CHAPTER III

MISCLASSIFICATION UNDER FULLY-CURTAILED DOUBLE SAMPLING PLAN BY ATTRIBUTES

3.1 In this chapter we consider the problem of misclassification under fully-curtailed DSP introduced in Section 2.3 of Chapter II. The MLEs of the fraction defective and the probability of misclassification are obtained. The asymptotic variances and covariance of the MLEs are also derived. Two particular cases for the extreme values of the probability of misclassification are also discussed at the end of the chapter.

3.2 Description of Misclassification :

3.2.1 In lot-by-lot acceptance sampling plan rejection of a lot involves sometimes a botheration. An immediate consequence of the rejection of a lot leads to the inspection of all the units of a lot, when screening is prevailing. Another consequence is that the rejection of a lot creates undue doubt about the quality of the units produced. For these reasons, the concerned person may avoid the rejection of a lot by misclassifying a defective as a nondefective.

This tendency may be obvious when the lot could be rejected because of finding just the minimum number of defectives necessary for the rejection.

3.2.2 Misclassification under Single Sampling Plan :

In usual single sampling plan, we know that a lot is rejected if the number of defectives observed in n inspected units is a + 1 (a being the acceptance number) or more. When an inspector finds exactly a + 1 defectives in n inspections, he will be inclined to classify a defective as a nondefective. This will lead to the acceptance of a rejectable lot. Cohen [6] has considered this type of misclassification. He has obtained the maximum likelihood estimates of the fraction defective and the probability of misclassification when data of uncurtailed single sampling plan are subject to this type of misclassification. He has also given the asymptotic variances and covariance of these estimates.

In case of curtailed single sampling plan the question of 100% inspection of a rejectable lot does not arise, hence the purpose of misclassification is to avoid undue doubt about the quality of units which may arise due to the rejection of a lot. Phatak [41] has studied the problem of misclassification under semi-curtailed and fully-curtailed single sampling plans and gave similar results.

3.2.3 Misclassification under Fully-Curtailed DSP :

We consider the problem of similar type of misclassification under fully-curtailed DSP introduced in chapter II (Section 2.3). In case of fully-curtailed DSP one may think of four possible situations under which an inspector may misclassify a defective as a nondefective which leads to the acceptance of a lot. These situations are given below :

(i) During the inspection of the first sample inspector observes (g_1-1) nondefectives in y inspections, $g_1-1 \leq y \leq n_1-2$ and a defective at the $(y+1)$th inspection.

(ii) During the inspection of first sample inspector observes (n_1-g_1) defectives in (n_1-1) inspections and a defective at the n_1th inspection.

(iii) During the inspection of second sample inspector observes (g_2-1) nondefectives in y inspections, $g_2-1 \leq y \leq n_1+n_2-2$, and a defective at the $(y+1)$th inspection.

(iv) During the inspection of second sample inspector observes (r_2-1) defectives in (n_1+n_2-1) inspections and a defective at the (n_1+n_2)th inspection.

In all these situations a lot will be accepted by misclassifying a defective as a nondefective which appears at the last inspection. Furthermore, misclassification of a defective as a nondefective in situations (i) and (iii) leads to the curtailment of the inspection at the acceptance stage during the inspection of first and second sample respectively. Misclassification under situation (ii) avoids the inspection of second sample. Misclassification under situation (iv) leads to the acceptance of a rejectable lot. We discard first three situations under the assumption that misclassification should not lead to any curtailment in the inspection or avoid the inspection of second sample. This argument is based on the fact that the inspector is not too disloyal to report a defective as a nondefective so that his misclassification would lead to the curtailment of the inspection or avoid the inspection of the second sample. Hence we consider only situation (iv) in the matter that follows. Let the inspector misclassify a defective as a nondefective with probability θ. Furthermore, it is assumed that the inspector gives complete information about the sampling inspection. We have obtained the MLE of the fraction defective and of θ, and asymptotic variance-covariance matrix of these estimators.

3.3 Probability Function of the Fully-Curtailed DSP under the Misclassification :

3.3.1 Statement of the Fully-Curtailed DSP under the Misclassification :

The statement of the fully-curtailed DSP under the misclassification described in situation (iv) is given below:

Consider an attributes acceptance plan in which individual units randomly selected from a lot of size N are inspected one at a time till one of the following six events occurs :

(e_1) g_1 nondefectives are observed and the number of units inspected is greater than n_0 and less than or equal to n_1,

(e_2) g_2 nondefectives are observed and the number of units inspected is greater than n_1 and less than or equal to (n_1+n_2-1),

(e_3) g_2 nondefectives are observed and the number of units inspected is equal to n_1+n_2 or r_2th defective is observed at (n_1+n_2)th inspection and it is misclassified as a nondefective,

(e_4) r_1 defectives are observed and the number of units inspected is greater than n_o and less than or equal to n_1,

(e_5) r_2 defectives are observed and the number of units inspected is greater than n_1 and less than or equal to (n_1+n_2-1),

(e_6) r_2th defective is observed at (n_1+n_2)th inspection.

Here n_o is assigned a value zero.

The decision rule is then to accept the lot if one of the three events, e_1, e_2 and e_2 occurs and to reject the lot if one of the three events, e_4, e_5 and e_6 occurs.

3.3.2 Probability Function :

Let the process average proportion of defectives be p and for sufficiently large lot it can be considered as the probability of selecting a defective in a single trial. Furthermore, let the probability p remain constant from trial to trial and the trials be stochastically independent. This applies to the Type B situation of Dodge and Romig [10], hence the lot size does not subsequently appear.

Let Y denote the number of units inspected when the inspection is stopped due to the occurrence of the events e_i, i=1,2,3,4,5,6. Let T_i (i=1,2,3,4,5,6) be the set of possible values attained by Y. Then

$$T_1 = \{g_1, g_1+1, \ldots, n_1\},$$

$$T_2 = \{g_2-g_1+n_1+1, \ldots, n_1+n_2-1\},$$

$$T_3 = \{n_1+n_2\},$$

$$T_4 = \{r_1, r_1+1, \ldots, n_1\},$$

$$T_5 = \{r_2-r_1+n_1+1, \ldots, n_1+n_2-1\},$$

$$T_6 = \{n_1+n_2\}.$$

Further define a random variable I as follows :

$$I = i \text{ if } e_i \text{ occurs, } i = 1,2,3,4,5,6.$$

Then the joint probability function of the random variables Y and I can be expressed as

$$P(Y=y,\ I=i) = t_i(y;p,\theta) \qquad y \in T_i,\ i=1,2,3,4,5,6$$
$$= 0 \text{ elsewhere} \qquad \ldots(3.3.1)$$

where

$$t_1(y;p,\theta) = \binom{y-1}{g_1-1} p^{y-g_1} q^{g_1} \qquad \ldots(3.3.2)$$

$$t_2(y;p,\theta) = \sum_{u=1}^{b_1} \binom{n_1}{g_1-u} \binom{y-n_1-1}{g_2-g_1+u-1} p^{y-g_2} q^{g_2} \qquad \ldots(3.3.3)$$

$$t_3(y;p,\theta) = \sum_{u=1}^{b_1} \binom{n_1}{g_1-u} \binom{n_2-1}{g_2-g_1+u-1} p^{n_1+n_2-g_2} q^{g_2} \left\{1+\frac{\theta p}{q}\right\} \qquad \ldots(3.3.4)$$

$$t_4(y;p,\theta) = \binom{y-1}{r_1-1} p^{r_1} q^{y-r_1} \quad \ldots(3.3.5)$$

$$t_5(y;p,\theta) = \sum_{u=1}^{b_1} \binom{n_1}{g_1-u}\binom{y-n_1-1}{b_2-u} p^{r_2} q^{y-r_2} \quad \ldots(3.3.6)$$

$$t_6(y;p,\theta) = (1-\theta) \sum_{u=1}^{b_1} \binom{n_1}{g_1-u}\binom{n_2-1}{b_2-u} p^{r_2} q^{n_1+n_2-r_2} \quad \ldots(3.3.7)$$

and $q=1-p$, $0 \leq p \leq 1$, $0 \leq \theta \leq 1$, $b_1=g_1+r_1-n_1-1$,
$b_2=g_1+r_2-n_1-1$.

The remark given below the probability function in Section 2.4.1 of Chapter II also holds here. That is while calculating the various terms of the summation involved in (3.3.3), (3.3.4), (3.3.6) and (3.3.7), $\binom{n}{x}$ is regarded as zero whenever x exceeds n or x is negative.

3.4 The Maximum Likelihood Estimates of the Fraction Defective and Probability of Misclassification :

Let m lots have undergone the inspection under fully-curtailed DSP subject to the misclassification described under situation (iv). Let m_i, for i=1,2,3, be the number of lots accepted and let m_i, for i=4,5,6, be the number of rejected lots. Clearly $m = \sum_{i=1}^{6} m_i$. The m pairs given by

$$(y_{ij},\ I=i) \qquad j=1,2,\ldots,m_i,\quad i=1,2,3,4,5,6. \quad \ldots(3.4.1)$$

where $y_{ij} \in T_i$, $j=1,2,\ldots,m_i$ for fixed i, can be considered as a random sample of size m from a bivariate distribution whose probability function is given by (3.3.1). The likelihood function, L, based on this sample can be expressed as

$$L = \prod_{i=1}^{6} \prod_{j=1}^{m_1} t_i(y;p,\theta)$$

$$=(\text{const.}) \prod_{j=1}^{m_1} [p^{y_{1j}-g_1} q^{g_1}] \prod_{j=1}^{m_2} [p^{y_{2j}-g_2} q^{g_2}]$$

$$\prod_{j=1}^{m_3} [q^{g_2} p^{n_1+n_2-g_2} \{ 1+ \frac{\theta p}{q} \}] \prod_{j=1}^{m_4} [p^{r_1} q^{y_{4j}-r_1}]$$

$$\prod_{j=1}^{m_5} [p^{r_2} q^{y_{5j}-r_2}] \prod_{j=1}^{m_6} [(1-\theta) p^{r_2} q^{n_1+n_2-r_2}] \qquad \ldots(3.4.2)$$

where we use (3.3.2) through (3.3.7) to obtain (3.4.2). Note that $y_{3j} = n_1+n_2$ for $j=1,2,\ldots m_3$ and $y_{6j}=n_1+n_2$ for $j=1,2,\ldots,m_6$. Taking logarithm of (3.4.2), we get

$$\log L = \log (\text{const.}) + (TD) \log p +(TND)\log q$$

$$+ m_3 \log (1+\theta p/q) + m_6 \log (1-\theta). \qquad \ldots(3.4.3)$$

where (TD) = Total number of defective units observed when m lots have undergone the inspection.

$$= \sum_{j=1}^{m_1} (y_{1j}-g_1) \sum_{j=1}^{m_2} (y_{2j}-g_2)+m_3(n_1+n_2-g_2)$$

$$+m_4 r_1+m_5 r_2 + m_6 r_2 \qquad \ldots(3.4.4)$$

(TND)= Total number of nondefective units observed when m lots have undergone the inspection.

$$= m_1 g_1 + m_2 g_2 + m_3 g_2 + \sum_{j=1}^{m_4} (y_{4j} - r_1)$$

$$+ \sum_{j=1}^{m_5} (y_{5j} - r_2) + m_6 (n_1 + n_2 - r_2) \quad \ldots(3.4.5)$$

Differentiating (3.4.3) with respect to p and θ and equating the partial derivatives to zero, the maximum likelihood estimates of p and θ are

$$\hat{p} = \frac{(TD) - m_6}{(TU) - (m_3 + m_6)} \quad \ldots(3.4.6)$$

$$\hat{\theta} = \frac{m_3 \hat{p} - m_6 (1 - \hat{p})}{\{m_3 + m_6\} \hat{p}} \quad \ldots(3.4.7)$$

where (TU) = Total number of units inspected during the inspection of m lots.

= (TD) + (TND).

3.5 Asymptotic Variances and Covariance of the MLEs :

We need the following expectations to compute the asymptotic variances and covariance of the MLEs.

$$E(TD)/m = p \{ ASN - t_3(n_1 + n_2; p, \theta)/A \} \quad \ldots(3.5.1)$$

$$E(TU)/m = ASN \quad \ldots(3.5.2)$$

where the expression for ASN is given by (2.4.16) of Chapter-II, and $A = p + q/\theta$.

Now

$$-E(\partial^2 \log L/\partial p^2)/m = \frac{ASN}{pq} + \frac{t_3(n_1+n_2;p,\theta)}{q^2 A}\left\{\frac{1}{A} - \frac{1}{p}\right\}$$

$$= \phi_{11} \qquad \ldots(3.5.3)$$

$$-E(\partial^2 \log L/\partial p \partial \theta) = -\frac{t_3(n_1+n_2;p,\theta)}{A^2\theta^2}$$

$$= \phi_{12} = \phi_{21} \qquad \ldots(3.5.4)$$

$$-E(\partial^2 \log L/\partial \theta^2)/m = \frac{t_6(n_1+n_2;p,\theta)}{(1-\theta)^2} + \frac{p^2 t_3(n_1+n_2;p,\theta)}{A^2\theta^2}$$

$$= \phi_{22} \qquad \ldots(3.5.5)$$

Consider matrix M as

$$M = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \qquad \ldots(3.5.6)$$

Hence the variance-covariance matrix of $\hat{p}$ and $\hat{\theta}$ is given by $(M^{-1})/m$. The asymptotic variances and covariance are

$$V(\hat{p}) = \phi_{22}/mD \qquad \ldots(3.5.7)$$

$$V(\hat{\theta}) = \phi_{11}/mD \qquad \ldots(3.5.8)$$

$$Cov(\hat{p}, \hat{\theta}) = -\phi_{12}/mD \qquad \ldots(3.5.9)$$

where $D = |M| = (\phi_{11}\phi_{22} - \phi_{12}^2)$

3.6 Particular Cases $\theta = 1$ and $\theta = 0$:

3.6.1 Case $\theta = 1$

When $\theta = 1$, misclassification is carried with certainty. In this case event e_6 will not occur and hence the observed frequency, m_6 will be zero. Here the problem will reduce to the estimation of only one parameter, p. The probability function can be obtained by substituting $\theta=1$ in (3.3.1). The maximum likelihood estimate of p and asymptotic variance of the MLE can be derived from this probability function in the usual way. They are given below :

$$\hat{p} = \frac{(TD)'}{(TU)'-m_3'} \qquad ...(3.6.1)$$

and

$$V(\hat{p}) = \frac{pq}{ASN-t_3'(n_1+n_2;p)} \qquad ...(3.6.2)$$

Where

(i) $(TU)' = (TD)' + (TND)'$

(ii) $$(TD)' = \sum_{j=1}^{m_1} (y_{1j}-g_1) + \sum_{j=1}^{m_2} (y_{2j}-g_2) + m_3'(n_1+n_2-g_2) + m_4 r_1 + m_5 r_2$$

(iii) $$(TND)' = m_1 g_1 + m_2 g_2 + m_3' g_2 + \sum_{j=1}^{m_4} (y_{4j}-r_1) + \sum_{j=1}^{m_5} (y_{5j}-r_2)$$

(iv) $t_3'(n_1+n_2;p) = t_3(n_1+n_2;p,1)$

(v) m_3' = frequency with which event e_3 occurs in this case.

3.6.2 Case $\theta = 0$

When $\theta=0$, it is the case of correct classification. In this case the problem of the estimation of parameter p reduces to the problem of the estimate of p under fully-curtailed DSP. This is already dealt in the Section 2.4 of Chapter II.

3.7 Numerical Example :

In this section we illustrate the results of this chapter by a numerical example. Table 3.1 gives the tabulation of 100 observations associated with the inspection of 100 lots under a fully-curtailed DSP. It is assumed that the data are subject to misclassification of the type discussed in this chapter. The plan is

$$n_1 = 5,\ n_2 = 10,\ r_1 = 3,\ r_2 = 5,\ g_1 = 4,\ g_2 = 11.$$

Table 3.1

Plan: $n_1=5$, $n_2=10$, $r_1=3$, $r_2=5$, $g_1=4$, $g_2=11$.

Event	Number of units inspected	Number of accepted lots	Event	Number of units inspected	Number of rejected lots
e_1	4	41	e_4	3	1
	5	33		4	2
				5	3
e_2	13	3			
	14	5	e_5	8	0
				9	0
e_3	15	5		10	1
				11	1
				12	1
				13	1
				14	1
			e_6	15	1

From the table above we find the following :

$m_1 = 74$ $m_4 = 6$ $(TD) = 122$

$m_2 = 8$ $m_5 = 5$ $(TU) = 614$

$m_3 = 5$ $m_6 = 1$

Substituting these values in the expressions (3.4.6) and (3.4.7) we get $\hat{p} = 0.199013$ and $\hat{\theta} = 0.162534$. It may be noted that the data of this example were obtained using model sampling with $\theta = 0.10$ and $p = 0.20$. Using these hypothetical values of p and θ we have

$$V(\hat{p}) = 0.0002558849$$

$$V(\hat{\theta}) = 0.6018047143$$

The absolute difference between $\hat{p}$ and p, and $\hat{\theta}$ and θ may be attributed due to sampling fluctuations since we observe that

$$|\hat{p} - p| = 0.061691 \text{ S.E. } (\hat{p})$$

and $|\hat{\theta} - \theta| = 0.080610$ S.E. $(\hat{\theta})$.

In practice, one may use $\hat{p}$ and $\hat{\theta}$ to compute the estimates of the asymptotic variances and covariance when one does not know the true values of p and θ. The Binomial probability distribution Tables [43] are used for the computation illustrated in this example.