4.1. INTRODUCTION

Mukhopadhyay (1982, 1984) and Mukhopadhyay and Bhattacharyya (1989) considered the problem of variance estimation and obtained the optimal strategy in the class of homogeneous quadratic p-unbiased predictors. In this chapter we generalize their results by considering classes of quadratic p-unbiased and all p-unbiased predictors under Model G and Model E, respectively.

Let $U = \{1, 2, \ldots, N\}$ be a finite population of identifiable units. We denote by $S$ the set of all subsets $s$ of $U$. A sampling design is a function $p$ on $S$ such that $p(s) \geq 0$ and $\Sigma_{s} p(s) = 1$, where $\Sigma_{s}$ denotes summation over $s \in S$. Let $\pi_{i} = \Sigma_{s \ni i} p(s)$ and $\pi_{i j} = \Sigma_{s \ni i, j} p(s)$ be the first- and second-order inclusion probabilities for unit $i$ and for units $i$ and $j$. Associated with the $i$th unit is a characteristic $Y_{i}, i = 1, \ldots, N$. The population variance is

given as

\[ S^2_Y = c \sum_{i \in U} Y_i^2 - c \sum_{i,j \in U} Y_i Y_j \]

where \( c = 1/N \), \( c = 1/(N(N-1)) \) and \( \sum_{U}, \sum_{U} \) denote summations over \( i \in U \) and \( i \neq j \in U \), respectively, and we are interested in inference about \( S^2_Y \) from a subset \( s \), termed a sample. We consider the following superpopulation models.

Model \( G \) (general transformation model). The class of probability measures \( \xi \) on \( \mathbb{R} \) such that \( Y_1, \ldots, Y_N \) are asymptotically normally distributed and for \( i, j = 1, \ldots, N \):

\[
\begin{align*}
\mathbb{E}(Y_i) &= \mu_i = a_i \mu + b_i \\
\mathbb{V}(Y_i) &= \sigma_i^2 = a_i^2 \sigma^2 \\
\mathbb{C}(Y_i, Y_j) &= a_i a_j \rho \sigma^2 \quad (i \neq j)
\end{align*}
\]

where \( a_i > 0, b_i (i = 1, \ldots, N) \) are known numbers, \( \mu, \sigma^2 > 0 \), \( -1/(N-1) \leq \rho < 1 \) are unknown. \( \mathbb{E}(\cdot), \mathbb{V}(\cdot) \) and \( \mathbb{C}(\cdot, \cdot) \) denote, respectively, \( \xi \)-expectation, \( \xi \)-variance and \( \xi \)-covariance.

Model \( G \) (regression model). The class of probability measures \( \xi \) on \( \mathbb{R} \) such that \( Y_1, \ldots, Y_N \) are independently normally distributed, and

\[
\begin{align*}
\mathbb{E}(Y_i) &= \beta x_i \\
\mathbb{V}(Y_i) &= \sigma^2 v(x_i) \\
(i = 1, \ldots, N)
\end{align*}
\]
where $\beta$ and $\sigma^2$ are unknown, $v(\cdot)$ is a known function and $x_1, \ldots, x_N$ are known positive numbers.

Model $G_{\text{MR}}$ (multiple regression model). The class of probability measures $\xi$ on $\mathbb{R}^N$ such that $Y_1, \ldots, Y_N$ are independently distributed, and

$$
\xi(Y_i) = \beta_i + \sum_{k=2}^{q} \beta_k x_{ik}
$$

$$
\nu(Y_i) = \sigma^2 v_i
$$

where $\beta_1, \ldots, \beta_q, \sigma^2$ are unknown, and $x_{i2}, \ldots, x_{iq}, v_i$ is a set of known numbers for every $i (i = 1, \ldots, N)$.

Model $E_T$ (exchangeable transformation model). This model defines the class of probability measures $\xi$ on $\mathbb{R}^N$ such that, for known numbers $a > 0, b_i (i = 1, \ldots, N)$ the random variables $Z_i = (Y_i - b_i)/a_i$ $(i = 1, \ldots, N)$ have an exchangeable, absolutely continuous distribution. The first and second moments of the $Y_i$ are given by $(*)$.

Model $E_{\text{RP}}$ (random permutation model). The class of distributions $\xi$ such that, for any fixed, unknown numbers $z_1, \ldots, z_N$, and for given numbers $a_i > 0, b_i (i = 1, \ldots, N)$ the random variables $Z_i = (Y_i - b_i)/a_i$ have an exchangeable distribution such that
\[ P(Z = z_1, \ldots, Z = z_N) = \frac{1}{N!} \]

for every permutation \( r_1, \ldots, r_N \) of \( 1, \ldots, N \).

The \( \xi \)-moments of the \( Y_i \) are, for \( i = 1, \ldots, N \):

\[
\begin{align*}
\xi_i(Y_i) &= a_i \mu_z + b_i \\
\gamma_i(Y_i) &= a_i^2 \sigma_z^2 \\
\text{Cov}(Y_i, Y_j) &= -a_i a_j \sigma_z^2/(N - 1) \\
\end{align*}
\]

(i \( \neq \) j)

where the unknown \( \mu_z \) and \( \sigma_z^2 \) are given by

\[
\begin{align*}
\mu_z &= \frac{1}{N} \sum_{i=1}^{N} z_i, \\
\sigma_z^2 &= \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \mu_z)^2.
\end{align*}
\]

A predictor \( Q \) of \( S_Y^2 \) is said to be quadratic predictor if

\[
Q = b_s + \sum_{s} b(s,ii)Y_i^2 + \sum_{s} b(s,ij)Y_i Y_j
\]

where \( b_s \), \( b(s,ii) \), \( b(s,ij) = b(s,ji) \) are constants not depending on the \( Y \)-values, \( \sum_{s} \) and \( \sum_{s} \) denote summation over \( i \in s \) and \( i \neq j \in s \).

Definition 4.1. \( Q \) is called a \( p \)-unbiased predictor of \( S_Y^2 \) if and only if, for given design \( p \), \( E(q) = S_Y^2 \) for all \( y \in \mathbb{R}^N \), where \( q \) is the value of \( Q \) for \( Y_i = y_i, i \in s \). The strategy \( (p, Q) \) is called \( p \)-unbiased if \( Q \) is a \( p \)-unbiased predictor under \( p \).
Definition 4.2. $Q$ is called a $\xi$-unbiased predictor of $S^2_Y$ if and only if, for a given $\xi$, $\mathbb{E}(Q - S^2_Y) = 0$ for all $s \in S$, the set of all samples.

Definition 4.3. $Q$ is called a $p_\xi$-unbiased predictor of $S^2_Y$ if and only if, for given $p$ and $\xi$, $\mathbb{E}(Q - S^2_Y) = 0$.

The $\xi$-expected $p$-mean square error, denoted by $\mathbb{MSE}$, of an arbitrary strategy $(p, Q)$ is defined as

$$\mathbb{MSE}(p, Q) = \mathbb{V}(Q - S^2_Y)$$

which, if $Q$ is $p$-unbiased, reduces to the $\xi$-expected $p$-variance, $\mathbb{V}(p, Q)$. Here $\mathbb{E}(\cdot)$, $\mathbb{V}(\cdot)$ and $\mathbb{MSE}(\cdot)$ denote, respectively, $p$-expectation, $p$-variance and $p$-mean square error with respect to design $p$.

Definition 4.4. If $Q_1$ and $Q_2$ are predictors such that, for the given design $p$, $\mathbb{MSE}(p, Q_1) \leq \mathbb{MSE}(p, Q_2)$ for all $\xi \in \mathcal{E}$, a given class of superpopulations, then we say that $Q_1$ is at least as good a predictor $Q_2$ for the design $p$. If strict inequality holds for at least one $\xi \in \mathcal{E}$, then $Q_1$ will be called better than $Q_2$.

Definition 4.5. If $(p_1, Q_1)$ and $(p_2, Q_2)$ are strategies such that $\mathbb{MSE}(p_1, Q_1) \leq \mathbb{MSE}(p_2, Q_2)$ for all $\xi \in \mathcal{E}$ than we shall
say that \((p_1, Q_1)\) is at least as good a strategy as \((p_2, Q_2)\).

If strict inequality holds for at least one \(\xi \in \mathcal{C}\), then we shall say that \((p_1, Q_1)\) is better than \((p_2, Q_2)\).

Our purpose is to find a predictor \(Q\) of \(s^2_y\) such that \(\mathbb{E}(Q - s^2_y)^2\) is minimised subject to the requirement of \(p\)-unbiasedness of \(Q\) viz. \(\mathbb{E}(Q) = s^2_y\).

Mukhopadhyay (1982,1984) and Mukhopadhyay and Bhattacharyya (1989) considered the above problem and obtained the optimal strategy \((p^*_o, Q^*_{HTO})\) where \(p^*_o\) is such that

\[
\pi^*_i = \frac{nf(a_i)}{\sum f(a_i)},
\]

\[
\pi^*_{ij} = \frac{n(n-1)f(a_i)f(a_j)}{\sum f(a_i)f(a_j)},
\]

and

\[
Q^*_{HTO} = \frac{\sum Y_i^2/N\pi^*_i - \sum Y_i Y_j/N(N-1)\pi^*_{ij}}{N\pi^*_{ij}},
\]

the Horvitz-Thompson type predictor, on restricting to the class of homogeneous quadratic predictors under Model \(G_T\) with \(i\) \(b_i = 0, a^2_i \propto f(a_i) \forall i, a\) known function of \(a_i, \mu = \rho = 0\) and \(ii\) \(b_i = 0 \forall i\) and \(\rho = 0\). Sengupta (1988) derived the strategy \((p^*_o, s^2_y)\) where \(p^*_o\) is any design with \(\pi^*_i = n/N, \pi^*_{ij} = n(n-1)/N(N-1)\) and \(s^2_y\) is the sample variance, as the optimal strategy under Model \(G_T\) with \(a_i = 1, b_i = 0 \forall\)
Further he proved that it is admissible. We note that Sengupta's (1988) model is a particular case of the model considered by Mukhopadhyay and Bhattacharyya (1989). However the former did not consider the normality.

4.2. OPTIMUM p-UNBIASED PREDICTION: MODEL G

We shall consider noninformative sampling design $p$ with FES(n). Let $Q_p$ be the class of all, homogeneous or nonhomogeneous, quadratic p-unbiased predictors and $S_p$ be the class of all p-unbiased predictors of $S_Y^2$. We prove the following theorem.

Theorem 4.1 Under Model $G_T$,

$$\mathbb{E}(V(p, Q)) \geq \mathbb{E}(V(p_o, Q_{o, o})) \geq 4\mu^2 \sigma^2 K(a, \rho) - \tau^2$$

where $p_o = p_o(s)$ is any FES(n) design such that

$$\pi_{i_0} = na_i^2 / \sum a_i^2; \quad \pi_{i,j_0} = n(n-1)a_i a_j / \sum a_i a_j \quad (2.1)$$

the predictor

$$Q_{o, o} = \hat{S}_o^2 + c_1 \sum_{i \in S} \frac{Y_{i}^2 - b_i^2}{\pi_{i_0}} - c_2 \sum_{i,j \in S} \frac{Y_{i} Y_{j} - b_i b_j}{\pi_{i,j_0}} \quad (2.2)$$
and the functions

\[ K(a, \rho) = \frac{c^2}{n} \left( \sum_{i=1}^{n} a_i^2 \right)^2 \left[ \rho^* + \frac{2}{n\mu} \rho^* \sum_{i} \frac{b_i}{a_i} + \frac{1}{n\mu^2} \left( \sum_{i} \frac{b_i^2}{a_i^2} + \rho \sum_{s} \frac{b_i b_j}{a_i a_j} \right) \right] \]

\[ + \frac{c^2}{n(n-1)} \left( \sum_{i=1}^{n} a_i a_j \right)^2 \left[ (1+\rho^*) + \frac{1}{n\mu^2} \sum_{i} \frac{b_i^2}{a_i^2} \right] \]

\[ + \frac{2}{n\mu} (1 + \rho^*) \sum_{i} \frac{b_i}{a_i} + \frac{\rho^{**}}{n(n-1)\mu^2} \sum_{s} \frac{b_i b_j}{a_i a_j} \]

\[ - \frac{2c^2}{n} \left( \sum_{i=1}^{n} a_i^2 \right) \left( \sum_{i=1}^{n} a_i a_j \right) \left[ \rho^* + \frac{2}{n\mu} \rho^* \frac{b_i}{a_i} \right] \]

\[ + \frac{\rho}{n\mu^2} \left( \sum_{s} \frac{b_i^2}{a_i^2} - \frac{1}{n-1} \sum_{s} \frac{b_i b_j}{a_i a_j} \right) + \frac{\rho^{**}}{n(n-1)\mu^2} \sum_{s} \frac{b_i b_j}{a_i a_j} \]

and

\[ \tau^2 = \gamma(S^2_Y) \tag{2.3} \]

where

\[ \rho^* = 1 + (n-1)\rho \quad \text{and} \quad \rho^{**} = 1 + 2(n-2)(n-3)\rho, \]

for any strategy \((p, Q)\) such that \(p\) is an \(FES(n)\) design and \(Q \in Q_p\), the class of all quadratic (homogeneous or nonhomogeneous) \(p\)-unbiased predictors of \(S^2_Y\); equality holds if and only if \((p, Q) = (p^*_o, Q^*_GDO)\). 

56
Proof. The requirement of p-unbiasedness is equivalent to a set of $N^2 + 1$ conditions given by

\[
\begin{align*}
\sum_{s} p(s) b &= 0 \\
\sum_{s \in \Omega} p(s) b(s,ii) &= c_i \quad (i = 1, \ldots, N) \\
\sum_{s \in \Omega} p(s) b(s,ij) &= -c_{i:j} \quad (i \neq j = 1, \ldots, N)
\end{align*}
\]

(2.4)

Note that p-unbiasedness implies $\phi$-unbiasedness. Hence the above $N^2 + 1$ conditions imply

\[
\begin{align*}
\sum_{s \in \Omega} p(s) \left[ b + \sum_{s} b(s,ii) b^2 + \sum_{s} b(s,ij) b b \right] &= S_b^2 \\
\sum_{u \in \Omega} a^2 \sum_{s \in \Omega} p(s) b(s,ii) &= c \sum_{u \in \Omega} a^2 \\
\sum_{u \in \Omega} a a \sum_{s \in \Omega} p(s) b(s,ij) &= -c \sum_{u \in \Omega} a^2 \\
(2.5a)
\end{align*}
\]

(2.5b) can be rewritten as

\[
\begin{align*}
\sum_{u \in \Omega} a b \sum_{s \in \Omega} p(s) b(s,ii) &= c \sum_{u \in \Omega} a b \\
\sum_{u \in \Omega} a b \sum_{s \in \Omega} p(s) b(s,ij) &= -c \sum_{u \in \Omega} a b \\
(2.5c)
\end{align*}
\]
where

\[
\begin{align*}
2h(s, ii) &= b(s, ii) a_i^2 \\
2h(s, ij) &= b(s, ij) a_i a_j
\end{align*}
\] (2.7)

As \( p(s) \) is noninformative and as \( Q \) is \( p \)-unbiased, we have

\[
\mathbb{E}\{V(P, Q)\} = \mathbb{E}\{V(Q)\} + \mathbb{E}\{\mathbb{E}(Q)\}^2 - V(Q)^2
\] (2.8)

where \( \mathbb{E}(Q) = \mathbb{E}(Q) - \mathbb{E}(Q)^2 \).

Without loss of generality, assume that \( b_i = 0 \) (\( i = 1, \ldots, N \)) in Model \( G \).

Analysing the terms of (2.8) we first find, using (10.12) and (10.13) of Kendall et.al. (1987), that

\[
\mathbb{E}\{V(Q)\}
\]

\[
= 4\mu^2 \sigma^2 \sum_s p(s) \left\{ \sum_i b^2(s, ii) a_i^4 + \rho \sum_i \sum_i b(s, ii) b(s, i'i') a_i^2 a_{i'}^2 \\
+ (1 + 3\rho) \sum_i \sum_i \sum_k b(s, ij) b(s, ik) a_i^2 a_j a_k \\
+ (1/2) (1+\rho) \sum_i \sum_i b^2(s, ij) a_i^2 a_j^2 \\
+ \rho \sum_i \sum_i \sum_i \sum_j b(s, ij) b(s, i'j') a_i a_j a_{i'} a_{j'} \\
+ 2 (1-\rho) \sum_i \sum_i b(s, ii) b(s, ij) a_i^3 a_j \\
+ 2\rho \sum_i b(s, ii) a_i^2 \sum_i b(s, ij) a_i a_j \right\}.
\]

Using (2.6) and (2.7), we can write the above \( \mathbb{E}V \) as
\[ E[\varphi(Q)] = 4 \mu^2 \sigma^2 \left[ \sum_s p(s) \left\{ (1-\rho) \sum h^2(s,ii) + \rho \left( \sum h(s,ii) \right) \right\}^2 \right. \\
\left. + \frac{(1-\rho)}{2} \sum h^2(s,ij) + \rho \left( \sum h(s,ij) \right) \right] \\
+ 2(1-\rho) \sum h(s,ii)h(s,ij) \right\} \\
- 2\rho \sum_{i} \sum_{u} (\sum a_i^2) \left( \sum_{u} a_i a_j \right) \] 

Let us minimize
\[
\phi = E[\varphi(Q)] - 8\mu^2 \sigma^2 \sum_s p(s) \sum h(s,ii) - c \sum a_i^2 \\
- 8\mu^2 \sigma^2 \sum_s p(s) \sum h(s,ij) + c \sum a_i a_j
\]

with respect to \( b(s,ii) \) and \( b(s,ij) \), where \( \lambda_1 \) and \( \lambda_2 \) are Langrangian multipliers. By equating the partial derivative of \( \phi \) with respect to \( h(s,ii) \) to zero, we find that
\[
\lambda_1 = \sum_s p(s) \left\{ (1-\rho) h(s,ii) + \rho \sum h(s,ii) \\
+ (1-\rho) \sum_{j \neq i} h(s,ij) \right\} \tag{2.9}
\]

where \( \bar{S} \) is the set of all subsets \( s \) such that \( p(s) > 0 \).

Summing over \( i \in s \) we have using (2.6) and the fixed-size property
\[
\lambda_1 = (1-\rho)(c/n) \sum a_i^2 + \rho \sum a_i^2 - (1-\rho)(c/n) \sum a_i a_j
\]

59
Inserting this value in (2.9) we find, with the help of (2.6), that

\[ h(s,ii) + \sum_{j \neq i \in s} h(s,ij) = \frac{c_1}{n} \sum_u a_i^2 - \frac{c_2}{n(n-1)} \sum_u a_i a_j \] (2.10)

Again, setting \( \partial \phi/\partial h(s,ij) \) equal to zero and summing over \( i \neq j \in s \) we get, after some simplification,

\[ h(s,ii) + h(s,ij) = \frac{c_1}{n} \sum_u a_i^2 - \frac{c_2}{n(n-1)} \sum_u a_i a_j \] (2.11)

Solving (2.10) and (2.11) for \( h(s,ii) \) and \( h(s,ij) \) under the condition (2.6), we get the unique solution as

\[
\begin{align*}
    h_n(s,ii) &= \frac{c_1}{n} \sum_u a_i^2 \\
    h_n(s,ij) &= -\frac{c_2}{n(n-1)} \sum_u a_i a_j 
\end{align*}
\] (2.12)

from which

\[
\begin{align*}
    b_n(s,ii) &= \frac{c_1}{n} (\sum_u a_i^2) - \frac{1}{a_i} = \frac{c_1}{\pi_{i0}} \\
    b_n(s,ij) &= -\frac{c_2}{n(n-1)} (\sum_u a_i a_j) - \frac{1}{a_i a_j} = -\frac{c_2}{\pi_{ij0}}
\end{align*}
\] (2.12)

and hence the optimal \( p^*_\xi \)-unbiased predictor as

\[
Q_{H\xi} = c \sum_{i \in s} \frac{Y_i^2}{\pi_{i0}} - c \sum_{i \in s} \sum_{j \neq i \in s} \frac{Y_i Y_j}{\pi_{ij0}}
\]

Using (2.12) in \( E[\psi(Q)] \) along with the assumption \( b_i = 0 \ \forall \ i \) we get the function
\[ K(\mathbf{a}, \rho) = \frac{c_i^2}{n} \left\{ 1 + (n-1) \rho \right\} (\sum_{i} a_i^2)^2 \]

\[ + \frac{c_i^2}{2n(n-1)} \left\{ 1 + [1+2(n-2)(n-3)]\rho \right\} (\sum_{i} a_i a_j)^2 \]

\[ - \frac{2c_i c_j}{n} \left\{ 1 + (n-1)\rho \right\} (\sum_{i} a_i^2) (\sum_{i} a_i a_j) \]

which when substituted in (2.8) will give the minimum value of \( \mathbb{E}[V(Q)] \).

Suppose that \( b_i \neq 0 \) for all or some \( i \). Then, as (2.12) satisfies the constraint (2.5c), we have from (2.5a),

\[ b_{os} = s^2 - c \sum_{i} \frac{b_i^2}{\pi_i \pi_{io}} + c \sum_{i} \frac{b_i b_j}{\pi_i \pi_{ij}} \quad (2.13) \]

and hence the optimal \( p^* \)-unbiased predictor as \( Q_{GDO} \).

Next, in (2.8) \( E[\mathbb{E}(Q)] \geq 0 \) where equality is attained in particular if \( b_s \) is given in (2.13), and \( b(s,ii), b(s,ij) \) are given in (2.12). Thus under (2.5) for any given FES(n) design \( p(s) \)

\[ E[V(Q)] + E[\mathbb{E}(Q)]^2 \geq 4\mu^2 \sigma^2 K(\mathbf{a}, \rho) \quad (2.14) \]

where equality holds if and only if \( Q = Q_{GDO} \).

Finally, it follows from (2.4) that (2.6) is
satisfied. Moreover \((p, Q) = (p_{\text{GDO}}, Q_{\text{GDO}})\) is \(p\)-unbiased, attains equality in (2.14), and therefore has minimum \(\mathcal{E}(V)\) among all \((p, Q)\) such that \(Q \in \mathcal{Q}_p\) and \(p\) has \(\text{FES}(n)\). Noting that \(\mathcal{E}(S^2_Y) = \tau^2\), it follows from (2.8) and (2.14) that

\[
\mathcal{E}(p_{\text{GDO}}, Q_{\text{GDO}}) = 4\mu^2\sigma^2 K(a, \rho) - \tau^2
\]

as claimed by the theorem.

Remark 4.1. It is easy to see from the proof that, for any given fixed size design \(p(s)\), \(Q_{\text{GDO}}\) minimizes \(\mathcal{E}(S - \bar{S}_Y^2)\), that is \(Q_{\text{GDO}}\) is best among quadratic, \(p\)-unbiased estimators of \(\mathcal{E}(S^2_Y)\). Any given \(\text{FES}(n)\) design is optimal for \(Q_{\text{GDO}}\).

Remark 4.2. Results of Sengupta (1988) is deducible from our result by taking a simple model with \(a_i = 1, b_i = 0\) \(\forall\ i\) and \(\rho = 0\) in Model \(G_T\).

Remark 4.3. In various interesting special cases of Model \(G_T\) the predictor \(Q_{\text{GDO}}\) given by (2.2) can be expressed as follows:

(i) If \(a_i = 1, b_i = 0\) \(\forall\ i\), then \((p_{\text{GDO}}, S^2_Y)\) is the best strategy, where \(p_{\text{GDO}}\) is such that \(\pi = n/N\) and \(\pi_{ij} = n(n-1)/N(N-1)\), \((i, j = 1, \ldots, N, i \neq j)\), which is satisfied for the design SRS, and the predictor is the sample variance.
(ii) If \( a = 1 \) for all \( i \), then \((p^*, Q^*_{\text{DO}})\) is the best strategy where

\[
Q^*_{\text{DO}} = s^2_Y + s^2_b - s^2
\]

is the difference type predictor.

(iii) If \( b = 0 \) for all \( i \), then \((p^*, Q^*_{\text{HTO}})\) is the best strategy, where

\[
Q^*_{\text{HTO}} = c\frac{\sum Y_i^2}{\pi} - c\frac{\sum Y_i Y_j}{\pi} \]

is the Horvitz-Thompson type predictor, which is extensively studied by Liu (1974a, 1974b) as an estimator of the population variance.

Corollary 4.1. Under Model \( G \), and for any design \( p \),

\[
\mathbb{E}\{V(p, Q_{\text{HT}})\} \geq \mathbb{E}\{V(p^*, Q^*_{\text{GDO}})\} \equiv 4\mu^2 \sigma^2 K(a, \rho) - \tau^2
\]

where \( Q_{\text{HT}} = c\frac{\sum Y_i^2}{\pi} - c\frac{\sum Y_i Y_j}{\pi} \). Equality holds if and only if \( p = p_o \) and \( b \propto a_i \) (\( i = 1, \ldots, N \)) in which case \( Q^*_{\text{GDO}} = Q^*_{\text{HTO}} \) so that the minimum \( \mathbb{E}V \) strategy becomes \((p^*, Q^*_{\text{GDO}})\).

4.3. OPTIMUM \( p \)-UNBIASED PREDICTION: MODEL \( E \)

In this section, relaxing the conditions of
'Tnormality' and 'quadraticity' but imposing the stronger condition 'exchangeability' the optimality of $Q_{GDO}$ is established in the wider class of all $p$-unbiased predictors, $\mathcal{A}_p$.

Theorem 4.2. Under Model E$_T$,

$$\mathbb{E}\{V(p,Q)\} \geq \mathbb{E}\{V(p_\circ, Q_{GDO})\} \cong 4\mu^2 \sigma^2 K(a, \rho) - \tau^2$$

for any strategy $(p, Q)$ such that $p$ is an FES(n) design and $Q \in \mathcal{A}_p$, the class of all (quadratic or nonquadratic) $p$-unbiased predictors of $S^2_{Y}$; equality holds if and only if $(p, Q) = (p_\circ, Q_{GDO})$, where $p_\circ$ and $Q_{GDO}$ are given by (2.1) and (2.2) respectively.

Proof. We have under Model E$_T$,

$$\mathbb{E}\{S^2_{Y}\} = \mathbb{E}\{S^2_{\circ}\} + c\sum \frac{(c^2 + \mu^2)}{u} a_i^2 - c\sum (\rho^2 + \mu^2) \sum u_i a_i a_j + 2\mu \left[ c \sum u_i a_i b_i - c \sum u_i a_i b_j \right]$$

(3.1)

We must find $Q$ to minimize the $p\xi$-variance

$$\mathbb{E}(Q - S^2_{Y})^2 = \mathbb{E}[\mathbb{E}(Q)] + \mathbb{E}[\mathbb{E}(Q)]^2$$

subject to $\mathbb{E}(Q) = \mathbb{E}(S^2_{Y})$.

Define $Z_i = (Y_i - b_i)/a_i$, $(i = 1, \ldots, N)$. Then proceeding on
the same line as in Lemma 4.5 of Cassel et al. (1977), one can easily show that the family of distributions of \( Z_s = \{Z_i : i \in s\} \) is complete. Thus we restrict our search to statistics \( Q \) that depend on data \( d = \{(i, Y_i) : i \in s\} \) only through \( Z_s \).

But as \( \bar{Z}_s = \frac{1}{n} \sum_{i \in s} Z_i \) and \( \sum_{s \subseteq i} Z_i Z_j / n(n-1) \) are respectively, the unique predictors which are functions of \( Z_s \) and \( p\xi\)-unbiased for \( \mu = \mathbb{E}(Z_i), \mu^2 + \sigma^2 = \mathbb{E}(Z_i^2) \) and \( \rho \sigma^2 + \mu^2 = \mathbb{E}(Z_i Z_j), (i, j = 1, \ldots, N, i \neq j) \), we then have from (3.1)

\[
Q_{\text{GDO}} = S^2 + \frac{2}{n} \left( \sum_a \frac{a^2}{\sigma^2} \right) \sum_{s \subseteq i} Z_i^2 - \frac{2}{n(n-1)} \left( \sum_{s \subseteq i} \sum_{s \subseteq j} \sum_{i \neq j} \frac{a_i a_j}{\sigma^2} \right) \sum_{s \subseteq i} \sum_{s \subseteq j} Z_i Z_j \\
+ 2 \left[ \sum_{s \subseteq i} \sum_{s \subseteq j} \frac{a_i a_j}{\sigma^2} \right] \bar{Z}_s
\]

which reduces, under the condition of \( p\xi\)-unbiasedness to \( Q_{\text{GDO}} \), given by (2.1). Thus \( Q_{\text{GDO}} \) is the minimum \( p\xi\)-variance \( p\xi\)-unbiased estimator of \( \mathbb{E} S^2_Y \).

Finally, note that formula (2.7) of Theorem 4.1 applies with \( \mathbb{E}(S^2_Y) = \tau^2 \). Now \( Q \) is \( p \)-unbiased for \( S^2_Y \) and therefore \( p\xi\)-unbiased for \( \mathbb{E}(S^2_Y) \).

As \( \mathbb{E}(Q - \mathbb{E}(S^2_Y)^2 \) is minimized, for any given FES(n) design, by \( Q_{\text{GDO}} \), and \( \mathbb{E}(Q - \mathbb{E}(S^2_Y))^2 \simeq 4\mu^2 \sigma^2 K(a, \rho) \). The latter equation holds in particular for \( p = p_0 \) given by (2.1) and \( Q_{\text{GDO}} \) is
p-unbiased for $p_0$. Therefore,

$$\frac{\%}{\%}E(Q - \%S_0^2)^2 \geq 4\%^2 \%^2 K(a, \%$$

for any p-unbiased strategy $(p, Q)$, with equality if only if $(p, Q) = (p_0, Q_{GDO})$.

Remark 4.4. Recently Bhattacharyya (1993) also relaxed the condition of normality and quadraticity and obtained optimality of their (1989) predictor in the entire class of p-unbiased predictors $\mathcal{M}_p$.

Remark 4.5. Theorem 4.2 can be shown to hold under the class of discrete distribution $\xi$ covered by the Random permutation model, Model $E_{RP}$, considered in Mukhopadhyay (1984).