)}{Var_n(M_\epsilon)} \equiv e_\epsilon^*, \tag{3.5}$$

which, when both variances can be approximated in terms of large-deviation rates for $\epsilon > 0$, will either be zero or infinity, unless the two large-deviation rates happen to coincide.

On the other hand, in the symmetric TEIFR case,

$$\lim_{n\to\infty} \frac{Var_n(\overline{X})}{Var_n(M)} \equiv e_0^*$$

equals $e_0$ as given by (3.1).

Typically, then, what will happen is that $e_\epsilon^*$ will be zero or infinite for $\epsilon$ in some open interval to the right of zero, with $e_0^*$ finite and non-zero, and $e_\epsilon^*$ therefore not right-continuous at zero.

The disparate behavior of $e_\epsilon$ and $e_\epsilon^*$ near zero is traceable to the fact that, for the "regular" order $n^{-1}$ large-sample variance $(4f^2(\mu)n)^{-1}$ and $\sigma^2/n$ of $M$ and $\overline{X}$, the two definitions of ARE (i.e., limiting ratio of equivalent sample sizes, and limiting variance ratio) coincide, while, for the "non-regular" exponential-order large-sample variances of $M_\epsilon$ and $\overline{X}_\epsilon$, especially of form $\exp(-n\theta_M(\epsilon))$ and $\exp(-n\theta_{\overline{X}}(\epsilon))$, they do not.

### 3.3. Coarse Grids

For $\epsilon$ not necessarily small, Condition C1 still covers what is needed for the analysis of $M_\epsilon$. However, Condition C1, while guaranteeing the existence of the moment generating function of $X_i - \mu$ in some neighborhood of the origin, which is the only requirement for our small $\epsilon$ analysis of $\overline{X}_\epsilon$ based on Lemma 2.4 of Bahadur [2], does not guarantee the existence of the moment generating function far enough away from the origin to accommodate an arbitrary $\epsilon$. What is needed for both the distribution $F(y)$ of $X_i - \mu$ and the distribution $G(y) \equiv 1 - F(-y)$ of $-(X_i - \mu)$ is the existence of a $t_{1,\epsilon}$ and $t_{2,\epsilon}$ satisfying the requirement stated in italics for $\tau$, between relation (2.3) and relation (2.4) of Bahadur and Rao [4], with their $a$ corresponding to our $\epsilon$. Such $t_{1,\epsilon}$ and $-t_{2,\epsilon}$ do exist for $\epsilon$ of any size, for both the standard Normal and standard Laplace distributions, and allow computing the large deviation rate

$$\theta_{\overline{X}}^{(N)}(\epsilon) \equiv \epsilon^2/2 \tag{3.6}$$

for the standard Normal (which, when (3.6) is set equal to (2.39), gives $\delta^{(N)}(\epsilon) \equiv 0$ trivially satisfying (2.31)), and, after some computation, the large deviation rate

$$\theta_{\overline{X}}^{(L)}(\epsilon) \equiv \ln 2 - 1 - 2\ln\epsilon + \ln[\sqrt{(1+\epsilon)} - 1] + \sqrt{(1+\epsilon)} \tag{3.7}$$

for the standard Laplace (which, when (3.7) is set equal to (2.39) gives $\delta^{(L)}(\epsilon) = (4/\epsilon^2)(\theta_{\overline{X}}^{(L)}(\epsilon) - \epsilon^2/4)$ also satisfying (2.31)). Also, as pointed out at the beginning of Lemma 2.4, the argument in that lemma applies to the present case. Hence (3.6) and (3.7) describe the behavior of $Var_m(\overline{X}_\epsilon)$ in the sense of (2.40).

In addition, both the standard Laplace distribution and the standard Normal distribution satisfy Condition C1 for $\epsilon$ of any size, so that $\theta_M(\epsilon)$, as given by the middle term of (2.21), i.e.,

$$\theta_M^{(N)}(\epsilon) = -\ln 2\sqrt{\Phi(\epsilon)\{1-\Phi(\epsilon)\}} \tag{3.8}$$

for the standard Normal, and

$$\theta_M^{(L)}(\epsilon) = -\frac{1}{2}\ln[\exp(-\epsilon)\{2-\exp(-\epsilon)\}] \tag{3.9}$$

for the standard Laplace, describe the behavior of $Var_n(M_\epsilon)$ in the sense of (2.21).

Finally, the first equality of relation (2.44), with (3.6) and (3.8) substituted in the case of the Normal, and (3.7) and (3.9) in the case of the Laplace, then gives the asymptotic relative efficiency of $M_\epsilon$ vs. $\overline{X}_\epsilon$ for arbitrary $\epsilon$, in the sense of (2.42) and (2.43).

We now use (3.7) and (3.9) to show that, in the case of the standard Laplace distribution (for which $\sigma^2 = 2$ and $f(\mu) = 1/2$), the quantity $e_\epsilon^{-1}$ is not of the same sign for all $\epsilon$ on the entire half-line $(0, +\infty)$. Thus, $M_\epsilon$ is asymptotically more efficient than $\overline{X}_\epsilon$ for some mesh size, and less efficient for others. Indeed, we learn from (3.1) and (3.2) that $e_\epsilon$ is near 2 for $\epsilon$'s near zero. On the other hand, as will be demonstrated in the next paragraph, $e_\epsilon < 1$ for large $\epsilon$, all of which says that $M_\epsilon$ is asymptotically more efficient than $\overline{X}_\epsilon$ when $\epsilon$ is small, but asymptotically less efficient than $\overline{X}_\epsilon$ when $\epsilon$ is large.

The assertion that $e_\epsilon < 1$ for large $\epsilon$ follows by noting that $\theta_{\overline{X}}(\epsilon) - \theta_M(\epsilon) = \theta_{\overline{X}}(\epsilon)(1-e_\epsilon)$ is positive for $\epsilon$ large. This last assertion is followed in turn because, using (3.7) and (3.9), one finds that the derivative

$$\frac{\sqrt{(1+\epsilon^2)}-1}{\epsilon} + \frac{\exp(-\epsilon)-1}{2-\exp(-\epsilon)}$$

of $\theta_{\overline{X}}^{(N)}(\epsilon) - \theta_M^{(N)}(\epsilon)$ tends to the positive number $\frac{1}{2}$ with increasing $\epsilon$.

The Normal distribution, with $\overline{X}_\epsilon$ maximum-likelihood, does not exhibit this equivocal behavior; indeed, the difference $\theta_{\overline{X}}^{(N)}(\epsilon) - \theta_M^{(N)}(\epsilon)$, as computed using (3.6) and (3.8), is bounded below, for all $\epsilon \geq 0$, by the increasing quartic

$$[\frac{\pi - 2}{2\pi}]\epsilon^2 + [\frac{1}{4\pi}]\epsilon^4, \tag{3.10}$$

which may be seen as follows:

Let $\Phi$ and $\phi$ be, respectively, the cumulative and density of the standard Normal distribution, and assume $\sigma^2 = 1$. With $\theta_M^{(N)}(\epsilon) = -\ln 2\sqrt{\Phi(\epsilon)\{1 - \Phi(\epsilon)\}}$, one has, for all positive $\epsilon$, that

$$\begin{aligned}
\frac{d}{d\epsilon}[\theta_{\overline{X}}^{(N)}(\epsilon) - \theta_M^{(N)}(\epsilon)] &= \epsilon + \frac{\phi(\epsilon)\{1 - 2\Phi(\epsilon)\}}{2\Phi(\epsilon)\{1 - \Phi(\epsilon)\}} \\
&\geq \epsilon + \frac{[\frac{1}{\sqrt{(2\pi)}}(1 - \frac{\epsilon^2}{2})][1 - 2(\frac{1}{2} + \frac{\epsilon}{\sqrt{2\pi}})]}{2(\frac{1}{4})} \\
&= [\frac{\pi - 2}{\pi}]\epsilon + \frac{1}{\pi}\epsilon^2,
\end{aligned}$$

which is positive.

Since, in addition, $\theta_{\overline{X}}^{(n)}(0) = \theta_M^{(N)}(0) = 0$, it follows that

$$\theta_{\overline{X}}^{(N)}(\epsilon) - \theta_M^{(N)}(\epsilon) \geq [\frac{\pi - 2}{2\pi}]\epsilon^2 + \frac{1}{4\pi}\epsilon^4 \tag{3.11}$$

for all positive $\epsilon$, where the RHS is the bound given by (3.10). This bound turns out to be equivalent to a certain quartic bound on the symmetric version

$$R(\epsilon)R(-\epsilon) \equiv \frac{\Phi(\epsilon)\Phi(-\epsilon)}{\phi(\epsilon)\phi(-\epsilon)}$$

of Mill's ratio, which can be demonstrated by noting that relation (3.11) holds if and only if

$$\frac{2\sqrt{\Phi(\epsilon)\{1-\Phi(\epsilon)\}}}{\exp(-\epsilon^2/2)} \geq \exp[(\frac{\pi-2}{2\pi})\epsilon^2 + (\frac{1}{4\pi})\epsilon^4],$$

which in turn holds if and only if

$$R(\epsilon)R(-\epsilon) \geq \frac{\pi}{2}\exp[(\frac{\pi-2}{\pi})\epsilon^2 + (\frac{1}{2\pi})\epsilon^4],$$

which is the quartic Mill's ratio bound referred in the above.

## 3.4. Testing Hypotheses

We close this chapter with a restatement, in the context of tests, of the discussion concerning the standard Laplace distribution in Section 3.3.

Let $\mu$ be the unknown center of a standard Laplace distribution, and consider testing $H_o : \mu = \mu_o$ vs, $H_1 : \mu = \mu_1 = \mu_o + 2\epsilon$ using one or the other of the following two tests:

A test $T_1$ rejecting $H_o$ if $M \geq c_1$, with $c_1$ chosen so as to minimize the sum of the errors of both kinds; and a test $T_2$ rejecting $H_o$ if $\overline{X} \geq c_2$ with $c_2$ chosen so as to minimize the sum of the errors of both kinds.

Then, since both $M$ and $\overline{X}$ will be symmetrically distributed about $\mu$, both $c_1$ and $c_2$ will equal $(\mu_1 + \mu_o)/2 = \mu_o + \epsilon$, and the error probabilities for $T_1$ and $T_2$ will respectively be given by

$$\Pr(M - \mu_o \geq \epsilon)$$

and

$$\Pr(\overline{X} - \mu_o \geq \epsilon),$$

and these two error probabilities will decay exponentially to first order, with respective rates $\theta_M(\epsilon)$ and $\theta_{\overline{X}}(\epsilon)$. The remarks of this section therefore imply that $T_1$ will be the better test when $\epsilon$ is small, but the poorer test when $\epsilon$ is large.

# 4. GRANULARITY AND EFFICIENCY IN THE $P$-VARIATE REGRESSION MODEL

## 4.1. Introduction

Hammersley [26] considered the sample variance of the rounded sample mean $\overline{X}_\epsilon$ as an estimate of a Normal population mean $\mu$ rounded to a uniform grid of mesh size $2\epsilon$. Chapter 2 and Chapter 3 extended Hammersley's work to a certain class of "two-sided extended IFR (TEIFR)" distributions, via the theory of large deviations (Bahadur and Rao [4]; Bahadur [2,3]). These chapters dealt with a sequence of *iid* random variables $\{X_i \; ; \quad i = 1, 2, \ldots, n\}$ and verified that $\lim_{n\to\infty} n^{-1} \ln Var_n[\overline{X}_\epsilon] = -\frac{\epsilon^2}{2\sigma^2}[1 + \delta(\epsilon)]$, where $\sigma^2 = Var(X)$ and $\delta(\cdot)$ is such that $\lim_{\epsilon\to 0} \delta(\epsilon) = 0$. It was asserted in Lemma 2.4 of Chapter 2 that $\Pr(|\overline{X} - \mu| \geq \epsilon)$, which is the first term in the infinite series for $(2\epsilon)^{-2} Var_n[\overline{X}_\epsilon]$, dominates the rest of this series.

This chapter will consider the asymptotic variances and covariances of rounded least-squares estimators (RLSEs) $b_{j,\epsilon_j}$ in Normal $p$-variate regression models, in terms of the large deviations of $\Pr(|b_j - \beta_j| \geq \epsilon_j)$ and $\Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \geq \epsilon_j)$, respectively, where $b_i$ is the least-squares estimator of the grid-valued parameter $\beta_i$ for $i = 1, 2, \ldots, p$, rounded to the grid. The asymptotic variances and covariances of these rounded regression estimates are analyzable in terms of their large deviations in a manner analogous to the analysis of the rounded sample mean $\overline{X}_\epsilon$.

Assume that the disturbance vector is distributed $N(\underline{0},\ \sigma^2\underline{I})$ and there is a positive definite matrix $\underline{Q}$ such that $\lim_{n\to\infty} n^{-1}\underline{X}'\underline{X} = \underline{Q}$. To implement the regression analysis, Section 4.3 contains some lemmas that extend to the $b_j$'s the treatment of large deviations of sample means in Bahadur and Rao [4] and Bahadur [2,3]. These lemmas yield large-deviation rates for regression parameter vectors $\underline{b}$:

$$n^{-1}\ln \Pr(\underline{b} - \underline{\beta} \geq \underline{\epsilon}) = n^{-1}(-\frac{1}{2}\underline{\epsilon}'\underline{V}\underline{\epsilon}) + o_n(1).$$

The asymptotic variances and covariances of the RLSEs are formulated in section 4.4:

The asymptotic variance $Var_n(b_{j,\epsilon_j})$ of the RLSE $b_{j,\epsilon_j}$, obtained by rounding to the grid $(2k\epsilon_j\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$, satisfies

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\ln Var_n(b_{j,\epsilon_j}) &= \lim_{n\to\infty} -\frac{\epsilon_j^2}{2n\sigma_{n,j}^2} \\ &= -\frac{\epsilon_j^2 Q_{j,j}}{2\sigma^2}, \qquad (4.1)\end{aligned}$$

where $Q_{j,j}$ is the $(j,j)$ element of $\underline{Q}$.

The asymptotic covariance $Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j})$ of the RLSEs $b_{i,\epsilon_i}$ and $b_{j,\epsilon_j}$, rounded the grids $(2k\epsilon_i,\ 2k\epsilon_j\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$ satisfy

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\ln sgn(corr_{i,j})Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) &= \lim_{n\to\infty} -\frac{1}{2n}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon} \\ &= -\frac{1}{2\sigma^2}\underline{\epsilon}'\underline{Q}\underline{\epsilon},\end{aligned}$$

where $corr_{i,j}$ is the correlation between $b_i$ and $b_j$, and

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

The analysis is illustrated with the simple regression model.

The joint asymptotic efficiency (JAE) may be defined as the determinant of the asymptotic variance-covariance matrix, $\underline{V}^a$, leading to the JAE of RLSEs in the form

$$\begin{aligned} \lim_{n \to \infty} n^{-1} \ln |\underline{V}^a| &= \lim_{n \to \infty} \sum_{i=1}^{n} n^{-1} \ln Varn(b_{i,\epsilon_i}) \\ &= -\sum_{i=1}^{n} \frac{\epsilon_i^2 Q_{i,i}}{2\sigma^2}, \end{aligned}$$

which is the summation of expression (4.1).

## 4.2. Assumptions and Notation

The model is an *iid* multiple regression model. The vector of sample observations $\underline{Y}$ is expressed as a linear combination of $p$ explanatory vectors $\underline{X}_j$ plus an *iid* disturbance vector $\underline{u}$:

$$\underline{Y} = \underline{X}_1\beta_1 + \underline{X}_2\beta_2 + \cdots + \underline{X}_p\beta_p + \underline{u}, \tag{4.2}$$

where each vector possesses $n$ elements, and the $n \times p$ matrix $\underline{X} = (\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_p)$

is assumed eventually to be of rank $p$.

Condition C2:

(1) The $n$ mutually independent elements of $\underline{u}$ are distributed $\mathbf{N}(0, \sigma^2)$.

(2) There exists a positive definite matrix $\underline{Q}$ such that $\lim_{n\to\infty} n^{-1}\underline{X}'\underline{X} = \underline{Q}$.

The Condition C2(2) implies that the $p$ averages $\overline{X}_{n,j} = \sum_{i=1}^{n} X_{ij}/n$ and sample variances $S^2_{x_j} = \sum_{i=1}^{n}(X_{ij} - \overline{X}_{n,j})^2/n$ are assumed, respectively, to converge to constants $\mu_{x_j}$ and $\sigma^2_{x_j}$, $(-\infty < \mu_{x_j} < \infty, \quad 0 < \sigma^2_{x_j} < \infty)$ as sample size $n$ increases.

Let $\underline{b} = (b_1, b_2, \ldots, b_p)'$ be the vector of least-squares estimators, which are non-*iid* random variables. For $i$ and $j$ such that $i \neq j$, and $\epsilon_i$ and $\epsilon_j$ constants, define

$$P_{n,\epsilon_j} \equiv \Pr(b_j - \beta_j \geq \epsilon_j) \tag{4.3}$$

and

$$P_{n,\epsilon_i,\epsilon_j} \equiv \max(P^+_{n,\epsilon_i,\epsilon_j}, \quad P^-_{n,\epsilon_i,\epsilon_j}), \tag{4.4}$$

where

$$P^+_{n,\epsilon_i,\epsilon_j} \equiv \Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \geq \epsilon_j)$$

and

$$P^{-}_{n,\epsilon_i,\epsilon_j} \equiv \Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \leq -\epsilon_j).$$

Each $b_j - \beta_j$, which is a linear function of *iid* random variables $u_1, u_2, \ldots, u_n$, has a moment generating function with the Condition C2(1). Also $b_i - \beta_i$ and $b_j - \beta_j$ have a joint moment generating function. Hence we can define, for real $t_i$ and $t_j$,

$$\begin{aligned} \psi n(t_j) &\equiv \exp(-\epsilon_j t_j)\phi n(t_j) \\ &\equiv \exp(-\epsilon_j t_j)E[\exp(b_j - \beta_j)t_j] \\ &= \exp(-\epsilon_j t_j + \frac{1}{2}\sigma^2_{n,j}t_j^2) \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} \psi n(t_i, t_j) &\equiv \exp(-\epsilon_i t_i - \epsilon_j t_j)\phi n(t_i, t_j) \\ &\equiv \exp(-\epsilon_i t_i - \epsilon_j t_j)E[\exp((b_i - \beta_i)t_i + (b_j - \beta_j t_j))] \\ &= \exp(-\underline{\epsilon}'\underline{t} + \frac{1}{2}\underline{t}'\underline{V}\underline{t}), \end{aligned} \tag{4.6}$$

where

$$\begin{aligned} \sigma^2_{n,j} &= Var(b_j), \\ \underline{\epsilon} &= (0, \ldots, 0, \epsilon_i, 0, \ldots, 0, \epsilon_j, 0, \ldots, 0)', \\ \underline{t} &= (0, \ldots, 0, t_i, 0, \ldots, 0, t_j, 0, \ldots, 0)' \end{aligned}$$

and

$$\underline{V} = Var(\underline{b}) = \sigma^2(\underline{X}'\underline{X})^{-1}.$$

In addition, there exist $\tau_{n,j}$ and $(\tau_{n,i}, \tau_{n,j})$, respectively, such that

$$\psi n(\tau_{n,j}) = \inf_{t_j} \psi n(t_j) \equiv \rho n(\epsilon_j) \tag{4.7}$$

and

$$\psi n(\tau_{n,i}\tau_{n,j}) = \inf_{t_i,t_j} \psi n(t_i, t_j) \equiv \rho n(\epsilon_i, \epsilon_j). \tag{4.8}$$

## 4.3. Modified Lemmas

Bahadur and Rao [4] and Bahadur [2,3] considered $\Pr(|\overline{X} - \mu| \geq \epsilon)$ under the situation for which $X_1, X_2, \ldots, X_n$ are *iid* random variables and the sample mean $\overline{X}$ is an estimator of the population mean $\mu$.

However, for every $i$ and $j$ such that $i \neq j$, both $P_{n,\epsilon_j}$ and $P_{n,\epsilon_i,\epsilon_j}$ under the model (4.2) involve the vector $\underline{b} - \underline{\beta} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{u}$ that is a linear function of the *iid* components of the vector $\underline{u}$. Therefore, the large-deviation treatment of $\overline{X}$ needs to be modified to apply to $\underline{b}$.

The large-deviation treatment of $|\overline{X} - \mu|$ in Chapter 2 relied on Lemma 2.4 of Bahadur [2], but Lemmas of Bahadur and Rao [4] and Bahadur [3] can be extended to cover the large-deviation behavior of $|b_j - \beta_j|$.

Let $Y_{n,j} = b_j - \beta_j - \epsilon_j$ and let $Fn(\cdot)$ and $Fn(\cdot,\cdot)$, respectively, be the distribution function of $Y_{n,j}$ and $(Y_{n,i}, Y_{n,j})$. Define $Gn(\cdot)$ and $Gn(\cdot,\cdot)$, respectively, as the distribution functions of random variables $Z_{n,j}$ and $(Z_{n,i},\ Z_{n,j})$ obtained from $Fn(\cdot)$ and $Fn(\cdot,\cdot)$ by "exponential centering", as explained on page 13; i.e.,

$$dGn(z_j) = \rho n(\epsilon_j)^{-1} \exp(\tau_{n,j} z_j) dFn(z_j),$$

and

$$dGn(z_i, z_j) = \rho n(\epsilon_i, \epsilon_j)^{-1} \exp(\tau_{n,i} z_i + \tau_{n,j} z_j) dFn(z_i, z_j),$$

where $\rho n$ and $\tau n$'s are defined in expressions (4.7) and (4.8). Note that the distributions $Fn$ and $Gn$ are, as a matter of fact, Normal distributions; however, the notation of Bahadur and Rao [4] is followed here.

**Lemma 4.1** *The moment generating functions of $Z_{n,j}$ and $(Z_{n,i},\ Z_{n,j})$ exist in neighborhoods of the origin, and we have*

$$E(Z_{n,j}) = 0, \qquad 0 < Var(Z_{n,j}) < \infty.$$

Proof:

The univariate case is covered in Bahadur and Rao [2]; to prove the existence of the moment generating function in the bivariate case, proceed as follows:
Let $\xi n(t_i, t_j)$ denote the m.g.f. of $(Z_{n,i}, Z_{n,j})$ for the real numbers $t_i, t_j$.
Then

$$\xi n(t_i, t_j) = \psi n(\tau_{n,i} + t_i, \tau_{n,j} + t_j) / \rho n(\epsilon_i, \epsilon_j).$$

Since $\psi_n(t_i, t_j) < \infty$ in a neighborhood of $(t_i, t_j) = (\tau_{n,i}, \tau_{n,j})$, it follows that $\xi_n(t_i, t_j) < \infty$ in a neighborhood of $(t_i, t_j) = (0, 0)$. This completes the proof of Lemma 4.1.

Let

$$U_j(n) = Z_{n,j}/\sigma_{n,j},$$

and note that

$$\sigma_{nj}^2 = Var(Z_{n,j}) = \phi_n''(\tau_{n,j}) - \epsilon_j^2.$$

Then define

$$H_{n,j}(x) = \Pr(U_j(n) \leq x), \qquad -\infty < x < \infty.$$

**Lemma 4.2** *Expression (4.3) equals* $\rho_n(\epsilon_j)I_{n,j}$, *where*

$$I_{n,j} = \sigma_{n,j}\tau_{n,j} \int_0^\infty \exp(-\sigma_{n,j}\tau_{n,j}x)[H_{n,j}(x) - H_{n,j}(0)]dx.$$

The proof is similar to that in Bahadur and Rao [4].

Now, for an arbitrary positive $\delta$, we can derive, by integration by parts,

$$[H_{n,j}(\delta) - H_{n,j}(0)]\exp(-\sigma_{n,j}\tau_{n,j}\delta) \leq I_{n,j} \leq 1. \qquad (4.9)$$

The conditions on the fixed explanatory variables described in Section 4.2 imply that $\sigma_{n,j}$ converges to zero as sample size increases, so that we obtain

$$n^{-1} \ln I_{n,j} = o_n(1)$$

and

$$n^{-1} \ln P_{n,\epsilon_j} = n^{-1} \ln \rho_n(\epsilon_j) + o_n(1). \tag{4.10}$$

Now consider the bivariate case. Let

$$V_{i,j}(n) = \underline{\tau}_n' \underline{Z}_n / [\underline{\tau}_n' \underline{V} \underline{\tau}_n]^{1/2} \equiv \underline{\tau}_n' \underline{Z}_n / \sigma_n^*,$$

where $\underline{\tau}_n = (\tau_{n,i}, \tau_{n,j})'$ and $\underline{Z}_n = (Z_{n,i}, Z_{n,j})'$, and define

$$K_{n,i,j}(x) = \Pr(V_{i,j}(n) \leq x), \qquad -\infty < x < \infty.$$

**Lemma 4.3** *$P^+_{n,\epsilon_i,\epsilon_j}$ in expression (4.4) equals $\rho_n(\epsilon_i, \epsilon_j) I_{n,i,j}$, where*

$$I_{n,i,j} = \sigma_n^* \int_0^\infty \exp(-\sigma_n^* x)[K_{n,i,j}(x) - K_{n,i,j}(0)]dx.$$

Proof:

$P^{+}_{n,\epsilon_i,\epsilon_j} = \Pr(b_i - \beta_i \geq \epsilon_i, b_j - \beta_j \geq \epsilon_j)$ equals

$$\begin{aligned}
&\Pr(Y_{n,i} \geq 0, Y_{n,j} \geq 0) \\
&= \rho_n(\epsilon_i, \epsilon_j) \int\int_{\{z_i \geq 0, z_j \geq 0\}} \exp(-\tau_{n,i} z_i - \tau_{n,j} z_j) dG_n(z_i, z_j) \\
&= \rho_n(\epsilon_i, \epsilon_j) \int_{\{x \geq 0\}} \exp(-\sigma_n^* x) dK_{n,i,j}(x) \\
&= \rho_n(\epsilon_i, \epsilon_j) \sigma_n^* \int_0^\infty \exp(-\sigma_n^* x)[K_{n,i,j}(x) - K_{n,i,j}(0)] dx,
\end{aligned}$$

where the last equality is followed by integration by parts and $\sigma_n^*$ was defined in Section 4.3. The proof is completed. Note that, actually, it provides as well the argument for Lemma 4.2.

.With an argument analogous to that for inequality (4.9) relating to $I_{n,j}$, the following can be obtained, for $\delta$,

$$[K_{n,i,j}(\delta) - K_{n,i,j}(0)] \exp(-\sigma_n^* \delta) \leq I_{n,i,j} \leq 1,$$

so that one finds

$$n^{-1} \ln P^{+}_{n,\epsilon_i,\epsilon_j} = n^{-1} \ln \rho_n(\epsilon_i, \epsilon_j) + o_n(1).$$

Now, by performing the standard first order large-deviation computation (Bahadur and Rao [4]) extended to bivariate case, we have

$$P^{-}_{n,\epsilon_i,\epsilon_j} \leq \rho_n(\epsilon_i, \epsilon_j)$$

for all $n$; in addition, for any $\delta_n^* > 0$,

$$P_{n,\epsilon_i,\epsilon_j}^{-} \geq \rho_n(\epsilon_i, \epsilon_j)(1 - \delta_n^*)$$

for $n$ large enough.

Therefore, we can say

$$n^{-1} \ln P_{n,\epsilon_i,\epsilon_j} = n^{-1} \ln \rho_n(\epsilon_i, \epsilon_j) + o_n(1). \tag{4.11}$$

We may consider the $p$-variate large deviation in the model 4.1, i.e.,

$$\Pr(\underline{b} - \underline{\beta} \geq \underline{\epsilon}),$$

where

$$\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$$

and

$$\underline{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_p)'.$$

$\psi_n(\underline{t})$ was set up in (4.6) with $\underline{t} = (t_1, t_2, \ldots, t_p)'$ and then $\frac{d\psi_n(\underline{t})}{d\underline{t}} = \underline{0}$ gives $\underline{\tau} = \underline{V}^{-1}\underline{\epsilon}$, so that

$$\psi_n(\underline{\tau}) = \exp(-\frac{1}{2}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon}). \tag{4.12}$$

Then we may conclude that

$$n^{-1} \ln \Pr(\underline{b} - \underline{\beta} \geq \underline{\epsilon}) = n^{-1}(-\frac{1}{2}\underline{\epsilon}'\underline{V}\underline{\epsilon}) + o_n(1). \tag{4.13}$$

### 4.4. Asymptotic Variances and Covariances of RLSEs

Hammersley [26] and this dissertation, respectively, considered the rounded sample mean $\overline{X}_\epsilon$ as an estimate of the population mean, $\mu$, in the Normal case and within a certain "TEIFR" class in the Chapter 2, restricted to an uniform grid of mesh size $2\epsilon$. Now we may consider rounded least-squares estimators (RLSEs) $b_{j,\epsilon_j}$ of $\beta_j$ in the regression model (4.2) rounded to a uniform grid of mesh size $2\epsilon_j$ for every $j$.

As we formulated the asymptotic variances of the rounded sample mean $\overline{X}_\epsilon$ in terms of the large-deviation behavior of $|\overline{X} - \mu|$ in Chapter 2, the asymptotic variances of RLSEs can be developed through the large-deviation behavior of $|b_j - \beta_j|$:

For every $j$ and positive $\epsilon_j$,

$$(2\epsilon_j)^{-2} Varn(b_{j,\epsilon_j}) = (2\epsilon_j)^{-2} E[(b_{j,\epsilon_j} - \beta_j)^2] \tag{4.14}$$

$$= \{\Pr(|b_j - \beta_j| \geq \epsilon_j) \tag{4.15}$$

$$- \Pr(|b_j - \beta_j| \geq 3\epsilon_j) \tag{4.16}$$

$$+ \sum_{k=2}^{\infty} \Pi_k(n, \epsilon_j)\}, \tag{4.17}$$

$$\tag{4.18}$$

where

$$\Pi_k(n, \epsilon_j) = k^2 \Pr\{(2k-1)\epsilon_j \leq |b_j - \beta_j| < (2k+1)\epsilon_j\}.$$

It was shown in Lemma 2.4 of the Chapter 2 that $\Pr(|\overline{X} - \mu| \geq \epsilon)$ dominates the infinite series of $(2\epsilon)^{-2} Var(\overline{X}_\epsilon)$. Also, it is not hard to prove that the (4.15) term dominates both (4.16) and (4.17) with Condition C2, and this will be clear after examining the analysis of the asymptotic covariances of RLSEs $b_{i,\epsilon_i}$ and $b_{j,\epsilon_j}$ shown below.

**Theorem 4.4** *Under the Condition C2 about the model (4.2) described in the section 4.2, the asymptotic variance $Varn(b_{j,\epsilon_j})$ of the RLSE $b_{j,\epsilon_j}$, rounded the grid $(2k\epsilon_j\ ;\quad k = 0, \pm 1, \pm 2, \ldots)$ satisfies*

$$\begin{aligned} \lim_{n\to\infty} n^{-1} \ln Varn(b_{j,\epsilon_j}) &= \lim_{n\to\infty} -\frac{\epsilon_j^2}{2n\sigma_{n,j}^2} \\ &= -\frac{\epsilon_j^2 \mathbf{Q}_{j,j}}{2\sigma^2}, \end{aligned} \tag{4.19}$$

*where $\mathbf{Q}_{j,j}$ is the $(j,j)$ element of $\underset{\sim}{Q}$.*

Proof:

Under Lemma 4.2 and equation (4.10), it is true that

$$\lim_{n\to\infty} n^{-1} \ln \Pr(|b_j - \beta_j| \geq \epsilon_j) = \lim_{n\to\infty} n^{-1} \ln \rho n(\epsilon_j),$$

so that we can write

$$\lim_{n\to\infty} n^{-1} \ln Varn(b_{j,\epsilon_j}) = \lim_{n\to\infty} n^{-1} \ln \rho n(\epsilon_j).$$

Now from expression (4.7),

$$\frac{d}{dt_j} \psi n(t_j) = 0$$

implies

$$\tau_{n,j} = \epsilon_j / \sigma^2_{n,j},$$

so that $\rho_n(\epsilon_j)$ in fact equals

$$\exp(-\epsilon_j^2 / 2\sigma^2_{n,j}).$$

Thus we have

$$\lim_{n\to\infty} n^{-1} \ln Var_n(b_{j,\epsilon_j}) = \lim_{n\to\infty} -\frac{\epsilon_j^2}{2n\sigma^2_{n,j}},$$

whose limits would be expressed in (4.19).

We now turn to the asymptotic covariance of the RLSEs $b_{i,\epsilon_i}$ and $b_{j,\epsilon_j}$, for every $i$ and $j$ ($i \neq j$), and positive $\epsilon_i$ and $\epsilon_j$. In this situation, it is helpful to consider the four quadrants of the plane, and we write the analogue of (4.14) as the sum of twenty terms:

$$(2\epsilon_i)^{-1}(2\epsilon_j)^{-1} Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j})$$

$$= (2\epsilon_i)^{-1}(2\epsilon_j)^{-1} E[(b_{i,\epsilon_i} - \beta_i)(b_{j,\epsilon_j} - \beta_j)]$$

$$= \{\Pr(b_i - \beta_i \geq \epsilon_i, b_j - \beta_j \geq \epsilon_j) \tag{4.20}$$

$$- \Pr(b_i - \beta_i \geq \epsilon_i, b_j - \beta_j \geq 3\epsilon_j) \tag{4.21}$$

$$- \Pr(b_i - \beta_i \geq 3\epsilon_i, b_j - \beta_j \geq \epsilon_j) \tag{4.22}$$

$$+ \Pr(b_i - \beta_i \geq 3\epsilon_i, b_j - \beta_j \geq 3\epsilon_j) \tag{4.23}$$

$$+ \sum_{l=1}^{\infty} \sum_{m=1 (l=m\neq 1)}^{\infty} \Pi^{+,+}_{l,m}(n, \epsilon_i, \epsilon_j)\} \tag{4.24}$$

$$-\{\Pr(b_i - \beta_i \geq \epsilon_i, b_j - \beta_j \leq -\epsilon_j) \quad (4.25)$$

$$- \Pr(b_i - \beta_i \geq \epsilon_i, b_j - \beta_j \leq -3\epsilon_j) \quad (4.26)$$

$$- \Pr(b_i - \beta_i \geq 3\epsilon_i, b_j - \beta_j \leq -\epsilon_j) \quad (4.27)$$

$$+ \Pr(b_i - \beta_i \geq 3\epsilon_i, b_j - \beta_j \leq -3\epsilon_j) \quad (4.28)$$

$$+\sum_{l=1}^{\infty} \sum_{m=1 (l=m\neq 1)}^{\infty} \Pi_{l,m}^{+,-}(n, \epsilon_i, \epsilon_j)\} \quad (4.29)$$

$$-\{\Pr(b_i - \beta_i \leq -\epsilon_i, b_j - \beta_j \geq \epsilon_j)$$

$$- \Pr(b_i - \beta_i \leq -\epsilon_i, b_j - \beta_j \geq 3\epsilon_j)$$

$$- \Pr(b_i - \beta_i \leq -3\epsilon_i, b_j - \beta_j \geq \epsilon_j)$$

$$+ \Pr(b_i - \beta_i \leq -3\epsilon_i, b_j - \beta_j \geq 3\epsilon_j)$$

$$+\sum_{l=1}^{\infty} \sum_{m=1 (l=m\neq 1)}^{\infty} \Pi_{l,m}^{-,+}(n, \epsilon_i, \epsilon_j)\}$$

$$+\{\Pr(b_i - \beta_i \leq -\epsilon_i, b_j - \beta_j \leq -\epsilon_j)$$

$$- \Pr(b_i - \beta_i \leq -\epsilon_i, b_j - \beta_j \leq -3\epsilon_j)$$

$$- \Pr(b_i - \beta_i \leq -3\epsilon_i, b_j - \beta_j \leq -\epsilon_j)$$

$$+ \Pr(b_i - \beta_i \leq -3\epsilon_i, b_j - \beta_j \leq -3\epsilon_j)$$

$$+\sum_{l=1}^{\infty} \sum_{m=1 (l=m\neq 1)}^{\infty} \Pi_{l,m}^{-,-}(n, \epsilon_i, \epsilon_j)\},$$

where

$$\Pi_{l,m}^{+,+}(n, \epsilon_i, \epsilon_j) = lm \Pr\{(2l-1)\epsilon_i \leq b_i - \beta_i < (2l+1)\epsilon_i,$$
$$(2m-1)\epsilon_j \leq b_j - \beta_j < (2m+1)\epsilon_j\},$$

$$\Pi_{l,m}^{+,-}(n,\epsilon_i,\epsilon_j) = lm\Pr\{(2l-1)\epsilon_i \le b_i - \beta_i < (2l+1)\epsilon_i, \\ -(2m+1)\epsilon_j < b_j - \beta_j \le -(2m-1)\epsilon_j\},$$

$$\Pi_{l,m}^{-,+}(n,\epsilon_i,\epsilon_j) = lm\Pr\{-(2l+1)\epsilon_i < b_i - \beta_i \le -(2l-1)\epsilon_i, \\ (2m-1)\epsilon_j \le b_j - \beta_j < (2m+1)\epsilon_j\},$$

$$\Pi_{l,m}^{-,-}(n,\epsilon_i,\epsilon_j) = lm\Pr\{-(2l+1)\epsilon_i < b_i - \beta_i \le -(2l-1)\epsilon_i, \\ -(2m+1)\epsilon_j < b_j - \beta_j \le -(2m-1)\epsilon_j\}.$$

**Lemma 4.5** *Verification that the term (4.20) dominates the terms (4.21) to (4.24) under the Condition C2.*

Proof:

Set $\hat{\beta}_i = b_i = a = \hat{\alpha}$, $\hat{\beta}_j = b_j = b = \hat{\beta}$, $\epsilon_i = \epsilon_0$ and $\epsilon_j = \epsilon_1$. Define, as in equation (4.6),

$$\phi_n(t_a,t_b) = E[\exp\{(a-\alpha)t_a + (b-\beta)t_b\}].$$

Then we can find $\tau_{n,a} > 0$ and $\tau_{n,b} > 0$ such that

$$\psi_n(\tau_{n,a},\tau_{n,b}) = \inf \exp(-\epsilon_0 t_a - \epsilon_1 t_b)\phi_n(t_a,t_b).$$

Using the expression (4.11), we then obtain

$$\lim_{n\to\infty} n^{-1} \ln \Pr(a-\alpha \geq \epsilon_0, \quad b-\beta \geq \epsilon_1) = \lim_{n\to\infty} n^{-1} \ln \psi_n(\tau_{n,a}, \tau_{n,b}),$$

from which we conclude that a lower bound for the term (4.20) is given:

$$\Pr(a-\alpha \geq \epsilon_0, \quad b-\beta \geq \epsilon_1) \geq \exp(-\epsilon_0\tau_{n,a} - \epsilon_1\tau_{n,b})\phi_n(\tau_{n,a}, \tau_{n,b}) \exp(1 - o_n(1)). \tag{4.30}$$

Now an upper bound for the term (4.24) will be given, for every $t_a > 0$ and $t_b > 0$:

$$\begin{aligned} \Pi^{+,+}_{l,m}(n, \epsilon_0, \epsilon_1) &\leq lm \Pr(a-\alpha \geq (2l-1)\epsilon_0, \quad b-\beta \geq (2m-1)\epsilon_1) \\ &\leq lm \exp[-(2l-1)\epsilon_0 t_a - (2m-1)\epsilon_1 t_b]\phi_n(t_a, t_b), \end{aligned} \tag{4.31}$$

so that, setting $t_a = \tau_{n,a}$ and $t_b = \tau_{n,b}$,

$$\begin{aligned} 0 &\leq \sum_{l=1}^{\infty}\sum_{m=1(l=m\neq 1)}^{\infty} \Pi^{+,+}_{l,m}(n, \epsilon_0, \epsilon_1) \\ &\leq \sum_{l=1}^{\infty}\sum_{m=1(l=m\neq 1)}^{\infty} lm \exp[-(2l-1)\epsilon_0\tau_{n,a} - (2m-1)\epsilon_1\tau_{n,b}]\phi_n(\tau_{n,a}, \tau_{n,b}). \end{aligned}$$

Hence, appealing to (4.30),

$$\begin{aligned} 0 &\leq \sum_{l=1}^{\infty}\sum_{m=1(l=m\neq 1)}^{\infty} \frac{\Pi^{+,+}_{l,m}(n, \epsilon_0, \epsilon_1)}{\Pr(a-\alpha \geq \epsilon_0, \quad b-\beta \geq \epsilon_1)} \\ &\leq \{\sum_{l=1}^{\infty}\sum_{m=1(l=m\neq 1)}^{\infty} lm \exp[-(2l-1)\epsilon_0\tau_{n,a} - (2m-1)\epsilon_1\tau_{n,b}]\} \exp(1 - o_n(1)). \end{aligned}$$

It remains to show that the first bracketed term tends to zero with $n$. This term can be written as

$$\sum_{l=0}^{\infty}\sum_{m=0(l=m\neq 0)}^{\infty} (l+1)(m+1) \exp(-2l\epsilon_0\tau_{n,a} - 2m\epsilon_1\tau_{n,b})$$

$$\leq \sum_{l=0}^{\infty} \sum_{m=0 (l=m\neq 0)}^{\infty} \exp[l(1-2\epsilon_0\tau_{n,a}) + m(1-2\epsilon_1\tau_{n,b})]$$

$$= \frac{1}{[1-\exp(1-2\epsilon_0\tau_{n,a})][1-\exp(1-2\epsilon_1\tau_{n,b})]} - 1,$$

which tends to zero as sample size $n$ increases to infinity since, when sample size increases, both $\tau_{n,a}$ and $\tau_{n,b}$ tend to infinity.

Comparing the terms (4.21), (4.22) and (4.23), to the term (4.24) is of course less involved, and requires just the inequality (4.30) along with the inequality (4.31) with $l = 1$ and/or $m = 1$. Thus the proof of the Lemma is completed.

In a similar way, it can be shown that the term (4.25) dominates the four other terms (4.26) to (4.29). Furthermore, using symmetry, we may write

$$(2\epsilon_i)^{-1}(2\epsilon_j)^{-1} Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j})$$

$$= 2\{\Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \geq \epsilon_j) - \Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \leq -\epsilon_j)\}$$

$$(1 + o_n(1)).$$

**Theorem 4.6** *Under the Condition C2 about the model (4.2) described in the section 4.2, the asymptotic covariance* $Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j})$ *of the RLSEs* $b_{i,\epsilon_i}$ *and* $b_{j,\epsilon_j}$, *rounded the grids* $(2k\epsilon_i,\ 2k\epsilon_j\ ;\ \ k = 0, \pm 1, \pm 2, \ldots)$ *satisfies*

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{i,j}) Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) = \lim_{n\to\infty} -\frac{1}{2n} \underline{\epsilon}' \underline{V}^{-1} \underline{\epsilon}$$

$$= -\frac{1}{2\sigma^2}\underline{\epsilon}'\underline{Q}\underline{\epsilon}, \qquad (4.32)$$

*where* $corr_{i,j}$ *is the correlation between* $b_i$ *and* $b_j$ *and*

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Proof:

If the correlation between $b_i$ and $b_j$ is positive, then $\Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \geq \epsilon_j)$ dominates $\Pr(b_i - \beta_i \geq \epsilon_i, \quad b_j - \beta_j \leq -\epsilon_j)$, and one finds

$$sgn(corr_{i,j})(2\epsilon_i)^{-1}(2\epsilon_j)^{-1}Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) = P_{n,\epsilon_i,\epsilon_j}(1 + o_n(1)).$$

Equation (4.11) allows the assertion that

$$\lim_{n\to\infty} n^{-1}\ln P_{n,\epsilon_i,\epsilon_j} = \lim_{n\to\infty} n^{-1}\ln \rho_n(\epsilon_i, \epsilon_j).$$

Then we can write

$$\lim_{n\to\infty} n^{-1}\ln sgn(corr_{i,j})Cov_n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) = \lim_{n\to\infty} n^{-1}\ln \rho_n(\epsilon_i, \epsilon_j).$$

Now in order to obtain $\psi_n(\tau_i, \tau_j)$ defined in (4.8), setting $\frac{d}{d\underline{t}}\psi_n(t_i, t_j)$ equal to $\underline{0}$ gives

$$\underline{\tau}_n^o = (0, \ldots, 0, \tau_{n,i}, 0, \ldots, 0, \tau_{n,j}, 0, \ldots, 0)' = \underline{V}^{-1}\underline{\epsilon},$$

so that we have

$$\rho n(\epsilon_i, \epsilon_j) = \exp(-\frac{1}{2}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon}),$$

with $\underline{\epsilon} = (0, \ldots, 0, \epsilon_i, 0, \ldots, 0, \epsilon_j, 0, \ldots, 0)$.

Therefore, all told, we have

$$\begin{aligned}\lim_{n\to\infty} n^{-1} \ln sgn(corr_{i,j}) Cov n(b_{i,\epsilon_i}, b_{j,\epsilon_j}) &= \lim_{n\to\infty} -\frac{1}{2n}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon} \\ &= -\frac{1}{2\sigma^2}\underline{\epsilon}'\underline{Q}\underline{\epsilon}.\end{aligned}$$

This completes the proof of Theorem 4.6.

To illustrate the above, consider a simple regression model whose conditions are described in the Condition C2;

$$Y_i = \beta_1 + \beta_2 X_i + u_i; \qquad i = 1, 2, \ldots, n.$$

With the variance and covariance matrix

$$\underline{V} = \sigma^2 \begin{pmatrix} \frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2} & \frac{-\bar{X}}{\sum(X_i - \bar{X})^2} \\ \frac{-\bar{X}}{\sum(X_i - \bar{X})^2} & \frac{1}{\sum(X_i - \bar{X})^2} \end{pmatrix}$$ and grid sizes $2\epsilon_1$ and $2\epsilon_2$,

equation (4.19) for the variance of $b_{1,\epsilon_1}$ tell us that

$$\lim_{n\to\infty} n^{-1} \ln Var n(b_{1,\epsilon_1}) = \lim_{n\to\infty} -\frac{\epsilon_1^2 \sum(X_i - \bar{X})^2}{2\sigma^2 \sum X_i^2}$$

$$= -\frac{\epsilon_1^2 \sigma_x^2}{2\sigma^2(\sigma_x^2 + \mu_x^2)}.$$

In a similar way, concerning $b_{2,\epsilon_2}$, we obtain

$$\lim_{n\to\infty} n^{-1} \ln Varn(b_{2,\epsilon_2}) = \lim_{n\to\infty} -\frac{\epsilon_2^2 \sum(X_i - \bar{X})^2}{2n\sigma^2}$$

$$= \lim_{n\to\infty} -\frac{\epsilon_2^2 S_x^2}{2\sigma^2}$$

$$= -\frac{\epsilon_2^2 \sigma_x^2}{2\sigma^2}.$$

Moreover, equation (4.32) for the covariance of $b_{1,\epsilon_1}$ and $b_{2,\epsilon_2}$ in this model becomes

$$\lim_{n\to\infty} n^{-1} \ln -Covn(b_{1,\epsilon_1}, b_{2,\epsilon_2}) = \lim_{n\to\infty} -\frac{1}{2\sigma^2}(\epsilon_1^2 + 2\bar{X}\epsilon_1\epsilon_2 + \frac{\sum X_i^2}{n}\epsilon_2^2)$$

$$= -\frac{1}{2\sigma^2}[\epsilon_1^2 + 2\mu_x\epsilon_1\epsilon_2 + (\sigma_x^2 + \mu_x^2)\epsilon_2^2].$$

## 4.5. Joint Asymptotic Efficiency

In order to consider the joint asymptotic efficiency (JAE), we may take the determinant of the asymptotic variance and covariance matrix, say $|\underline{V}^a|$. From the illustration with the simple regression model above, the determinant of the asymptotic variance and covariance matrix of the RLSEs $b_{1,\epsilon_1}$ and $b_{2,\epsilon_2}$ is as follows:

$$\exp\{-\frac{n\sigma_x^2}{2\sigma^2}(\frac{\epsilon_1^2}{\sigma_x^2+\mu_x^2}+\epsilon_2^2)+o_n(n)\}-\exp\{-\frac{n}{\sigma^2}[\epsilon_1^2+2\mu_x\epsilon_1\epsilon_2+(\sigma_x^2+\mu_x^2)\epsilon_2^2]+o_n(n)\}.$$

Since

$$\frac{\sigma_x^2}{2}(\frac{\epsilon_1^2}{\sigma_x^2+\mu_x^2}+\epsilon_2^2) < \epsilon_1^2+2\mu_x\epsilon_1\epsilon_2+(\sigma_x^2+\mu_x^2)\epsilon_2^2,$$

the product of the asymptotic variances dominates the square of the asymptotic covariance, so that we may say that, as sample size increases, the asymptotic covariances will be negligible. Therefore, the JAE of this illustration is

$$\lim_{n\to\infty} n^{-1}\ln|\underline{V}^a| = -\frac{\sigma_x^2}{2\sigma^2}(\frac{\epsilon_1^2}{\sigma_x^2+\mu_x^2}+\epsilon_2^2).$$

Now we can go back to the $p$-variate regression model to conclude that the JAE of the RLSEs is as follows:

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\ln|\underline{V}^a| &= \lim_{n\to\infty}\sum_{i=1}^{n} n^{-1}\ln Var_n(b_{i,\epsilon_i})\\ &= -\sum_{i=1}^{n}\frac{\epsilon_i^2 Q_{i,i}}{2\sigma^2},\end{aligned}$$

which is the summation of expression (4.19).

# 5. MULTIVARIATE GRANULARITY AND EFFICIENCY

## 5.1. Introduction and Assumptions

For a possible illustration involving an ordinary grid-valued location parameter, there comes to mind the grid of nameless Iowa dirt county roads whose purpose it is to frame the mile-square "section," and which therefore criss-cross the countryside a mile apart. They are made hazardous at their intersections in late summer by the tall corn that blocks the lines of sight. Since the roads are nameless and the distinguishing landmarks few, initial intersection accident reports to the sheriff, to the extent that they would be vectored in terms of hastily approximated distances from hastily chosen reference points not necessarily adjacent to the scene of the accident, could perhaps be thought of as furnishing a random sample from a bivariate density of location centered at the intersection accident site. Thus it may be of interest to consider the bivariate, and also the multivariate, cases.

Chapter 2 considered the rounded sample mean and the rounded sample median via Bahadur's treatment of the behavior of the probability of large deviations, i.e., $\Pr(\overline{X} - \mu \geq \epsilon)$ and $\Pr(M - \mu \geq \epsilon)$. In addition, the rounded least-square estimators $b_{j,\epsilon_j}$ for the $p$-variate regression model were considered in Chapter 4, where $\Pr(b_j - \beta_j \geq \epsilon_j)$ and $\Pr(b_i - \beta_i \geq \epsilon_i,\ \ b_j - \beta_j \geq \epsilon_j)$ were examined in order to formulate the asymptotic variances and covariances of the rounded least-squares estimators. Now, the present chapter will consider not-necessarily-Normal multivariate rounded sample means and rounded sample medians, treated by bivariate large deviations similar to those introduced for the regression analysis. For $j = 1, 2, \ldots, p$, let $\{X_{ij}\}$ be a random sample from a certain $p$-variate distribution $F$, satisfying the condition that $\mu_j = M_j$, plus Condition C1 described in Section 2.2. Also assume $Var(X_{ij}) = \sigma_j^2$, and $X_{.j}$ and $X_{.j'}$ ( $j \neq j'$ ) are not necessarily independent. Each $\overline{X}_j$ and $M_j$ are rounded to the nearest point of the uniform grid $2\epsilon_j$, yielding $\overline{X}_{j,\epsilon_j}$ and $M_{j,\epsilon_j}$, respectively. Then the asymptotic variances and asymptotic covariances of $\underline{\overline{X}_\epsilon} = ( \overline{X}_{1,\epsilon_1}, \overline{X}_{2,\epsilon_2}, \ldots, \overline{X}_{p,\epsilon_p} )$ and $\underline{M_\epsilon} = ( M_{1,\epsilon_1}, M_{2,\epsilon_2}, \ldots, M_{p,\epsilon_p} )$ are considered. In order to examine the asymptotic covariances, the bivariate distributions are assumed to be such that both $\Pr(X_1 - \mu_1 \geq \epsilon_1,\ X_2 - \mu_2 \geq \epsilon_2) \equiv P_{\epsilon_1,\epsilon_2}$ and $\Pr(X_1 - \mu_1 \geq \epsilon_1,\ X_2 - \mu_2 \leq \epsilon_2) \equiv Q_{\epsilon_1,\epsilon_2}$ are appropriately quadrant-symmetric, and bivariate log-concave in positive $(\epsilon_1,\ \epsilon_2)$, a property that we can call the "Two-sided Extended Bivariate Increasing Failure Rate (TEBIFR)". Thus the $p$-variate analysis needs to extend Condition C1 described in Section 2.2. Let us define $\gamma_{\epsilon_1} \equiv \sup\{\epsilon_1 : P_{\epsilon_1,\epsilon_2} > 0\}$, $\gamma_{\epsilon_2} \equiv \sup\{\epsilon_2 : P_{\epsilon_1,\epsilon_2} > 0\}$, $\gamma'_{\epsilon_1} \equiv \sup\{\epsilon_1 : Q_{\epsilon_1,\epsilon_2} > 0\}$, and $\gamma'_{\epsilon_2} \equiv \sup\{\epsilon_2 : Q_{\epsilon_1,\epsilon_2} > 0\}$. Then Condition C3 is as follows:

<u>Condition C3:</u>

(0) $\Pr(X_1 - \mu_1 \geq \epsilon_1,\ X_2 - \mu_2 \geq \epsilon_2) = \Pr(X_1 - \mu_1 \leq \epsilon_1,\ X_2 - \mu_2 \leq \epsilon_2)$
$\equiv P_{\epsilon_1,\epsilon_2}$ and

$$\Pr(X_1 - \mu_1 \geq \epsilon_1,\ \ X_2 - \mu_2 \leq \epsilon_2) = \Pr(X_1 - \mu_1 \leq \epsilon_1,\ \ X_2 - \mu_2 \geq \epsilon_2)$$
$$\equiv Q_{\epsilon_1,\epsilon_2}.$$

(1) $P_{\epsilon_1,\epsilon_2}$ is log-concave for $\{(\epsilon_1, \epsilon_2) : \epsilon_1 \in [0,\ \gamma_{\epsilon_1}),\ \epsilon_2 \in [0,\ \gamma_{\epsilon_2})\}$, and $Q_{\epsilon_1,\epsilon_2}$ is log-concave for $\{(\epsilon_1, \epsilon_2) : \epsilon_1 \in [0,\ \gamma'_{\epsilon_1}),\ \epsilon_2 \in [0,\ \gamma'_{\epsilon_2})\}$.

(2) $(\epsilon_1,\ \epsilon_2)$ is such that $R_{\epsilon_1,\epsilon_2} \equiv \max(\ P_{\epsilon_1,\epsilon_2},\ \ Q_{\epsilon_1,\epsilon_2}\ ) > 0$.

(3a) If $(\epsilon_1,\ \epsilon_2)$ is such that $P_{\epsilon_1,\epsilon_2} > 0$,
then $P_{\epsilon_1,\epsilon_2} > \max(\ P_{\epsilon_1,3\epsilon_2},\ P_{3\epsilon_1,\epsilon_2},\ P_{3\epsilon_1,3\epsilon_2}\ )$.

(3b) If $(\epsilon_1,\ \epsilon_2)$ is such that $Q_{\epsilon_1,\epsilon_2} > 0$,
then $Q_{\epsilon_1,\epsilon_2} > \max(\ Q_{\epsilon_1,3\epsilon_2},\ Q_{3\epsilon_1,\epsilon_2},\ Q_{3\epsilon_1,3\epsilon_2}\ )$.

We refer to Condition C3(1), as specifying that the distribution of $(X_1 - \mu_1,\ X_2 - \mu_2)$ be the TEBIFR. Condition C3(2) implies Condition C1(2), which guarantees a finite and non-zero variance. Later on, this Condition C3(2) is a requirement for being able to describe the behavior of $(\overline{X}_{i\epsilon_i},\ \overline{X}_{j\epsilon_j})$ in terms of bivariate large deviation. Finally, the rationale for requiring Condition C3(3) is as follows: Suppose that $P_{\epsilon_1,\epsilon_2} > 0$. Then since the medians of $(X_1 - \mu_1)$ and $(X_2 - \mu_2)$ are zero in view of the above condition that $\mu_j = M_j$, so that $P_{\epsilon_1,\epsilon_2} \leq 1/2$, Condition C3(3a) insures that $P_{\epsilon_1,\epsilon_2}(1 - P_{\epsilon_1,\epsilon_2})$ is greater than

$$\max\{\ P_{3\epsilon_1,\epsilon_2}(1 - P_{3\epsilon_1,\epsilon_2}),\ P_{\epsilon_1,3\epsilon_2}(1 - P_{\epsilon_1,3\epsilon_2}),\ P_{3\epsilon_1,3\epsilon_2}(1 - P_{3\epsilon_1,3\epsilon_2})\ \}.$$

This chapter extends the univariate analyses of Chapters 2 and 3, based on Bahadur and Rao's treatment of univariate large deviations, to the multivariate case for sample means and to the bivariate case for sample medians:

$$n^{-1} \ln \Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}) = \ln \rho_{\underline{\bar{X}}}(\underline{\epsilon}) + o_n(1),$$

where $\rho_{\underline{\bar{X}}}(\underline{\epsilon})$ is defined as $\min_{\underline{t}}[\exp(-\sum_{j=1}^{p} \epsilon_j t_j)\phi_{\underline{X}-\underline{\mu}}(\underline{t})]$, and

$$n^{-1} \ln \Pr(M_1 - \mu_1 \geq \epsilon_1,\ \ M_2 - \mu_2 \geq \epsilon_2) = \ln \rho_{\underline{M}}(\epsilon_1, \epsilon_2) + o_n(1),$$

where $\rho_{\underline{M}}(\epsilon_1, \epsilon_2)$ is defined as $\min_{(t_1,t_2)} [\exp(-\frac{t_1}{2} - \frac{t_2}{2})\phi_{Y_{\epsilon_1}, Y_{\epsilon_2}}(t_1, t_2)]$.

The asymptotic variances of the rounded sample mean and the rounded sample median have been already considered in Chapter 2. The preceding relations make possible the computations of the asymptotic covariances of such estimators:

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Cov_n(\overline{X}_{i,\epsilon_i},\ \overline{X}_{j,\epsilon_j}) = \ln \rho_{\underline{\bar{X}}}(\epsilon_i, \epsilon_j)$$

and

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Cov_n(M_{i,\epsilon_i},\ M_{j,\epsilon_j}) = \ln \rho_{\underline{M}}(\epsilon_i, \epsilon_j),$$

where $corr_{i,j}$ is the correlation between $b_i$ and $b_j$, and

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

We thus obtain the matrices of the asymptotic variance and covariance of the rounded sample means and the rounded sample medians, leading to the possibility of a comparison of $\underline{\overline{X}_\epsilon}$ and $\underline{M_\epsilon}$, in terms of the joint asymptotic efficiencies (JAEs), which were defined in Chapter 4. It is found that the JAEs of $\underline{\overline{X}_\epsilon}$ and $\underline{M_\epsilon}$ are the summations of univariate large-deviation rates:

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{\bar{X}})| = \sum_{j=1}^{p} \rho_{\bar{X}}(\epsilon_j)$$

and

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{M})| = \sum_{j=1}^{p} \rho_M(\epsilon_j).$$

These computations are illustrated using the Normal and Laplace distributions.

## 5.2. Multivariate Large-deviation Behavior

First of all, we begin with the multivariate large-deviations of sample means, i.e.,

$$\Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}). \tag{5.1}$$

Lemma 2.3 showed in Section 2.4 that a distribution which belongs to the TEIFR class has a moment generating function. Analogously, if the distribution $F$ of $\underline{X} - \underline{\mu}$ is a "two-sided multivariate IFR (TEMIFR)", then $F$ also possesses a moment generating function. Since, for given $j$, $\{X_{ij}\ ;\ \ i = 1, 2, \ldots, n\}$ are independent and have a TEMIFR distribution $F$, an upper bound of (5.1) will be obtained by using Markov's inequality explained in Section 1.2:

$$\begin{aligned}\min_{\underline{t}} [\exp(-\sum_{j=1}^{p} \epsilon_j t_j)\phi_{\underline{X}-\underline{\mu}}(\underline{t})]^n &\equiv \min_{\underline{t}} [\psi_{\underline{\bar{X}}}(\underline{t})]^n \\ &\equiv \rho^n_{\underline{\bar{X}}}(\underline{\epsilon}).\end{aligned}$$

Hence we may conclude that, by multivariate Chernoff's Theorem and Bahadur and Rao's [4] Lemma 2.4 (especially, equation (1.2)),

$$n^{-1} \ln \Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}) = \ln \rho_{\underline{\bar{X}}}(\underline{\epsilon}) + o_n(1).$$

If $X_{.j}$ and $X_{.j'}$ ( $j \neq j'$ ) are independent, an upper bound of (5.1) turns out to be

$$\begin{aligned}
&\Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}) \\
&= \prod_{j=1}^{p} \Pr(\overline{X}_j - \mu_j \geq \epsilon_j) \\
&\leq \prod_{j=1}^{p} \min_{t_j}[\exp(-\epsilon_j t_j)\phi_{x_j - \mu_j}(t_j)]^n \\
&= \prod_{j=1}^{p} \rho^n_{\underline{\overline{X}}}(\epsilon_j),
\end{aligned}$$

so that

$$n^{-1} \ln \Pr(\underline{\overline{X}} - \underline{\mu} \geq \underline{\epsilon}) = \sum_{j=1}^{p} \ln \rho_{\underline{\overline{X}}}(\epsilon_j) + o_n(1).$$

Next, consider the multivariate large-deviations of sample medians, i.e.,

$$\Pr(\underline{M} - \underline{\mu} \geq \underline{\epsilon}).$$

Without loss of generality, we assume sample size $n$ is an odd number, and take the simple case, $p$=2. Then, as we already showed in Section 2.3,

$$\begin{aligned}
&\Pr(M_1 - \mu_1 \geq \epsilon_1, \quad M_2 - \mu_2 \geq \epsilon_2) \\
&= \Pr(\overline{Y}_{\epsilon_1} \geq \frac{1}{2}, \ \overline{Y}_{\epsilon_2} \geq \frac{1}{2}), \qquad (5.2)
\end{aligned}$$

where $Y_{\epsilon_1}$ is the indicator of the event $X_1 - \mu_1 \geq \epsilon_1$.

In a similar way, an upper bound of (5.2) can be obtained as follows using the moment generating function of the bivariate binomial distribution:

$$\min_{(t_1,t_2)} [\exp(-\frac{t_1}{2} - \frac{t_2}{2})\phi_{Y_{\epsilon_1},Y_{\epsilon_2}}(t_1,t_2)]^n \equiv \min [\psi_{\underline{M}}(t_1,t_2)]^n$$
$$\equiv \rho^n_{\underline{M}}(\epsilon_1,\epsilon_2). \quad (5.3)$$

The bivariate binomial probability function was found by Wishart [41], viz., the probability of $x_1$ successes of the first type ( and $n - x_1$ failures ) and of $x_2$ successes of the second type ( and $n - x_2$ failures ) is given by the following function of the two variables $x_1$ and $x_2$:

$$f(x_1,x_2) = \binom{n}{x_1} \sum_{i=0}^{\min(x_1,\, x_2)} \{\binom{n-x_1}{x_2-i}\binom{x_1}{i} p_{11}^i \, p_{10}^{x_1-i} \, p_{01}^{x_2-i} \, p_{00}^{n-x_1-x_2+i}\},$$

where

$$p_{11} = \Pr(\text{ 1st event is success 2nd event is success })$$
$$p_{10} = \Pr(\text{ 1st event is success 2nd event is failure })$$
$$p_{01} = \Pr(\text{ 1st event is failure 2nd event is success })$$
$$p_{00} = \Pr(\text{ 1st event is failure 2nd event is failure })$$

We define $p = p_{11} + p_{10}$, $p' = p_{11} + p_{01}$, $q = 1 - p$, $q' = 1 - p'$ and $p_{00} = 1 - p_{10} - p_{01} - p_{11}$. We note also that $p_{00}\, p_{11} - p_{10}\, p_{01} = p_{11} - pp'$.
Then the moment generating function can be computed as

$$p_{00} + p_{10}e^{t_1} + p_{01}e^{t_2} + p_{11}e^{t_1+t_2},$$

so that $\psi_M(t_1,t_2)$ turns out to equal

$$\exp(-\frac{t_1}{2} - \frac{t_2}{2})(p_{00} + p_{10}e^{t_1} + p_{01}e^{t_2} + p_{11}e^{t_1+t_2}).$$

In order to obtain an upper bound (5.3),

$d\psi_M(t_1, t_2)/dt_1 = 0$ and $d\psi_M(t_1, t_2)/dt_2 = 0$ imply

$$(p_{10} + p_{11}e^{t_2})e^{t_1} = p_{00} + p_{01}e^{t_2}$$

and

$$(p_{01} + p_{11}e^{t_1})e^{t_2} = p_{00} + p_{10}e^{t_1},$$

so that we get

$$\hat{t}_1 = \frac{1}{2}\ln\frac{p_{00}\ p_{01}}{p_{10}\ p_{11}} \equiv \tau_1$$

and

$$\hat{t}_2 = \frac{1}{2}\ln\frac{p_{00}\ p_{10}}{p_{01}\ p_{11}} \equiv \tau_2.$$

Then we obtain:

$$\begin{aligned} \psi_M(\tau_1, \tau_2) &= 2[\sqrt{p_{00}\ p_{11}} + \sqrt{p_{10}\ p_{01}}] \\ &\equiv \rho_{\underline{M}}(\epsilon_1, \epsilon_2). \end{aligned}$$

The upper bound (5.3) will be $\{2[\sqrt{p_{00}\ p_{11}} + \sqrt{p_{10}\ p_{01}}]\}^n$.

Therefore,

$$\begin{aligned} & n^{-1}\ln\Pr(M_1 - \mu_1 \geq \epsilon_1,\quad M_2 - \mu_2 \geq \epsilon_2) \\ &= \ln\{2[\sqrt{p_{00}\ p_{11}} + \sqrt{p_{10}\ p_{01}}]\} + o_n(1) \\ &\equiv \ln\rho_{\underline{M}}(\underline{\epsilon}) + o_n(1). \end{aligned} \tag{5.4}$$

If $Y_1$ and $Y_2$ are independent (e.g., $X_{.1}$ and $X_{.2}$ are independent), then $p_{11} = pp'$ and $p_{00} = qq'$ imply $p_{00}\ p_{11} = p_{10} - p_{01} = pqp'q'$, so that we obtain an upper bound (5.3):

$$\rho^n_{\underline{M}}(\epsilon_1, \epsilon_2) \equiv [2^2\sqrt{pqp'q'}]^n. \tag{5.5}$$

### 5.3. Asymptotic covariances

We will consider the asymptotic covariances of rounded sample means and rounded sample medians, which rely on the $p$-variate behaviors described in the previous section. However, we need only the bivariate analysis to examine the asymptotic covariances (i.e., $p = 2$).

Now we begin with the asymptotic covariance of rounded sample means. It is also helpful to consider the four quadrants of the plane, and we can write the asymptotic covariance of rounded sample means as the sum of the twenty terms similar to those found in the analysis of the RLSEs in the $p$-variate regression model defined in Chapter 4. Lemma 4.5, showed that the first term of a certain series dominates the sum of the rest of the series in each quadrant. That sort of domination argument, together with Condition C3, can be easily applied to the sample means, so that we may write

$$(2\epsilon_i)^{-1}(2\epsilon_j)^{-1}Cov_n(\overline{X}_{i,\epsilon_i},\ \overline{X}_{j,\epsilon_j})$$

$$\begin{aligned} &= 2\{\Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \geq \epsilon_j\ ) \\ &\quad - \Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \leq -\epsilon_j\ )\}(1 + o_n(1)). \end{aligned}$$

Similar arguments, appearing in the proof of Theorem 4.6 about the asymptotic covariance of RLSEs, allow the assertion that:

$$sgn(corr_{ij})(2\epsilon_i)^{-1}(2\epsilon_j)^{-1}Cov_n(\overline{X}_{i,\epsilon_i}, \overline{X}_{j,\epsilon_j})$$

$$= 4P_{n,\epsilon_i,\epsilon_j}(\underline{\overline{X}})(1 + o_n(1)),$$

where

$$\begin{aligned} &P_{n,\epsilon_i,\epsilon_j}(\underline{\overline{X}}) \\ &= \max\{\Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \geq \epsilon_j\ ), \\ &\qquad \Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \leq -\epsilon_j\ )\}, \end{aligned}$$

$corr_{i,j}$ is the correlation between $b_i$ and $b_j$, and

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

As we already discussed equation (4.10) in Section 4.3, it is true that

$$n^{-1} \ln P_{n,\epsilon_i,\epsilon_j}(\overline{X}) = \ln \rho_{\underline{X}}(\epsilon_i, \epsilon_j) + o_n(1),$$

which leads to

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Cov_n(\overline{X}_{i,\epsilon_i}, \overline{X}_{j,\epsilon_j}) = \ln \rho_{\underline{X}}(\epsilon_i, \epsilon_j).$$

If $X_{.j}$ and $X_{.j'}$ $(j \neq j')$ are independent, then $\Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \geq \epsilon_j\ )$ is equal to $\Pr(\ \overline{X}_i - \mu_i \geq \epsilon_i,\ \ \overline{X}_j - \mu_j \leq -\epsilon_j\ )$, so that $Cov_n(\overline{X}_{i,\epsilon_i}, \overline{X}_{j,\epsilon_j})$ equals zero.

Next, consider the asymptotic covariances of rounded sample medians. With arguments similar to those yielding the asymptotic covariances of rounded sample means, one may, considering the four quadrants of the plane, express the asymptotic covariance of rounded sample median as the sum of twenty terms. Then Lemma 4.5 can be applied to sample medians with Condition C3, to obtain

$$(2\epsilon_i)^{-1}(2\epsilon_j)^{-1}Cov_n(M_{i,\epsilon_i},\ M_{j,\epsilon_j})$$

$$\begin{aligned} &= 2\{\Pr(\ M_i - \mu_i \geq \epsilon_i,\ \ M_j - \mu_j \geq \epsilon_j\ ) \\ &\quad - \Pr(\ M_i - \mu_i \geq \epsilon_i,\ \ M_j - \mu_j \leq -\epsilon_j\ )\}(1 + o_n(1)), \end{aligned}$$

so that

$$\begin{aligned} &sgn(corr_{ij})(2\epsilon_i)^{-1}(2\epsilon_j)^{-1}Cov_n(M_{i,\epsilon_i},\ M_{j,\epsilon_j}) \\ &\qquad = 4P_{n,\epsilon_i,\epsilon_j}(\underline{M})\exp(1 + o_n(1)), \end{aligned}$$

where

$$\begin{aligned} &P_{n,\epsilon_i,\epsilon_j}(\underline{M}) \\ &\quad = \max\{\Pr(\ M_i - \mu_i \geq \epsilon_i,\ \ M_j - \mu_j \geq \epsilon_j\ ), \\ &\qquad\qquad \Pr(\ M_i - \mu_i \geq \epsilon_i,\ \ M_j - \mu_j \leq -\epsilon_j\ )\}. \end{aligned}$$

Hence we conclude that

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij})Cov_n(M_{i,\epsilon_i},\ M_{j,\epsilon_j}) = \ln \rho_{\underline{M}}(\epsilon_i, \epsilon_j),$$

since $n^{-1} \ln P_{n,\epsilon_i,\epsilon_j}(M) = \ln \rho_{\underline{M}}(\epsilon_j, \epsilon_j) + o_n(1)$.

Chapter 3 examined the asymptotic variances of the rounded sample mean and the rounded sample median in the univariate Normal and Laplace cases. Here we shall treat only the multivariate Normal case.

First of all, we need to find $\rho_{\bar{\underline{X}}}$ in the $p$-variate Normal case. Recall expression (4.5) from Chapter 4:

$$\psi_{\bar{\underline{X}}}(\underline{t}) = \exp(-\underline{\epsilon}'\underline{t} + \frac{1}{2}\underline{t}'\underline{V}\underline{t}),$$

where $\underline{V}$ is the variance and covariance matrix of $\sqrt{n}(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p)'$.
Now, $d\psi_{\bar{\underline{X}}}(\underline{t})/d\underline{t} = \underline{0}$ gives $\underline{\tau} = \underline{V}^{-1}\underline{\epsilon}$, so that we are led to equation (4.12)

$$\psi_{\bar{X}}(\underline{\tau}) = \exp(-\frac{1}{2}\underline{\epsilon}'\underline{V}'\underline{\epsilon}),$$

and thence to equation (4.13)

$$n^{-1}\ln \rho_{\bar{\underline{X}}}(\underline{\epsilon}) = -\frac{1}{2}\underline{\epsilon}'\underline{V}^{-1}\underline{\epsilon}. \tag{5.6}$$

More specifically, when $p = 2$, the RHS of (5.6) turns out to equal

$$-\frac{1}{2}\frac{\sigma_j^2\epsilon_i^2 - 2\sigma_{ij}\epsilon_i\epsilon_j + \sigma_i^2\epsilon_j^2}{\sigma_i^2\sigma_j^2 - \sigma_{ij}^2},$$

which implies that the asymptotic covariance of rounded sample means in the $p$-variate Normal distribution, satisfies

$$\lim_{n\to\infty} n^{-1}\ln sgn(corr_{ij})Covn(\overline{X}_{i,\epsilon_i}, \overline{X}_{j,\epsilon_j})$$

$$= -\frac{1}{2}\frac{\sigma_j^2\epsilon_i^2 - 2\sigma_{ij}\epsilon_i\epsilon_j + \sigma_i^2\epsilon_j^2}{\sigma_i^2\sigma_j^2 - \sigma_{ij}^2}.$$

Secondly, we discussed the asymptotic covariance of rounded sample median in the Normal case. We can define, in equation (5.5), $p \equiv \Pr(X_i - \mu_i \geq \epsilon_i) = 1 - \Phi(\epsilon_i/\sigma_i)$, $p' \equiv \Pr(X_j - \mu_j \geq \epsilon_j) = 1 - \Phi(\epsilon_j/\sigma_j)$, $q = 1 - p$, and $q' = 1 - p'$. Also assume $\sigma_i = \sigma_j = 1$, then

$$\begin{aligned} p_{00} &= \Pr(X_i - \mu_i < \epsilon_i,\ \ X_j - \mu_j < \epsilon_j) = \Phi(\epsilon_i, \epsilon_j), \\ p_{01} &= \Pr(X_i - \mu_i < \epsilon_i,\ \ X_j - \mu_j \geq \epsilon_j) = \Phi(\epsilon_i) - \Phi(\epsilon_i, \epsilon_j), \\ p_{10} &= \Pr(X_i - \mu_i \geq \epsilon_i,\ \ X_j - \mu_j < \epsilon_j) = \Phi(\epsilon_j) - \Phi(\epsilon_i, \epsilon_j), \\ p_{11} &= \Pr(X_i - \mu_i \geq \epsilon_i,\ \ X_j - \mu_j \geq \epsilon_j) = 1 - \Phi(\epsilon_i) - \Phi(\epsilon_j) + \Phi(\epsilon_i, \epsilon_j). \end{aligned}$$

Note that

$$\begin{aligned} p_{00}\, p_{11} &= \Phi(\epsilon_i, \epsilon_j)[1 - \Phi(\epsilon_i) - \Phi(\epsilon_j) + \Phi(\epsilon_i, \epsilon_j)], \\ p_{01}\, p_{10} &= [\Phi(\epsilon_i) - \Phi(\epsilon_i, \epsilon_j)][\Phi(\epsilon_j) - \Phi(\epsilon_i, \epsilon_j)]. \end{aligned}$$

And note that, if $X_{.j}$ and $X_{.j'}$ are independent (i.e., $p_{11} = pp'$),

$$p_{00}\, p_{11} = p_{01}\, p_{10} = \Phi(\epsilon_i)[1 - \Phi(\epsilon_i)]\Phi(\epsilon_j)[1 - \Phi(\epsilon_j)].$$

In any event, (5.4) yields

$$\begin{aligned} &\rho_{\underline{M}}(\epsilon_i, \epsilon_j) \\ &= 2\{\sqrt{\Phi(\epsilon_i, \epsilon_j)[1 - \Phi(\epsilon_i) - \Phi(\epsilon_j) + \Phi(\epsilon_i, \epsilon_j)]} \\ &\quad + \sqrt{[\Phi(\epsilon_i) - \Phi(\epsilon_i, \epsilon_j)][\Phi(\epsilon_j) - \Phi(\epsilon_i, \epsilon_j)]}\}. \end{aligned}$$

Hence we conclude that the asymptotic covariance of rounded sample medians, in the Normal case, satisfies

$$\lim_{n\to\infty} n^{-1} \ln sgn(corr_{ij}) Covn(M_{i,\epsilon_i}, M_{j,\epsilon_j})$$

$$= \ln 2\{\sqrt{\Phi(\epsilon_i,\epsilon_j)[1-\Phi(\epsilon_i)-\Phi(\epsilon_j)+\Phi(\epsilon_i,\epsilon_j)]}$$
$$+ \sqrt{[\Phi(\epsilon_i)-\Phi(\epsilon_i,\epsilon_j)][\Phi(\epsilon_j)-\Phi(\epsilon_i,\epsilon_j)]}\}$$

## 5.4. Joint Asymptotic Efficiency

After obtaining the asymptotic variance and covariance matrix of rounded sample means and rounded sample medians from the symmetric TEBIRF distribution, we might consider the issue of efficiency. As we did in Section 4.5, the joint asymptotic efficiency (JAE) will be considered, based on the determinant of the asymptotic variance-covariance matrix, say $|\underline{V}^a|$.

Let us begin with the rounded sample means. In Section 4.5, it turned out that, as sample size increases, the asymptotic covariances will be negligible and so the determinant will equal the product of the asymptotic variances. Hence

$$|\underline{V}^a(\underline{\bar{X}})| = \prod_{j=1}^{p} Varn(\bar{X}_{j,\epsilon_j}) + o_n(1)$$
$$= \prod_{j=1}^{p} \exp[n\rho_{\bar{X}}(\epsilon_j)] + o_n(1),$$

so that

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{\bar{X}})| = \sum_{j=1}^{p} \rho_{\bar{X}}(\epsilon_j).$$

Here $\rho_{\bar{X}}(\cdot)$ is analogous to $-\theta_{\bar{X}}(\cdot)$ defined in (2.39).

Similarly, regarding the rounded sample medians,

$$\begin{aligned} |\underline{V}^a(\underline{M})| &= \prod_{j=1}^{p} Var_n(M_{j,\epsilon_j}) + o_n(1) \\ &= \prod_{j=1}^{p} \exp[n\rho_M(\epsilon_j)] + o_n(1), \end{aligned}$$

so that

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}^a(\underline{M})| = \sum_{j=1}^{p} \rho_M(\epsilon_j).$$

Here again $\rho_M(\cdot)$ is analogous to $-\theta_M(\cdot)$ defined in (2.21).

The asymptotic covariances formulated in the previous section are negligible when the JAE is considered. With the results obtained above, the JAEs of the rounded sample means and the rounded sample medians for the Normal distribution, and also for the Laplace distribution, can be given as follows:

1. The JAE of the rounded sample means in the Normal distribution.

Since $\rho_{\bar{X}}^N(\epsilon_j) = -(\epsilon_j^2/2\sigma_j^2)$,

$$\lim_{n\to\infty} n^{-1} \ln |\underline{V}_N^a(\bar{X})| = -\sum_{j=1}^{p} \frac{\epsilon_j^2}{2\sigma_j^2}.$$

2. The JAE of the rounded sample means in the Laplace distribution.

Since

$$\rho_{\bar{X}}^{L}(\epsilon_j) = \{\frac{\epsilon_j^2 \exp[\frac{\sigma_j^2 - \sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2}}{\sigma_j^2}]}{[\sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2} - \sigma_j^4]}\},$$

$$\begin{aligned}
\lim_{n\to\infty} n^{-1} \ln |\underline{V}_L^a(\bar{X})| &= \sum_{j=1}^{p} \ln \frac{\epsilon_j^2 \exp[\frac{\sigma_j^2 - \sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2}}{\sigma_j^2}]}{[\sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2} - \sigma_j^4]} \\
&= 2\sum_{j=1}^{p} \ln \epsilon_j \\
&\quad + \sum_{j=1}^{p} [\frac{\sigma_j^2 - \sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2}}{\sigma_j^2}] \\
&\quad - \sum_{j=1}^{p} \ln[\sqrt{\sigma_j^4 + \epsilon_j^2\sigma_j^2} - \sigma_j^4].
\end{aligned}$$

3. The JAE of the rounded sample medians in the Normal distribution.

Since $\rho_M^N(\epsilon_j) = 2\sqrt{\Phi(\epsilon_j/\sigma_j)[1-\Phi(\epsilon_j/\sigma_j)]}$,

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\ln|\underline{V}_N^a(M)| &= \sum_{j=1}^{p} \ln 2\sqrt{\Phi(\epsilon_j/\sigma_j)[1-\Phi(\epsilon_j/\sigma_j)]} \\ &= \sum_{j=1}^{p} \ln 2 \\ &\quad + \frac{1}{2}\sum_{j=1}^{p} \ln\{\Phi(\epsilon_j/\sigma_j)[1-\Phi(\epsilon_j/\sigma_j)]\}.\end{aligned}$$

4. The JAE of the rounded sample medians in the Laplace distribution.

Since $\rho_M^L(\epsilon_j) = \sqrt{[e^{-\epsilon_j/\sigma_j}(2-e^{-\epsilon_j/\sigma_j})]}$,

$$\begin{aligned}\lim_{n\to\infty} n^{-1}\ln|\underline{V}_L^a(M)| &= -\sum_{j=1}^{p} \frac{\epsilon_j}{2\sigma_j} \\ &\quad + \frac{1}{2}\sum_{j=1}^{p} \ln(2-e^{-\epsilon_j/\sigma_j}).\end{aligned}$$

## 6. OTHER ROUNDINGS

Three kinds of rounding were mentioned in Section 1.1. Chapter 2 to Chapter 5 considered rounding sample estimates which is the first case, i.e., compute $X_i$ exactly and then round the sample estimates.

Now we may consider the second case; in other words, round $X_i$'s, say $X_{i\epsilon}$, and compute the sample mean. Kendall and Stuart [31] mention "Sheppard's corrections for grouping" with $h = 2\epsilon$, and find

$$Var(X_{i\epsilon}) \doteq Var(X_i) + \frac{1}{3}\epsilon^2,$$

so that the variance of the sample mean, $\overline{(X\epsilon)}$, computed from the rounded individual values, will be:

$$Varn(\overline{(X_\epsilon)}) \doteq Varn(\overline{X}) + \frac{1}{3n}\epsilon^2.$$

In other words, for all $n$ and positive $\epsilon$, one has that, approximately,

$$Varn(\overline{(X_\epsilon)}) > Varn(\overline{X}). \qquad (6.1)$$

But Sheppard's corrections break down when certain conditions on "high-order contact" are violated. For example, Sheppard's corrections fail in the Normal case.

We might consider another way to think about this kind of rounding. Define

$$X_\epsilon = X + X_\epsilon - X = X + \delta,$$

where $\delta = X_\epsilon - X$ is the "round-off quantity." Then

$$Var(X_\epsilon) = Var(X) + Var(\delta) + 2COV(X, \delta).$$

Since $Var(\delta) = \frac{\epsilon^2}{3}$,

$$(\sqrt{Var(X)} - \frac{\epsilon}{\sqrt{3}})^2 \leq Var(X_\epsilon) \leq (\sqrt{Var(X)} + \frac{\epsilon}{\sqrt{3}})^2.$$

Hence we obtain

$$\frac{1}{n}(\sqrt{Var(X)} - \frac{\epsilon}{\sqrt{3}})^2 \leq Var_n(\overline{(X_\epsilon)}) \leq \frac{1}{n}(\sqrt{Var(X)} + \frac{\epsilon}{\sqrt{3}})^2.$$

Whereas Sheppard's corrections tell us that $Var_n(\overline{(X_\epsilon)})$ is always larger than $Var_n(\overline{X})$, (6.1) is not always true, but depends on the size of $\epsilon$. This raises the problem of the nature of the correlation between the random variable $X$ and the round-off quantity $\delta$ which is unknown.

We might also consider a third kind of rounding, that is, rounding the sample estimates after rounding individual $X_i$'s. For example, the case of the sample mean will be examined.

Let $X_i$ be an *iid* random sample from $N(\mu, \sigma^2)$. Let $\overline{(X_{\epsilon_1})}_{\epsilon_2}$ be the rounded sample mean, restricted to grid values with a rounding interval $2\epsilon_2$, which is the average of the individual values restricted to grid values corresponding to another rounding interval $2\epsilon_1$. Then it is possible to derive

$$Var_n(\overline{(X_{\epsilon_1})}_{\epsilon_2}) = (2\epsilon_2)^2 \Pr(|\overline{(X_{\epsilon_1})} - \mu| \geq \epsilon_2) \exp(1 + o_n(1)).$$

With a condition analogous to Condition C1, it seems plausible that one can obtain the following:

$$\begin{aligned} n^{-1}\ln \Pr(|\overline{(X_{\epsilon_1})} - \mu| \geq \epsilon_2) &= \ln\{\min_t [e^{-\epsilon_2 t} \sum_m e^{2m\epsilon_1 t} \int_{(2m-1)\epsilon_1}^{(2m+1)\epsilon_1} dN(\mu,\ \sigma^2)]\} \\ &\equiv \ln\{\min_t e^{-\epsilon_2 t} \phi_{\epsilon_1}(t)\} \\ &\equiv \ln \rho(\epsilon_1, \epsilon_2), \end{aligned}$$

so that it is likely to be true that

$$n^{-1} \ln Var_n(\overline{(X_{\epsilon_1})}_{\epsilon_2}) = \ln \rho(\epsilon_1, \epsilon_2) + o_n(1).$$

$\rho(\epsilon_1, \epsilon_2)$ would not be easy to find in such a situation, but we can say that $e^{-\epsilon_2 t}\phi_{\epsilon_1}(t)$ is a convex function of $t$. A general method, possibly useful in this case, for finding the minimum of a convex function can be based on the following lemma:

**Lemma 6.1** *For a real number $a$ and a positive $\epsilon$, let $f(x)$ be a convex function on $[a-\epsilon,\ a+\epsilon]$ with $f(a-\epsilon) > f(a)$ and $f(a+\epsilon) > f(a)$. Let $f^-$ be the straight line through $f(a-\epsilon)$ and $f(a)$, and $f^+$ be the straight line through $f(a+\epsilon)$ and $f(a)$. Then $f^-(a+\epsilon) = 2f(a) - f(a-\epsilon)$ and $f^+(a-\epsilon) = 2f(a) - f(a+\epsilon)$. Hence the minimum of $f(x)$ exists on $(a-\epsilon,\ a+\epsilon)$, and $f(x) \geq \min\{f^-(a+\epsilon),\ f^+(a-\epsilon)\}$.*

Proof:

Since $f(\cdot)$ is a convex function,

$$f'(a) > \frac{f(a) - f(a-\epsilon)}{\epsilon}$$

and

$$f'(a) < \frac{f(a+\epsilon) - f(a)}{\epsilon}.$$

We may define, respectively,

$$f^{-}(x) = \frac{f(a) - f(a-\epsilon)}{\epsilon}(x-a) + f(a)$$

and

$$f^{+}(x) = \frac{f(a+\epsilon) - f(a)}{\epsilon}(x-a) + f(a),$$

so that

$$f^{-}(a+\epsilon) = 2f(a) - f(a-\epsilon)$$

and

$$f^{+}(a-\epsilon) = 2f(a) - f(a+\epsilon).$$

The minimum value of $f(x)$ exists on $(a-\epsilon,\ a+\epsilon)$ by its convexity. So for $x$ belonging to $(a-\epsilon,\ a)$,

$$\begin{aligned} f(x) &- f^{+}(x) \\ &= [\frac{f(x) - f(a)}{x-a} - \frac{f(a+\epsilon) - f(a)}{\epsilon}](x-a) \\ &\geq 0, \end{aligned}$$

and for $x$ belonging to $[a,\ a+\epsilon)$,

$$\begin{aligned} f(x) &- f^{-}(x) \\ &= [\frac{f(x)-f(a)}{x-a} - \frac{f(a)-f(a-\epsilon)}{\epsilon}](x-a) \\ &\geq 0, \end{aligned}$$

so that

$$f(x) \geq \min\{f^{-}(a+\epsilon),\ \ f^{+}(a-\epsilon)\},$$

which completes the proof of Lemma 6.1.

# 7. BIBLIOGRAPHY

[1] Bahadur, R. R. (1967). "Rates of convergence of estimates and test statistic." Ann. Math. Statist., 38, 303-324.

[2] ———. (1960). "On the asymptotic efficiency of tests and estimates." Sankhyā, 22, 229-252.

[3] ———. (1971). *Some Limit Theorems in Statistics*, Regional Conference Series in Applied Mathematics, 4. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

[4] Bahadur, R. R. and Rao, R. R. (1960). "On deviations of the sample mean." Ann. Math. Statist., 31, 1015-1027.

[5] Bahadur, R. R. and Zabell, S. L. (1979). "Large deviations of the sample mean in general vector spaces." Ann. Prob., 7(4), 587-621.

[6] Barrodale, I. (1968). "$L_1$ approximation and the analysis of data." Applied Statist., 17, 51-57.

[7] ———. (1973). "An improved algorithm for discrete $L_1$ linear approximation." Society for Industrial and Applied Mathematics, J. Numer. Anal., 10(5), 839-848.

[8] Barrodale, I. and Young, A. (1966). "Algorithms for best $L_1$ and $L_\infty$ linear approximations on a discrete set." Numerische Mathematik, 8, 295-306.

[9] Bártfai, P. (1978). "Large deviations of the sample mean in Euclidean spaces." Stat. Dept., Division of Mathematical Sciences, Mimeograph Series No. 78-13, Purdue University.

[10] Bassett, G., Jr. and Koenker, R. (1978). "Asymptotic theory of least absolute error regression." J. Am. Statist. Assoc., 73(363), 618-622.

[11] Berkson, J. (1950). "Are there two regressions?" J. Am. Statist. Assoc., 45, 164-180.

[12] Berman, S. M. (1962). "A law of large numbers for the maximum in a stationary Gausian sequence." Ann. Math. Statist., 33, 93-97.

[13] ________. (1964). "Limiting distribution of the maximum of a diffussion process." Ann. Math. Statist., 35, 319-329.

[14] Billingsley, P. (1979). *Probability and Measure.* Wiley, New York.

[15] Blackwell, D. and Hodges, J. L., Jr. (1959). "The probability in the extreme tail of a convolution." Ann. Math. Statist., 30, 1113-1120.

[16] Chapman, D. G. and Robbins, H. (1951). "Minimum variance estimation without regularity assumptions." Ann. Math. Statist., 22, 581-586.

[17] Chernoff, H. (1952). "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations." Ann. Math. Statist., 23,

493-507.

[18] Chung, K. L. (1974). *A Course in Probability Theory.* 2nd ed. Academic Press, New York.

[19] David, H. A. (1980). *Order Statistics.* 2nd ed. Wiley, New York.

[20] David, H. A. and Mishriky, R. S. (1968). "Order statistics for discrete populations and for grouped samples." J. Am. Statist. Assoc., 63, 1390-1398.

[21] Ghosh, M. (1974). "Admissibility and minimaxity of some maximum likelihood estimators when the parameter space is restricted to integers." J. Roy. Statist. Soc., B, 37, 264-271.

[22] Ghosh, M. and Meeden, G. (1978). "Admissibility of the MLE of the Normal integer mean." Sankhyā, B, 40, Pts. 1 and 2, 1-10.

[23] Giesbrecht, F. and Kempthorne, O. (1976). "Maximum likelihood estimation in the three-parameter Lognormal distribution." J. Roy. Statist. Soc., B, 38, 257-264.

[24] Groeneboom, P. and Shorack, G. R. (1981). "Large deviations of goodness of fit statistics and linear combinations of order statistics." Ann. Prob., 9(6), 971-987.

[25] Groeneboom, P., Oosterhoff, J. and Ruymgaart, F. H. (1979). "Large deviation theorems for empirical probability measures." Ann. Prob., 7(4), 553-586.

[26] Hammersley, J. M. (1950). "On estimating restricted parameters (with discussion). " J. Roy. Statist. Soc., B, 12(2), 192-240.

[27] Hodges, J. L. and Lehmann, E. L. (1956). "The efficiency of some non-parametric competitors of the $t$ test." Ann. Math. Statist., 27, 324-335.

[28] Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions, Vol.1 and Vol.2: Continuous Multivariate Distributions.* Houghton Mifflin, Boston.

[29] Karst, O. J. (1958). "Linear curve fitting using least deviations." J. Am. Statist. Assoc., 53, 118-132.

[30] Kempthorne, O. (1966). "Some aspects of experimental inference." J. Am. Statist. Assoc., 61, 11-34.

[31] Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics.* Vol.1 and Vol.2., 3rd ed. Griffin, London.

[32] Khan, R. A. (1973). "On some properties of Hammersley's estimator of an integer mean." Ann. Statist., 1(4), 756-762.

[33] ———. (1978). "A note on the admissibility of Hammersley's estimator of an integer mean." Can. J. Statist., 6(1), 113-119.

[34] Lambert, J. A. (1969). "Estimation of parameters in the four-parameter Lognormal distribution." Aust. J. Statist., 12(1), 33-43.

[35] ———. (1972). "Likelihoods with rounded data." Aust. J. Statist., 14(3), 204-210.

[36] Lehmann, E. L. (1986). *Testing Statistical Hypotheses.* 2nd ed. Wiley, New York.

[37] Rosenberg, B. and Carlson, D. (1977). "A simple approximation of the sampling distribution of least absolute residuals regression estimates." Comm. Statist., Sim. Comp., B, 6(4), 421-437.

[38] Sacks, J. (1975). "An asymptotically efficient sequence of estimators of a location parameter." Ann. Statist., 3(2), 285-298.

[39] Sankoorikal, J. T., Danofsky, R. A., David, H. T., Hendrickson, R. A. and Tollefson, M. (1985). "Statistical estimation of vibration parameters of nuclear core component." Trans. Am. Nucl. Soc., 49, 428-433.

[40] Schwarz, G. (1968). "Asymptotic shapes for the sequential testing of truncation parameters." Ann. Math. Statist., 39(6), 2038-2043.

[41] Wishart, J. (1949). "Cumulants of multivariate multinomial distributions." Biometrika, 36, 47-58.

# 8. ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my major professor Dr. H. T. David to whom I am greatly indebted. Dr. H. T. David, who is the most warm-hearted and hard-working scholar I have ever met, suggested the topic of this dissertation and he has been a constant source of guidance. To make a long story short, I have a dream to be like Dr. H. T. David.

Also I would like to thank the committee members, professor W. Q. Meeker, J. C. Kegley, Oscar Kempthorne, Jerome Sacks and Vince Sposito for helping me complete this dissertation. Especially, I have benefited from the useful comments of Oscar Kempthorne, Glen Meeden and Vince Sposito.

I have benefited greatly from all the professors in the Department of Statistics, Iowa State University, for educating and letting me complete the Ph.D. program here without any trouble. I had been a teaching assistant in the Department of Statistics from 1985 to 1986 and, from that time on, I have been a research assistant of Dr. Jerome Sacks who is a collaborator in USDA. I thank them for giving me an opportunity to work and, especially, thank Dr. Sacks for teaching me various data analysis techniques.

I give my appreciation to professors K. S. Chung, M. Y. Huh, P. Namkoong and K. W. Oh, at the Department of Statistics, Sung Kyun Kwan University in Seoul, Korea, for recommending me to study abroad, especially at Iowa State University, and constant encouragement.

Special thanks and gratitude should go to my parents, Sa Mock Hong and J. S. Chung for their constant encouragement and moral support.

Okhee, my wife, who has been always a good advisor and my fellow traveler, is expecting a baby right now, and also is trying to complete her study. She did not complain much even though I was too busy to help her, and did endure her pain well with her busy school work. I thank her for it and love her very much. I hope that she and the baby are healthy and strong.