5.1. INTRODUCTION

A general theory of prediction that includes quadratic form (the population variance in particular) has been formulated, under a more general model, by Rodrigues et al. (1985). They have derived, assuming normality, optimal predictors in the class of homogeneous quadratic $\xi$-unbiased predictors. In this chapter, we extend this class of predictors under Model $G_T$ and show that the optimal $p\xi$-unbiased predictor, obtained in Chapter 4, can be improved upon if only $\xi$-unbiasedness, and not $p$-unbiasedness, is required.

We propose an optimal predictor, that is, predictor which minimizes

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

(1.1)

for $S^2_Y$ by choosing the sampling design in advance. Here we restrict ourselves to the class of quadratic $\xi$-unbiased
predictors but the sampling design, though fixed in advance, need not be taken to be of fixed size.

5.2. OPTIMAL $\xi$-UNBIASED PREDICTION: MODEL $G_T$

We shall assume that noninformative sampling design $p$ with expected effective sample size $n(s)$ is given, that is, the sample is given. Hence for given $s \in S$, we can write

$$S^2_Y = A(s, Y) + B(s, Y) \quad (2.1)$$

where

$$A(s, Y) = c \sum_{i \in s} Y^2_i - c \sum_{s \in S} \sum_{i \in j} Y_i Y_j$$

$$B(s, Y) = c \sum_{i \in s} Y^2_i - c \left( \sum_{s \in S} \sum_{i \in j} Y_i Y_j - \sum_{s \in S} \sum_{i \in j} Y_i Y_j \right)$$

and

$$s = \{1, \ldots, N\} - s.$$

Theorem 5.1. Among all quadratic $\xi$-unbiased predictors, under Model $G_T$, of $S^2_Y$ based on a given $p$, the one, say $Q^* = Q^*(s, Y)$, which minimizes (1.1) is given by

$$Q^* = A(s, Y) + U^*(s, Y), \quad s \in S \quad (2.2)$$

where

$$U^*(s, Y) = \frac{1}{n(s)} \left( \sum_{i \in s} a_i^2 \right) \sum_{i \in s} Z^2_i$$
\[
\frac{C_2}{n(s)[n(s)-1]} \left( \sum_{i,j} a_{ij} - \sum_{i} a_i a_i \right) \sum_{i,j} Z_i Z_j \\
+ 2 \left[ c \sum_{i,j} a_{ij} b_{ij} - c \left( \sum_{i} a_i b_i - \sum_{i,j} a_{ij} b_{ij} \right) \right] \overline{Z} + B(\bar{s}, \bar{b})
\]

and

\[
Z_i = (Y_i - b)/a_i \quad (i = 1, \ldots, N)
\]

\[
\overline{Z} = \frac{\sum Z_i}{n(s)}
\]

\[
B(\bar{s}, \bar{b}) = c \sum_{i} b_i^2 - c \left( \sum_{i,j} b_{ij} - \sum_i b_i b_i \right)
\]

**Proof.** Without loss of generality any quadratic predictor \(Q(s, Y)\) for \(S^2_Y\) can be expressed, for any given \(s \in S\), as

\[
Q(s, Y) = A(s, Y) + U(s, Y)
\]

with

\[
U(s, Y) = u_0 + \sum_{i} u(s, ii)Y_i^2 + \sum_{i,j} u(s, ij)Y_i Y_j
\]

where \(u_0\), \(u(s, ii)\) and \(u(s, ij) = u(s, ji)\) are constants. Thus the purpose is to predict \(B(\bar{s}, \bar{Y})\) from the given data using the predictor \(U(s, Y)\). Since \(Q(s, Y)\) is assumed to be a \(\xi\)-unbiased for \(S^2_Y\), we have

\[
\sum_{s} \left[ U(s, Y) - B(\bar{s}, \bar{Y}) \right] = 0, \quad s \in S
\]

from which using definition of \(B(\bar{s}, \bar{Y})\) and moments of \(Y_i\) under Model \(G\) we obtain
\[
\mathcal{E}(U) = (\sigma^2 + \mu^2) \left( \sum_{i \in S} a_i^2 - (\rho \sigma^2 + \mu^2) \left( \sum_{i \in S} a_i - \sum_{s \in \mathcal{S}} a_s \right)^2 \right) \\
+ 2\mu [c \sum_{i \in S} a_i b_i - c \left( \sum_{i \in S} a_i b_i - \sum_{s \in \mathcal{S}} a_s b_s \right)] \\
+ B(s, b). \\
(2.6)
\]

Note that (1.1) with the help of (2.1) and (2.4) simplifies to

\[
\mathcal{E}(\mathcal{Q} - \mathcal{S}^2) = E[\mathcal{Y}(U) + \mathcal{V}(B(s, Y)) - 2\mathcal{Cov}(U, B(s, Y))] \\
(2.7)
\]

We now analyse the terms of the above expression. To show that \( \mathcal{Cov}(U, B(s, Y)) \) is a design independent quantity let us assume without loss of generality that \( b_i = 0 \quad \forall \ i \) in Model \( G_1 \). Then (2.5) and (2.6) yield

\[
\sum_{s} u(s, ii) a_i^2 = c \sum_{i \in S} a_i^2 \\
\sum_{s} u(s, ij) a_i a_j = - c \left( \sum_{i \in S} a_i - \sum_{s \in \mathcal{S}} a_s \right)^2 \\
\]

which in turn implies, using (10.13) of Kendall et al. (1987), that

\[
\mathcal{Cov}(U, B(s, Y)) \cong 4\mu^2 \sigma^2 \rho \left[ c \sum_{i \in S} a_i^2 - c \left( \sum_{i \in S} a_i - \sum_{s \in \mathcal{S}} a_s \right)^2 \right] \\
(2.8)
\]

which is independent of the choice \( u \)'s. Therefore, the \( \mathcal{MSE} \) given by (2.7) is minimized if, for every fixed \( s \in S \), we choose \( U \) to minimize \( \mathcal{V}(U) \), subject to (2.6), where \( \mu, \sigma^2 \) and
\( \rho \) are the unknown quantities.

By generalized least square theory, assuming \( \rho \) to be known, the uniformly minimum \( \xi \)-variance, \( \xi \)-unbiased estimators of \( \mu \), \( \mu^2 \) and \( \sigma^2 \) (see Arnold (1979)) are

\[
\hat{\mu} = \frac{\bar{Z}}{s}
\]

\[
\hat{\mu}^2 = \frac{2}{\bar{Z}} - \frac{1 + \{n(s)-1\} \rho}{n(s)(1-\rho)} \frac{\bar{Z}^2}{s}
\]

\[
\hat{\sigma}^2 = \frac{1}{1-\rho} \frac{\bar{Z}^2}{s}
\]

where

\[
\frac{\bar{s}^2}{\bar{Z}} = \frac{1}{n(s)} \sum_{s} \frac{Z^2}{s} - \frac{1}{n(s)(n(s)-1)} \sum_{s} \sum_{i,j} Z_i Z_j
\]

and consequently the uniformly minimum \( \xi \)-variance \( \xi \)-unbiased estimators of \( \sigma^2 + \mu^2 \) and \( \rho \sigma^2 + \mu^2 \) (see Rao (1973), p. 318) are

\[
\sigma^2 + \mu^2 = \frac{1}{n(s)} \sum_{s} \frac{Z^2}{s}
\]

\[
\rho \sigma^2 + \mu^2 = \frac{1}{n(s)(n(s)-1)} \sum_{s} \sum_{i,j} Z_i Z_j
\]

Substituting these estimators in (2.6), we get \( U = U^*(s, Y) \), as given in (2.3), the optimal predictor of \( B(s, Y) \) and hence this results in the predictor \( Q^*(s, Y) \) as stated in (2.2). As \( Q^* \) ultimately does not contain \( \rho \), it does not matter
whether $\rho$ is known or not known. Clearly, $Q^*$ minimizes $\mathbb{E}(Q - S^2_Y)^2$ for every $s \in S$ and hence also $\mathbb{E}(Q - S^2_Y)$. Thus the theorem is proved.

Remark 5.1. In various interesting special cases of Model $G_T$, the predictor $Q^*$ given by (2.2) can be expressed as follows:

(i) If $a_i = 1, b_i = 0 \ \forall \ i$, then the optimal predictor is

$$Q^*_i = s^2_Y,$$

the sample variance.

(ii) If $a_i = 1 \ \forall \ i$, then $Q^*$ reduces to

$$Q^*_i = s^2_Y + s^2_b - s^2_b - \frac{2(N-n)}{N-1} \{s_Y - s^2_b\}$$

where

$$s^2_b = \frac{1}{n(s)} \sum_{s} Y_{b,i} - \frac{1}{n(s)(n(s)-1)} \sum_{s,i,j} Y_{b,i} Y_{b,j}$$

(iii) If $b_i = 0 \ \forall \ i$, then

$$Q^*_i = \Lambda(s, Y) + \frac{c}{n(s)} \left( \sum_{s} a^2_i \sum_{s} \frac{Y^2_i}{a^2_i} \right) - \frac{c^2}{n(s)(n(s)-1)} \left( \sum_{s} a^2_i \sum_{s} a^2_j \sum_{s} \frac{Y_{i,j}}{a_i a_j} \right)$$

which is the predictor obtained by Mukhopadhyay and Bhattacharyya (1989).
5.3. OPTIMALITY OF THE SUGGESTED PREDICTOR

We show below that the predictor

$$Q_0 = \frac{c_1}{n(s)} \left(\sum a_i^2 \right) \frac{Y_i^2}{a_i^2} - \frac{c_2}{n(s)(n(s)-1)} \left(\sum \sum a_i a_j \right) \frac{Y_i Y_j}{a_i a_j}$$

(3.1)

which is $\xi$-unbiased for $S_Y^2$, can be improved upon, for any fixed design, if $\xi$-unbiasedness, but not $p$-unbiasedness is required. Note that if $p$ is fixed effective size $n$ design with $\pi_i \propto a_i^2$ and $\pi_{ij} \propto a_i a_j$, then $Q_0$ is optimal in the class of all homogeneous quadratic $p$-unbiased predictors of $S_Y^2$ (see Remark 4.3 (iii) of Chapter 4).

Theorem 5.2. Under Model $G$, with $b_i = 0 \ \forall \ i$, for any design $p$,

$$\mathbb{E}(Q_0 - S_Y^2)^2 - \mathbb{E}(Q_0^* - S_Y^2)^2 \geq 0$$

where $Q_0$ and $Q_0^*$ are given by (3.1) and (2.9) respectively.

Proof. Since $b_i = 0$, we have

$$Z_i = Y_i/a_i \quad i = 1, \ldots, N$$

Define

$$w_{ij} = a_i a_j - \frac{1}{n(s)(n(s)-1)} \sum \sum a_i a_j \quad i, j \in s \quad (3.2)$$

73
Note that
\[ \sum_{s \in t} w_i = 0 \quad \text{and} \quad \sum_{s \in ij} w_i = 0 \quad (3.3) \]

Using (10.12) and (10.13) of Kendall et al. (1987), we have, under Model \( G \) with \( b_i = 0 \) for all \( i \),

\[
\begin{align*}
Y(Z_i^2) &= 4\mu^2 \sigma^2 \\
\mathbb{E}(Z_i^2, Z_j^2) &= 4\mu^2 \sigma^2 \rho \\
Y(Z_i, Z_j) &= 2\mu^2 \sigma^2 (1 + \rho) \\
\mathbb{E}(Z_i, Z_j, Z_k) &= \mu^2 \sigma^2 (1 + 3\rho) \\
\mathbb{E}(Z_i^2, Z_j^2) &= 4\mu^2 \sigma^2 \rho \\
\mathbb{E}(Z_i, Z_j, Z_k) &= 4\mu^2 \sigma^2 \rho \\
\mathbb{E}(Z_i^2, Z_j, Z_k) &= 2\mu^2 \sigma^2 (1 + \rho)
\end{align*}
\]

\{ (3.4) \}

Observe that

\[
Q^*_0 = Q_0 + \Lambda(s, Y) - \left[ \frac{c_1}{n(s)} \sum_s a_s^2 \sum_s Y_i^2 \right]
\]

\[
- \frac{c_2}{n(s)[n(s)-1]} \left( \sum_s a_s \sum_s Y_i Y_j \right)
\]

\[
= Q_0 + c \sum_{i \in s} \left[ a_i^2 - \frac{1}{n(s)} \sum_s a_s^2 \right] Z_i^2
\]

\[
- c_2 \sum_s \left[ a_{ij} - \frac{1}{n(s)[n(s)-1]} \sum_{s \in ij} a_s \right] Z_i Z_j
\]

\[
= Q_0 + Q_i \quad \text{say} \quad (3.5)
\]

where

\[
Q_i = c_1 \sum_{s \in ij} Z_i^2 - c_2 \sum_{s \in ij} Z_i Z_j
\]
From (3.5) we have

\[ \mathbb{E}(Q_3^* - S_Y^2)^2 = \mathbb{E}(Q_0 - S_Y^2)^2 + \mathbb{E}(Q_1^2) + 2\mathbb{E}(Q_1 Q_0) - 2\mathbb{E}(Q_1 S_Y^2) \]

Using the fact that the moments of transformed variables Z's, given in (3.4), are functions of parameters and noting (3.3), one can see that

\[ \mathbb{E}(Q_0 Q_1) = 0 \quad \text{and} \quad \mathbb{E}(Q_1^2) = \mathcal{V}(Q_1) \]

Thus, using (2.1), we have

\[
\begin{align*}
\mathbb{E}(Q_3^* - S_Y^2)^2 &= \mathbb{E}(Q_0 - S_Y^2)^2 + \mathbb{E}(Q_1^2) - 2\mathbb{E}(Q_1 A(s, Y)) - 2\mathbb{E}(Q_1 B(s, Y)) \\
&= \mathbb{E}(Q_1^2) - 2\mathbb{E}(Q_1 A(s, Y)) - 2\mathbb{E}(Q_1 B(s, Y)) \\
&= \mathbb{E}(Q_1^2) \\
&= 3.6 \tag{3.6}
\end{align*}
\]

Again using (3.3) and (3.4), we have

\[
\begin{align*}
\mathbb{E}\left[ Q_1 A(s, Y) \right] &= \mathbb{E}\left[ Q_1 + \frac{c_1}{n(s)} \left( \sum a_i^2 \right) \sum Z_i^2 \\
&\quad - \frac{c_2}{n(s)[n(s)-1]} \left( \sum a_i a_j \right) \sum Z_i Z_j \right] \\
&= \mathbb{E}(Q_1^2) \\
&= \mathbb{E}(Q_1^2) \\
&= 3.7 \tag{3.7}
\end{align*}
\]

Next, proceeding as in (2.8) and noting \( \sum w_{ii} = 0 \), \( \sum w_{ij} = 0 \), one can see that

\[ \mathbb{E}(Q_1 B(s, Y)) = 0 \quad \tag{3.8} \]
Thus, using (3.7) and (3.8) in (3.6), we have

\[ \mathbb{E}(Q_0 - S^2)^2 - \mathbb{E}(Q_* - S^2)^2 = \mathbb{E} \mathcal{Y}(Q) \geq 0 \]

as \( \mathcal{Y}(Q) = \mathcal{Y}(Q_1) \). Which proves the theorem.

The expression for \( \mathcal{Y}(Q_1) \) can be obtained as under

\[
\mathcal{Y}(Q_1) = c^2 \left[ \sum_{s \in I} w^2_{s} \mathcal{Y}(Z^2) + \sum_{s \in I} w_{s} \text{cov} (Z^2, Z^2) \right]
\]

\[- c^2 \left[ \sum_{s \in I} w_{s} \mathcal{Y}(Z) + \sum_{s \in I} w_{s} \text{cov} (Z, Z) \right]
\]

\[- 2c \sum_{s \in I} w_{s} \text{cov} (Z, Z) \]

Using (3.4) and noting

\[ 0 = \left( \sum_{s \in I} w_{s} \right)^2 = \sum_{s \in I} w^2_{s} + \sum_{s \in I} w_{s} \]

\[ 0 = \left( \sum_{s \in I} w_{s} \right)^2 = 2 \sum_{s \in I} w^2_{s} + \sum_{s \in I} w_{s} \text{cov} \]

\[ 0 = \left( \sum_{s \in I} w_{s} \right) \left( \sum_{s \in I} w_{s} \right) = \sum_{s \in I} w_{s} \text{cov} \]

\( \mathcal{Y}(Q_1) \) is given by
\[ Y(Q_1) = c_i^2 \left( 4 \mu^2 \sigma^2 \sum_{s} w^2_{i} - 4 \mu^2 \sigma^2 \rho \sum_{s} w^2_{ii} \right) \]

\[ + c_i^2 \left[ 4 \mu^2 \sigma^2 (1+\rho) \sum_{s} w^2_{ij} + 4 \mu^2 \sigma^2 (1+3\rho) \sum_{s} w_{ij} w_{ik} \right. \]

\[ - 4 \mu^2 \sigma^2 \rho (2 \sum_{s} w^2_{ij} + 4 \sum_{s} w_{ij} w_{ik}) \]

\[ - 2c_i c_j \left[ 6 \mu^2 \sigma^2 (1+\rho) \sum_{s} w_{ii} w_{ij} - 2 \mu^2 \sigma^2 (1+5\rho) \sum_{s} w_{ii} w_{ij} \right] \]

\[ = 4 \mu^2 \sigma^2 (1-\rho) \left[ c_i^2 \sum_{s} w^2_{ii} + c_j^2 \sum_{s} w^2_{ij} + c_j^2 \sum_{s} w_{ij} w_{ik} \right. \]

\[ - 2c_i c_j \sum_{s} w_{ii} w_{ij} \]

5.4. OPTIMAL \( \xi \)-UNBIASED PREDICTION: MODEL \( G_R \)

We now consider the Model \( G_R \), assuming that \( x_i > 0 \) (i = 1, ..., N) are known auxiliary variable values. The following theorem gives the \( \xi \)-best quadratic unbiased predictor (\( \xi \)-BQU) of \( S^2_Y \) for the model under consideration.
Theorem 5.3. Under Model $G$, the $\xi$-BQU predictor of $S^2_Y$ is,
for any design $p$, given by

$$Q^* = A(s, Y) + \beta^2 B(s, x) + \sigma^2 c \sum_i v(x_i)$$  \hfill (4.1)

where

$$\beta = \frac{\sum_s x_i Y_i}{\sum_s v(x_i)}$$  \hfill (4.2)

$$\hat{\beta}^2 = \beta^2 - \frac{\sigma^2}{\sum_s x_i^2 / v(x_i)}$$  \hfill (4.3)

$$\hat{\sigma}^2 = \frac{1}{n(s)-1} \sum_s \frac{(Y_i - \hat{\beta}x_i)^2}{v(x_i)}$$  \hfill (4.4)

Proof. Note that

$$\delta(S^2_Y) = \delta(A(s, Y) + \beta^2 B(s, x) + \sigma^2 c \sum_i v(x_i))$$  \hfill (4.5)

where $\beta^2$ and $\sigma^2$ are unknown quantities.

By generalized least squares theory, the minimum $\xi$-variances quadratic $\xi$-unbiased estimators of $\beta^2$ and $\sigma^2$, for a given $s$, are given by $\hat{\beta}^2$ and $\hat{\sigma}^2$, formulae (4.3) and (4.4) respectively. Substitution of these estimators results in the predictor $Q^*$ as stated in (4.1).

It follows from Theorem 5.3 that $\delta \text{MSE}(p, Q^*) \leq$
\( \text{MSE}(p, Q_{\text{greg}}) \) for any fixed design \( p \) common to the two strategies, where \( Q_{\text{greg}} \) is given by (1.4) of Chapter 6.

The above result can be seen as a special case of Rodrigues et al. (1985).

Remark 5.2. If \( v(x_i) = x_i^2 \) for all \( i \), (4.1) becomes

\[
Q^*_{\text{I}} = A(s, Y) + \frac{c_1}{n(s)} \left( \sum_{i} x_i^2 \right) \sum_{i} \frac{Y_i^2}{x_i^2} - \frac{c_2}{n(s)[n(s)-1]} \left( \sum_{i} x_i \sum_{i} x_i - \sum_{i} x_i \sum_{i} x_i \right) \sum_{i} \frac{Y_i Y_i}{x_i x_i}
\]

which is the predictor obtained by Mukhopadhyay et al. (1990-91).

Remark 5.3. If \( p \) is an FES(n) design then the predictor \( Q^*_{\text{I}} \) given by (2.2), which is the minimum \( \text{MSE} \) quadratic \( \xi \)-unbiased predictor for a given design \( p \) under Model \( G_T \), is identical to \( Q^*_{\text{I}} \) provided \( a_i = x_i \) and \( b_i = 0 \) for \( i = 1, \ldots, N \) in Model \( G_T \).