A NOTE ON MINIMUM VARIANCE IN UNBIASED ESTIMATION

By SUJIT KUMAR MITRA

Indian Statistical Institute, Calcutta

1. Introduction

In a preceding publication Basu and Mitra (1952) have applied a useful technique due to Bhattacharyya (1947), Hammersley (1950), and Robbins and Chapman (1951) in finding the minimum attainable variance for unbiased estimators for the binomial population. A similar technique is applied here to investigate the lower limits of variance in two broad classes of problems :

(A) non-regular estimation of the continuous type for which the Cramér-Rao limit is not calculable, and

(B) cases of regular estimation of the continuous type, where the Cramér-Rao limit can be computed but the lower limit is not attained for the best unbiased estimate.

For the latter case a sufficient condition is stated, under which Hammersley's technique is expected to improve the situation. In case (A), however, though the actual limit determined is not very satisfactory, it is important inasmuch as it gives a valuable indication of the order of variances attainable in large samples by unbiased estimators in such cases.

2. Non-regular estimation of the continuous type

We shall consider two situations under case (A) as defined above.

(i) Consider first a distribution of the continuous type with the frequency function $f(x,\theta)$ where

$$f(x,\theta) = f(x)/g(\theta) \text{ for } 0 < x \leqslant \theta,$$
$$= 0 \text{ otherwise.}$$

Suppose $(x_1, x_2, \ldots, x_n)$ are n independent observations on the chance variable x and that $t(x_1, x_2, \ldots, x_n) \equiv t(\underset{\sim}{x})$ is an unbiased estimate for θ.

Let A stand for the n-dimensional Euclidean region defined by

$$0 < x_i \leqslant \theta - \delta \qquad (i = 1, 2, \ldots, n)$$

and $A+B$ for the region:

$$0 \leqslant x_i < \theta \qquad (i = 1, 2, \ldots, n).$$

We have
$$\left.\begin{aligned} \int_{A+B} t(\underline{x}) \prod_{i=1}^{n} f(x_i,\, \theta)\, dv &= \theta, \\ \int_{A} t(\underline{x}) \prod_{i=1}^{n} f(x_i,\, \theta-\delta)\, dv &= \theta-\delta, \end{aligned}\right\} \quad \ldots \quad (2.1)$$

where $dv = dx_1 dx_2 \ldots dx_n$.

Or, subtracting

$$\int_{A+B} \{ t(\underline{x})-\theta \} \left[\frac{\left\{ \prod_{i=1}^{n} f(x_i) \right\}}{g^n(\theta)} \right]^{\frac{1}{2}} \frac{\left[\frac{1}{g^n(\theta)} - \frac{F(\underline{x})}{g^n(\theta-\delta)} \right]}{\left[\frac{\prod_{i=1}^{n} f(x_i)}{g^n(\theta)} \right]^{\frac{1}{2}}} \prod_{i=1}^{n} f(x_i)\, dv = \delta \quad \ldots \quad (2.2)$$

where $F(\underline{x})$ is the characteristic function of the set A i.e. $F(\underline{x}) = 1$ whenever $x \epsilon A$, and $= 0$ otherwise.

Using Cauchy-Schwartz's inequality we get

$$\text{Var}_\theta\, t \geqslant \frac{\delta^2}{\int_{A+B} g^n(\theta) \left[\frac{1}{g^n(\theta)} - \frac{F(\underline{x})}{g^n(\theta-\delta)} \right]^2 \prod_{1}^{n} f(x_i)\, dv} \quad \ldots \quad (2.3)$$

The denominator of (2.3) can be written as

$$\int_{A+B} \frac{\prod_{i=1}^{n} f(x_i) dv}{\{g^n(\theta)\}} - 2 \int_{A} \frac{\prod_{i=1}^{n} f(x_i) dv}{\{g^n(\theta-\delta)\}} + \int_{A} \frac{\{g^n(\theta)\}}{\{g^n(\theta-\delta)\}^2} \prod_{i=1}^{n} f(x_i)\, dv. \quad \ldots \quad (2.4)$$

Since $\int_0^\theta f(x_i) dx = g(\theta)$, (2.4) simplifies to

$$\left[\frac{g(\theta)}{g(\theta-\delta)} \right]^n - 1 \quad \ldots \quad (2.5)$$

Therefore,
$$\text{Var}_\theta\, t \geqslant \underset{\delta}{\text{Sup}} \frac{\delta^2}{\left[\frac{g(\theta)}{g(\theta-\delta)} \right]^n - 1} \quad \ldots \quad (2.6)$$

Let us substitute $\delta h(\theta) = \frac{\alpha}{n} + o\left(\frac{1}{n}\right)$ in (2.6), where $h(\theta) = g'(\theta)/g(\theta)$.

We get $$\text{Var}_\theta t \geqslant \underset{\alpha}{\text{Sup}} \frac{1}{n^2h^2(\theta)} \frac{\alpha^2}{\left[1-\frac{\alpha}{n}+o\left(\frac{1}{n}\right)\right]^n-1}. \qquad \ldots \quad (2.7)$$

For large values of n the r.h.s. of (2.7) reduces to

$$\underset{\alpha}{\text{Sup}} \frac{1}{n^2h^2(\theta)} \cdot \frac{\alpha^2}{e^\alpha-1}. \qquad \ldots \quad (2.8)$$

Differentiating the logarithm of (2.8) w.r.t. α we get the equation

$$\frac{2}{\alpha} - \frac{e^\alpha}{e^\alpha-1} = 0,$$

or, $$p(\alpha) \equiv \alpha-2(1-e^{-\alpha}) = 0 \text{ if } \alpha \neq 0. \qquad \ldots \quad (2.9)$$

Let α_0 be the solution of $e^\alpha = 2$. Then it can be shown that $p(x)$ diminishes uniformly from $\alpha = 0$ (at which however $\alpha^2/(e^\alpha-1) = 0$) to $\alpha = \alpha_0$ and then increases uniformly for all positive values of α. Thus $p(\alpha_0) < 0$ which indicates that (2.9) admits a unique non-zero solution which is incidentally $> \alpha_0$.

Solving (2.9) by iteration we get the solution $\alpha = 1.5936$.

Threfore, $$\text{Var}_\theta t \geqslant \frac{0.6476}{n^2} \frac{1}{h^2(\theta)}, \text{ asymptotically.} \qquad \ldots \quad (2.10)$$

(ii) Next let us consider the following problem:

Suppose the chance variable x has a uniform distribution over the range θ to 2θ. As before let $(x_1, x_2, \ldots, x_n)$ constitute a random sample from the above population.

Let $A+C$ be the n-dimensional Euclidean region defined by

$$\theta+\delta \leqslant x_i < 2(\theta+\delta), \ (i = 1, 2, \ldots, n)$$

and $B+C$ the region

$$\theta-\delta \leqslant x_i \leqslant 2(\theta-\delta), \ (i = 1, 2, \ldots, n)$$

and similarly C: $$\theta+\delta \leqslant x_i \leqslant 2(\theta-\delta), \ (i = 1, 2, \ldots, n).$$

Let $t(x_1, x_2, \ldots, x_n) \equiv t(\underline{x})$ be an unbiased estimator for θ. We have

$$\int_{A+C} t(\underline{x}) \frac{dv}{(\theta+\delta)^n} = \theta+\delta,$$

$$\int_{B+C} t(\underline{x}) \frac{dv}{(\theta-\delta)^n} = \theta-\delta;$$

or,

$$\int_{A+B+C} \{t(\underline{x})-\theta\} \left\{ \frac{F_1(\underline{x})}{(\theta+\delta)^n} - \frac{F_2(\underline{x})}{(\theta-\delta)^n} \right\} dv = 2\delta,$$

where $F_1(\underline{x})$ and $F_2(\underline{x})$ are the characteristic functions of the regions $A+C$ and $B+C$ respectively; that is, in symbols,

$$\int_{A+B+C} \{t(\underline{x})-\theta\} \sqrt{\frac{F_1(\underline{x})}{(\theta+\delta)^n} + \frac{F_2(\underline{x})}{(\theta-\delta)^n}} \cdot \frac{\left\{ \frac{F_1(\underline{x})}{(\theta+\delta)^n} - \frac{F_2(\underline{x})}{(\theta-\delta)^n} \right\}}{\sqrt{\frac{F_1(\underline{x})}{(\theta+\delta)^n} + \frac{F_2(\underline{x})}{(\theta-\delta)^n}}} dv = 2\delta. \quad \ldots \quad (2.11)$$

By Cauchy-Schwartz's inequality we get

$$E_{\theta+\delta}(t-\theta)^2 + E_{\theta-\delta}(t-\theta)^2 \geqslant \frac{4\delta^2}{\int_{A+B+C} \frac{\{F_1(\underline{x})/(\theta+\delta)^n - F_2(\underline{x})/(\theta-\delta)^n\}^2}{\{F_1(\underline{x})/(\theta+\delta)^n + F_2(\underline{x})/(\theta-\delta)^n\}} dv}. \quad \ldots \quad (2.12)$$

The denominator of (2.12) is given by

$$\int_{A+B+C} \left\{ \frac{F_1(\underline{x})}{(\theta+\delta)^n} + \frac{F_2(\underline{x})}{(\theta-\delta)^n} \right\} dv - 4 \int_{A+B+C} \frac{F_1(\underline{x})F_2(\underline{x})}{\{F_1(\underline{x})(\theta-\delta)^n + F_2(\underline{x})(\theta+\delta)^n\}} dv.$$

$$= 2 - \frac{4}{\{(\theta-\delta)^n + (\theta+\delta)^n\}} \int_C dv.$$

$$= 2 - \frac{4(\theta-3\delta)^n}{\{(\theta-\delta)^n + (\theta+\delta)^n\}} . \quad \ldots \quad (2.13)$$

Therefore, $$\mathrm{Var}_{\theta+\delta}\, t + \mathrm{Var}_{\theta-\delta}\, t \geqslant \sup_{\delta} 2\delta^2 \left\{ \frac{1}{1 - \dfrac{2(\theta-3\delta)^n}{(\theta-\delta)^n+(\theta+\delta)^n}} - 1 \right\}$$

$$= \sup_{\delta} \frac{4\delta^2(\theta-3\delta)^n}{(\theta-\delta)^n+(\theta+\delta)^n-2(\theta-3\delta)^n}. \qquad \dots \quad (2.14)$$

As before if we put $\dfrac{\delta}{\theta} = \dfrac{\alpha}{n} + o\left(\dfrac{1}{n}\right)$, we get as $n \to \infty$

$$2\mathrm{Var}_{\theta} t \geqslant \sup_{\alpha} \frac{\theta^2}{n^2} \frac{4\alpha^2}{e^{2\alpha}+e^{4\alpha}-2}, \text{ asymptotically.} \qquad \dots \quad (2.15)$$

We get, on differentiating logarithmically, the equation

$$\frac{2}{\alpha} - \frac{2e^{2\alpha}+4e^{4\alpha}}{e^{2\alpha}+e^{4\alpha}-2} = 0. \qquad \dots \quad (2.16)$$

Solving this equation by iteration we get $\alpha = 0.446$ (correct to 3 decimal places). Hence we get for sufficiently large values of n the limit

$$2\mathrm{Var}_{\theta} t \geqslant 0.1244\theta^2/n^2,$$

or, $$\mathrm{Var}_{\theta} t \geqslant 0.0622\theta^2/n^2, \text{ asymptotically.} \qquad \dots \quad (2.17)$$

3. Regular estimation of the continuous type : the Cramér-Rao limit not attained

Let the density function of the chance variable x be given by

$$f(x, \theta)\, dx = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1}\, dx, \; (0 \leqslant x < \infty)$$

where p is a given positive constant, while θ is the unknown scale parameter.

In such a case it has been shown by Rao (1948) that the minimum variance unbiased estimate for θ exists and its variance is greater than the lower limit of Cramér-Rao.

Applying Hammersley's technique we get

$$\mathrm{Var}_{\theta} t \geqslant \sup_{\delta} \frac{\delta^2}{\left[E_{\theta}\left\{\dfrac{f(x, \theta+\delta)}{f(x, \theta)}\right\}^2\right]^n - 1} \qquad \dots \quad (3.1)$$

where $t(\underset{\sim}{x})$ is any unbiased estimator for θ based on a random sample $(x_1, x_2, \dots, x_n)$ of size n from the above population.

$$\text{Now} \quad E_\theta\left\{\frac{f(x,\theta+\delta)}{f(x,\theta)}\right\}^2 = \left(\frac{\theta+\delta}{\theta}\right)^{2p}\left(\frac{\theta}{\theta+2\delta}\right)^p.$$

Therefore, (3.1) reduces to

$$\text{Var}_\theta t \geqslant \operatorname*{Sup}_{\delta} \frac{\delta^2}{\left\{\left(1+\frac{\delta}{\theta}\right)^{2pn}\Big/\left(1+\frac{2\delta}{\theta}\right)^{pn}\right\}-1}. \qquad \ldots \quad (3.2)$$

Let us choose $\delta = \theta\frac{\alpha}{n}$ where α is a constant not depending on n.

Then (3.2) reduces to

$$\text{Var}_\theta t \geqslant \operatorname*{Sup}_{\alpha} \frac{\theta^2}{n^2}\frac{\alpha^2}{\left\{\left(1+\frac{\alpha}{n}\right)^{2pn}\Big/\left(1+\frac{2\alpha}{n}\right)^{pn}\right\}-1}. \qquad \ldots \quad (3.3)$$

$$\text{Now} \ \log\left\{\left(1+\frac{\alpha}{n}\right)^{2pn}\Big/\left(1+\frac{2\alpha}{n}\right)^{pn}\right\}$$

$$= 2pn\log\left(1+\frac{\alpha}{n}\right) - pn\log\left(1+\frac{2\alpha}{n}\right)$$

$$= \frac{p\alpha^2}{n} - \frac{2p\alpha^3}{n^2} + O\left(\frac{1}{n^3}\right).$$

Hence the r.h.s. of (3.3) simplifies to

$$\operatorname*{Sup}_{\alpha}\frac{\theta^2}{n^2}\cdot\frac{\alpha^2}{\exp\left\{\frac{p\alpha^2}{n}-\frac{2p\alpha^3}{n^2}+O\left(\frac{1}{n^3}\right)\right\}-1}$$

$$= \operatorname*{Sup}_{\alpha}\theta^2\frac{1}{np-\frac{(p^2\alpha^2-4p\alpha)}{2}+O\left(\frac{1}{n}\right)}.$$

Now we always have, $p^2\alpha^2-4p\alpha \geqslant 4$, the sign of equality holding only when $\alpha = 2/p$. Hence choosing α to our advantage we get

$$\text{Var}_\theta t \geqslant \frac{\theta^2}{np-2+O\left(\frac{1}{n}\right)}. \qquad \ldots \quad (3.4)$$

The minimum variance of the best unbiased estimate is $\theta^2/(np-2)$ while the Cramér-Rao limit is θ^2/np.

4. Comparison with the Cramér-Rao inequality

It will be shown that under certain general regularity conditions on the frequency function $f(x,\theta)$ Hammersley's technique will lead to a better lower limit to the variance of unbiased estimators than the Cramér-Rao limit.

Let us assume that the range of the chance variable x is independent of θ, and that, except on a set of points of measure zero

$$f(x,\theta+\delta)=f(x,\theta)+\delta f'(x,\theta)+\frac{\delta^2}{2!}f''(x,\theta)+\frac{\delta^3}{3!}f'''(x,\theta(\delta)) \qquad \ldots \quad (4.1)$$

where $\theta < \theta(\delta) \leqslant \theta+\delta.$

Under certain further regularity conditions on $f(x,\theta)$ regarding the permissibility of differentiation under the sign of integration and the existence of certain integrals, it is possible to write

$$E_\theta\left\{\frac{f(x,\theta+\delta)}{f(x,\theta)}\right\}^2=1+\delta^2J_{11}+\delta^3J_{12}+O(\delta^4) \qquad \ldots \quad (4.2)$$

where $$J_{11}=E_\theta\left\{\frac{f'(x,\theta)}{f(x,\theta)}\right\}^2, \qquad J_{12}=E_\theta\left\{\frac{f'(x,\theta)f''(x,\theta)}{f(x,\theta)f(x,\theta)}\right\}.$$

We also assume $J_{12}\neq 0$.

Now Hammersley's limit is given by

$$\text{Var}_\theta t \geqslant \underset{\delta}{\text{Sup}}\,[E_\theta(J\,|\,\theta,\delta)]^{-1} \qquad \ldots \quad (4.3)$$

where
$$E(J\,|\,\theta,\delta)=\frac{1}{\delta^2}\left[\left\{E_\theta\left(\frac{f(x,\theta+\delta)}{f(x,\theta)}\right)^2\right\}^n-1\right]$$

$$=\frac{1}{\delta^2}\left[\{1+\delta^2J_{11}+\delta^3J_{12}+O(\delta^4)\}^n-1\right]$$

$$=\frac{1}{\delta^2}\left[n(\delta^2J_{11}+\delta^3J_{12})+O(\delta^4)\right]$$

$$=nJ_{11}+n\delta J_{12}+O(\delta^2). \qquad \ldots \quad (4.4)$$

We thus get

$$\text{Var}_\theta t \geqslant \sup_\delta \frac{1}{nJ_{11}+n\delta J_{12}+O(\delta^2)} \qquad \dots \quad (4.5)$$

whereas the Cramér-Rao limit is

$$\text{Var}_\theta t \geqslant \frac{1}{nJ_{11}}. \qquad \dots \quad (4.6)$$

Now let $J_{12}>0$; then from (4.5) it appears that for any fixed n, choosing a sufficiently small $\delta<0$ the denominator in (4.5) can be made $<nJ_{11}$. Similarly for $J_{12}<0$ by assigning δ an arbitrarily small value >0 the denominator in (4.5) can be made $<nJ_{11}$, which proves the result.

Note that $\frac{1}{n(J_{11}-J^2_{12}/J_{22})}$, where $J_{22} = E_\theta\left\{\frac{f''(x,\theta)}{f(x,\theta)}\right\}^2$, gives a better lower limit (Bhattacharyya, 1946) for the variances of the unbiased estimators for θ than $\frac{1}{nJ_{11}}$ when $J_{12}\neq 0$.

In the example we were considering $J_{12}=-2p/\theta^3$.

My thanks are due to Shri D. Basu for his constant encouragement and helpful criticisms during the preparation of this paper.

REFERENCES

BASU, D. and MITRA, S. K. : A note on unbiased estimation of the binomial proportion. To be published in *Proceedings, National Institute of Sciences*, India.

BHATTACHARYYA, A. (1946–47) : On some analogues of the amount of information and their uses in statistical estimation. *Sankhyā*, 8, 1–14, 201–218, 315–328.

CHAPMAN, D. G. and ROBBINS, H. (1951) : Minimum variance estimation without regularity assumptions. *Ann. Math. Stat.*, 22, 581–586.

CRAMÉR, H. (1946) : *Mathematical Methods of Statistics*, Princeton University Press.

HAMMERSLEY, J. M. (1950) : On estimating restricted parameters. *J. Roy. Stat. Soc.*, Series B, 12, 192–220.

RAO, C. R. (1948) : Sufficient statistics and minimum variance estimates. *Proc. Camb. Phil. Soc.*, 45.

Paper received : August, 1952.