CHAPTER - VI
BAYES ESTIMATION FOR THE VARIANCE OF THE BEST ESTIMATOR
IN THE $T_1$-CLASS

6.0 Summary.

Prabhu-Ajgaonkar (1965) derived the best estimator for the $T_1$-class when Ikeda-Sen's (Midzuno's 1952) sampling procedure is employed. He (1971b) demonstrated that the variance function of a linear estimator can be expressed into a quadratic form and accordingly he noted that there exists nine principal classes of estimators out of which one principal class was examined by him in detail. He obtained expression for a unique estimator variance of the best estimator in the $T_1$-class for the Ikeda-Sen's sampling procedure. In this chapter we have derived a lower bound for the Bayes risk of an unbiased variance estimator for the variance of the best estimator $t$ in the $T_1$-class.

6.1. Introduction.

Let $\mathbb{U} = \{U_1, U_2, \ldots, U_N\}$ denote a finite population (set) of units (elements) $U$. On $\mathbb{U}$ is defined a real variate (function) $Y, Y(U)$ being its value for the unit $U$. If $\mathbb{Y}$ is a class of all real variates $Y$ defined on $\mathbb{U}$ and $f$ is a real valued
function on $\mathbf{Y}$ then the general problem in sampling is to estimate $f(\mathbf{Y})$ by observing the values $Y(U)$ for just those units $U$ belonging to specified sub-set $s$ of $\mathbf{U}$, given the sampling design. A function $T$ on $\mathbf{Y}$ is called the population total if for every $Y \in \mathbf{Y}$,

$$T(Y) = \sum_{U \in \mathbf{U}} Y(U).$$

Thus here the problem is to estimate $T(\mathbf{Y})$ by observing those values $Y(U)$ for which $U \in s$, where $s \subseteq \mathbf{U}$. A subset $s$ of $\mathbf{U}$ is called a sample $(s \subseteq \mathbf{U})$. $S$ denotes the set of all possible samples $s$, $S = \{ s_i \}$. On $S$ defined a function $p$ such that $1 \geq p(s) > 0$, for all $s \in S$ and $\sum_{s \in S} p(s) = 1$.

A sampling design $D = (S, p)$ is defined by attaching a probability of selection $p(s)$ to each $s \in S$ so that

$$p(s) \geq 0, \sum_{s \in S} p(s) = 1.$$

Prabhu-Ajgaonkar (1971b) demonstrated that whenever for the estimation of the population total, a linear estimator is employed its sampling variance can be expressed into a quadratic form, the general expression for which, is given by

$$Q = \sum_{i=1}^{N} A_{ii}^2 Y_i^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} Y_i Y_j$$

$$= \sum_{i=1}^{N} A_{ii}^2 Y_i^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} Y_i Y_j$$

(6.1.1)
where $A_{ij} (i,j=1,2,...,N)$ depend on the probabilities, the sample size and population size. Consequently the problem of estimating $Q$ reduces to the estimation of the two functions

$$Q_1 = \sum_{i=1}^{N} A_{ii} Y_i^2$$

and

$$Q_2 = \sum_{i \neq j=1}^{N} A_{ij} Y_i Y_j$$  \hspace{1cm} (6.1.2)$$

He further noted that the estimation of $Q_1$ is independent of $Q_2$. Therefore there exists nine principal classes of quadratic estimators referred to as

$$\hat{Q}(x)(y)$$  \hspace{1cm} (6.1.3)$$

where $x,y=1,2,3$; if for the estimation of $Q_1$ and $Q_2$ the estimators based on the axioms present in $x$ and $y$ principal classes respectively are employed.

6.2. Bayes Approach.

Prabhu-Ajgaonkar (1965) demonstrated that the best estimator for the $T_1$-class when Ikeda-Sen's (Midzuno's, 1952) sampling procedure employed, is

$$t = y_1 + \frac{N-1}{n-1} \sum_{r=2}^{n} y_r$$  \hspace{1cm} (6.2.1)$$
and the expression for its variance is given by

\[
\text{Var}(t) = \frac{N-n}{n-1} \left[ \sum_{i=1}^{N} (1-p_i)Y_i^2 - \frac{N}{N-2} \sum_{i \neq j=1}^{N} Y_i Y_j \right] \tag{6.2.2}
\]

where \( p_i \) represents the probability of drawing the \( i \)th element \( Y_i \) of the population at the first draw. A succinct description of Ikeda-Sen's sampling procedure is as follows:

In this sampling procedure elements are selected with varying probabilities on the first draw and on subsequent draws they are selected with equal probabilities without replacement; the sample size being always greater than or equal to two.

For a given sampling design \( (S,p) \) an estimator \( v \) is said to be unbiased for the variance, \( \text{Var}(t) \) if

\[
E[v(s,y)] = \text{Var}(t) \tag{6.2.3}
\]

A quadratic unbiased estimator \( \hat{Q} \) for variance \( \text{Var}(t) \) for a given sampling design \( (S,p) \) is a real valued function with property that

\[
\hat{Q} = \sum_{r=1}^{n} a_{rr} Y_r^2 + \sum_{r \neq k=1}^{n} a_{rk} Y_r Y_k \tag{6.2.4}
\]
where \( a_{rk} (r,k=1,2,\ldots,n) \) represents the weight associated with \( y_r y_k (r,k=1,2,\ldots,n) \) and

\[
\sum_{r=1}^{n} a_{rr} p_{ir} = \frac{N-n}{n-1} (1-p_i),
\]

for \( i=1,2,\ldots,N \)

and

\[
\sum_{r \neq k=1}^{n} a_{rk} p(ir,jk) = \frac{-(N-n)}{(N-2)(n-1)} (1-p_i p_j)
\]

for \( i \neq j=1,2,\ldots,N \).

where \( p_{ir} \) stands for the probability of selecting the \( i \)th population element at the \( r \)th draw and \( p(ir,jk) \) represents the probability of drawing the \( i \)th and \( j \)th population elements at the \( r \)th and \( k \)th draws respectively.

As a special case of the estimator \( \hat{Q} \) there is an estimator given by Prabhu-Ajgaonkar (1971b) for Ikeda-Sen's sampling procedure as follows:

\[
\hat{Q}(1)(1) = \frac{(N-n)(N-1)}{(n-1)^2} \left[ \sum_{r=2}^{n} y_r - \sum_{r \neq k=2}^{n} \frac{y_r y_k}{n-2} \right]
\]

which is the best estimator among the estimators belonging to the principal class \( \hat{Q}(1)(1) \).
The \( \text{Var}[\nu(s,y)] \) of the estimator \( \nu(s,y) \) is given by

\[
\text{Var}[\nu(s,y)] = \sum_{s \in S} \nu^2(s,y)p(s) - \left[ \text{Var}(t) \right]^2 \tag{6.2.7}
\]

It will be of interest to consider the problem of Bayes risk for the unbiased estimator \( \nu(s,y) \) with the loss function \( \text{Var}[\nu(s,y)] \). This problem has its own importance for the people who think that all our prior knowledge, in the present case about the population or variate under study can be formulated in some sort of prior distribution \( \alpha \). We use the same approach for the estimator of the population variance as was used by Godambe and Joshi (1965), and was used by Godambe (1955), Hajek (1959) and Aggrawal (1959) for the estimator of population mean.

Let \( Y_1, Y_2, \ldots, Y_N \) be independent random variables with zero mean and with joint distribution \( \alpha \) over \( Y \).

\( \mathcal{E}[\nu(s,y)], \text{Var}[\nu(s,y)], \text{Cov}[\nu(s,y), h(s,y)] \)

would denote the corresponding expectation, Variance and covariance with respect to \( \alpha \), \( s \) being held fixed,

Let \( \text{Var}[\nu(s,y)] \) denote the variance of an estimator \( \nu(s,y) \), then the Bayes risk of \( \nu(s,y) \) is

\[
\mathcal{E}\text{Var}[\nu(s,y)] = \int_{Y} \text{Var}[\nu(s,y)]d\alpha \tag{6.2.8}
\]
Theorem 6.2.1. Let $Y_1, Y_2, \ldots, Y_N$ be independent random variables with zero mean with joint distribution $\alpha$ over $\underline{Y}$, then for a sampling design $(S, p)$ for which $0 \leq p(s) \leq 1$, the Bayes risk of the unbiased estimator $v(s, \underline{y})$ has the lower bound given by

$$\mathbb{E} \text{Var}[v(s, \underline{y})] \geq \sum_{s \in S} p(s) \text{Var}[^*\hat{\xi}(1)(1)] - \text{Var}[\text{Var}(t)] \quad (6.2.9)$$

Proof: For a sampling design $(S, p)$, the variance of an unbiased estimator $v(s, \underline{y})$ for $\text{Var}(t)$ is

$$\text{Var}[v(s, \underline{y})] = \sum_{s \in S} v^2(s, \underline{y}) p(s) - [\text{Var}(t)]^2 \quad (6.2.10)$$

Hence the Bayes risk is

$$\mathbb{E} \text{Var}[v(s, \underline{y})] = \sum_{s \in S} p(s) \mathbb{E} \left[ v^2(s, \underline{y}) \right] - \mathbb{E} \left[ \text{Var}(t) \right]^2$$

$$= \sum_{s \in S} p(s) \left[ \mathbb{E} v(s, \underline{y}) \right]^2 + \sum_{s \in S} p(s) \text{Var}^* \left[ v(s, \underline{y}) \right]$$

$$- \mathbb{E} \left[ \text{Var}(t) \right]^2 \quad (6.2.11)$$

because

$$\text{Var}^* \left[ v(s, \underline{y}) \right] = \mathbb{E} \left[ v(s, \underline{y}) \right]^2 - \left\{ \mathbb{E} \left[ v(s, \underline{y}) \right] \right\}^2.$$
Now we obtain a lower bound for $\mathbb{E} \text{Var}[v(s,y)]$ as follows:

(i) Substituting the condition of unbiasedness from equation (6.2.3) into

$$\sum_{s \in S} p(s) \left[ \mathbb{E} v(s,y) - \text{Var}(t) \right]^2 \geq 0.$$ 

We have

$$\sum_{s \in S} p(s) \mathbb{E}^2 \left[ v(s,y) \right] \geq \mathbb{E}^2 \text{Var}(t) \quad (6.2.12)$$

(ii) Let

$$v(s,y) = \hat{Q}(1)(1) + \left[ v(s,y) - \hat{Q}(1)(1) \right]$$

$$= \hat{Q}(1)(1) + h(s,y) \quad (6.2.13)$$

Then we have from equation (6.2.13)

$$\text{Var}[v(s,y)] = \text{Var}(\hat{Q}(1)(1)) + \text{Var}[h(s,y)] + 2 \text{Cov}[\hat{Q}(1)(1), h(s,y)]$$

$$\quad (6.2.14)$$

Multiplying equation (6.2.14) by $p(s)$ and summing over all possible samples $s \in S$, we get
\[
\sum_{s \in S} \varstar[v(s, y)] p(s) = \sum_{s \in S} \varstar[\hat{Q}(1)(1)] p(s) + \sum_{s \in S} \varstar[h(s, y)] p(s) + 2 \sum_{s \in S} \covstar[\hat{Q}(1)(1), h(s, y)] p(s) \tag{6.2.15}
\]

To show that the last term in the right hand side of equation (6.2.15) vanishes we have

\[
\cov[\hat{Q}(1)(1), h(s, y)] = \mathcal{E}\left\{[\hat{Q}(1)(1) - \mathcal{E}\hat{Q}(1)(1)] [h(s, y) - \mathcal{E}h(s, y)]\right\}
\]

\[
= \mathcal{E}\left\{[\hat{Q}(1)(1) - \mathcal{E}\hat{Q}(1)(1)] h(s, y)\right\}
\]

\[
= \mathcal{E}\left\{\left[\frac{(N-n)(N-1)}{(n-1)^2} \left(\sum_{r=2}^{n} \frac{y^2}{y_r} - \sum_{r\neq k=2}^{n} \frac{y_r y_k}{n-2}\right)\right] h(s, y)\right\}
\]

\[
= \frac{(N-n)(N-1)}{(n-1)^2} \sum_{r=2}^{n} \mathcal{E}\left[(y^2_r - \mathcal{E}y^2_r) h(s, y)\right] - \frac{(N-n)(N-1)}{(n-1)^2} \sum_{r\neq k=2}^{n} \mathcal{E}\left[(y_r y_k - \mathcal{E}y_r y_k) h(s, y)\right] \tag{6.2.16}
\]

Multiplying both sides of equation (6.2.16) by \(p(s)\) and summing over all possible samples \(s \in S\) we get
\[ \sum_{s \in S} p(s) \text{ Cov} \left[ \hat{Q}(1)(1), h(s, y) \right] = \frac{(N-n)(N-1)}{(n-1)^2} \sum_{i=1}^{N} \xi_i \left( y_i^2 - \frac{1}{N} \sum y_i \right) \sum_{s \ni i} p(s) h(s, y) \]

- \frac{(N-n)(N-1)}{(n-1)^2(n-2)} \sum_{i \neq j}^{N} \xi_i \xi_j \left( y_i y_j - \frac{1}{N} \sum y_i \right) \sum_{s \ni i, j} p(s) h(s, y) \]

where \( s \ni i \) denotes the samples \( s \) which contains unit \( i \) and similarly \( s \ni i, j \) denotes the samples \( s \) which contains both the units \( i \) and \( j \).

Now since the estimator \( v(s, y) \) is unbiased we have from equation (6.2.13)

\[ \sum_{s \in S} p(s) h(s, y) = 0 \quad \text{for all} \quad y \in Y \]  \hspace{1cm} (6.2.18)

i.e.

\[ \sum_{s \ni i} p(s) h(s, y) = - \sum_{s \ni i} p(s) h(s, y) \]  \hspace{1cm} (6.2.19)

where \( \sum \) denotes the samples \( s \) which do not contain unit \( i \) and

\[ \sum_{s \ni i, j} p(s) h(s, y) = - \sum_{s \ni i, j} p(s) h(s, y) - \sum_{s \ni i} p(s) h(s, y) \]

- \( \sum_{s \ni i} p(s) h(s, y) \)  \hspace{1cm} (6.2.20)

\[ \sum_{s \ni j} p(s) h(s, y) \]
where \( \Sigma \) denotes all samples \( s \) which do not contain units \( s \neq i, j \).

Noting that zero mean and independence of \( Y_i \), we have

\[
\mathbb{E}\left\{ \left[ Y_i^2 - \mathbb{E} Y_i^2 \right] \sum_{s \neq i} p(s) h(s, y) \right\} = \mathbb{E} \left[ (Y_i^2 - \mathbb{E} Y_i^2) \right] \mathbb{E} \left[ - \sum_{s \neq i} p(s) h(s, y) \right] = 0 \quad (6.2.21)
\]

and

\[
\mathbb{E}\left\{ \left[ Y_i Y_j - \mathbb{E} Y_i \mathbb{E} Y_j \right] \sum_{s \neq i, j} p(s) h(s, y) \right\} = \mathbb{E} Y_i Y_j \mathbb{E} \left[ - \sum_{s \neq i, j} p(s) h(s, y) \right] + \mathbb{E} Y_j \mathbb{E} \left[ -Y_i \sum_{s \neq i} p(s) h(s, y) \right] + \mathbb{E} Y_i \mathbb{E} \left[ -Y_j \sum_{s \neq j} p(s) h(s, y) \right] = 0. \quad (6.2.22)
\]

Thus from equations (6.2.21) and (6.2.22) we have

\[
\sum_{s \in S} p(s) \, \text{Cov} \left[ \hat{Q}_i(1), h(s, y) \right] = 0 \quad (6.2.23)
\]
Now if we put this value in equation (6.2.15) we have the expression

$$
\sum_{s \in S} p(s) \var^*[v(s,y)] = \sum_{s \in S} p(s) \var^*[\hat{Q}(1)(1)] + \sum_{s \in S} p(s) \var^*[h(s,y)]
$$

(6.2.24)

i.e.

$$
\sum_{s \in S} p(s) \var^*[v(s,y)] \geq \sum_{s \in S} p(s) \var^*[\hat{Q}(1)(1)]
$$

(6.2.25)

Finally substituting the conditions (6.2.12) and (6.2.25) into the equation (6.2.11) we obtain the lower bound

$$
\mathcal{E}\var[v(s,y)] \geq \sum_{s \in S} p(s) \var^*[\hat{Q}(1)(1)] - \mathcal{E}[\var(t)]^2 + \mathcal{E}^2[\var(t)]
$$

$$
\geq \sum_{s \in S} p(s) \var^*[\hat{Q}(1)(1)] - \var^*[\var(t)]
$$

Hence the theorem. .