CHAPTER – VII

PASP BY VARIABLES FOR A CLASS OF SYMMETRICAL THICK TAILED DISTRIBUTIONS

7.1. INTRODUCTION

Most of the literatures on acceptance sampling plans by variables are derived under the assumption that the quality characteristic of interest is normally distributed. The assumption of normality may not be realistic for certain populations. In such a case, the estimation of percentage defectives based on the sample mean and standard deviation, using a normal sampling plan, will not be the same when the distribution is non-normal with a thicker tail. (cf. Montgomery 1991).

Duncan (1965) proposed a procedure for using a variable sampling plan when the distribution of the quality characteristic follows an exponential distribution. Kocherlakota and Balakrishnan (1984) derived sampling plans based on modified maximum likelihood (MML) estimators of the location (µ) and scale (σ) parameters of the normal distribution, when the sample is censored. Since the MML estimators of µ and σ coincide with \( \bar{x} \) and S respectively when a complete sample is available, the plans will not be robust against departures from normality.
Situations may arise where the quality characteristic has a symmetric distribution that is not normal. The data may exhibit a somewhat thick tailed distribution. We have come across one such situation in a study carried out at a leading two-wheeler manufacturing company in India by Deshmukh (1995). The company produces CAM BUSH (24111011) for scooters. The raw material for this comes with oversized outer diameter (OD) and required inner diameter (ID). The outer diameter is reduced to required dimension. It was found that there are variations in OD of CAM after machining, which creates problems in further assembly. A statistical study was conducted to identify the cause of variations. It was found that the raw material received do not satisfy the hardness specification along with several other findings. The investigator had suggested to include an acceptance sampling plan as a check for the hardness of the raw material. Further, it was found that the hardness of the raw material follows a distribution which is symmetric and some what heavy tailed. Hence, to implement the investigators suggestion, one cannot use the normal sampling plan. But, a sampling plan based on a symmetric distribution which is some what thick-tailed need to be considered.

A location-scale family of such symmetric distributions is considered whose probability density function is given as.

\[ f(x) = \frac{1}{\sigma \sqrt{m \beta}} \left[ 1 + \frac{(x-\mu)^2}{m \sigma^2} \right]^{-p} - \alpha < x < \alpha \quad (7.1) \]

Where \( m = 2p - 3, \ p \geq 2 \); \( \beta(\ldots) \) denotes the beta function and \( \mu \) and \( \sigma \) being the location and scale parameters respectively. Note that this family includes a range of symmetric distributions such as
This chapter provides for a situation when the assumption of normality in sampling plan by variables is violated. Based on a more general symmetrical thick tailed family of distributions, as described earlier, section 7.3 provides for the estimation of parameters using the modified maximum likelihood (MML) procedures introduced by Tiku and Suresh (1992). Section 7.2 proposes a sampling plan with a pricing approach. The extent to which a sampling plan protects the interests of the producer and the consumer is measured by a pay-off ratio which has been used in determining the parameters of a sampling plan. Section 7.4 provides two different approaches for the determination of the plan parameters \( n \) and \( t_0 \).

### 7.2. STATEMENT OF THE PLAN

#### 7.2.1. Notations

The variable \( X \) has a density function \( f(x) \) given in (7.1) with \( \mu \) as the location parameter and known value of the scale parameter \( \sigma \). An item is called defective whenever \( x \leq 1 \), otherwise it is non-defective. For given \( l \) and \( \sigma \), the quality of the lot, as measured by the proportion of defective items in it, clearly depends on \( \mu \).

Let

\[ P_T(\mu) = \text{True price of a lot with mean } \mu. \]

\[ E(\mu) = \text{Average price paid by the consumer for the lot with mean } \mu \text{ on the basis of sample observations.} \]

\[ R(\mu) = E(\mu) / P_T(\mu) = \text{Pay-off ratio for a lot with mean } \mu. \]
7.2.2. The sampling plan

Draw a random sample of size n from the lot and obtain the estimate of the location parameter \( \mu \) using the method suggested by Tiku and Suresh (1992).

If \( \hat{\mu} \geq t_0 \), every item of the lot is paid \( P_1 \), otherwise pay price \( P_2 \), where \( t_0 \) is a suitable defined quantity.

7.3. THE MML ESTIMATION OF \( \mu \)

In several situations the maximum likelihood (ML) equation \( \frac{\partial \log L}{\partial \mu} = 0 \) have no explicit solution, which may be due to the fact that a few terms in the ML equation is intractable; see, for example Tiku (1967, 1968, 1988). This problem was overcome by Tiku and Suresh (1992) where they first expressed the ML equations in terms of order statistics and then replaced the intractable terms by their Taylor series expansion, to obtain the modified maximum likelihood (MML) equations.

Let \( x_1, x_2, \ldots, x_n \) be a random sample of size n from (7.1) and \( x_{(1)}, x_{(2)}, \ldots, x_{(n)} \) is the corresponding order statistic.

Then the MML estimator of \( \mu \) is given by

\[
\hat{\mu} = \sum_{i=1}^{n} \beta_i x_{(i)} / \sum_{i=1}^{n} \beta_i
\]

(7.2)

Where \( \beta_i = \frac{1 - t_i^2 / m}{[1 + t_i^2 / m]^2} \), \( i = 1, 2, \ldots, n \)

\[
t_i = E \left\{ \frac{x_{(i)} - \mu}{\sigma} \right\}
\]
The values of $t_{(i)}$ are tabulated for different values of $p$ in Tiku and Kumura (1991). David and Johnson (1954) have used the approximation of moments of order statistics to obtain $t_{(i)}$.

Consequently, the MML estimator $\hat{\mu}$ given above is asymptotically equivalent to the ML estimator and thus $\hat{\mu}$ is asymptotically unbiased and efficient. However, from the symmetricity of (7.1), $\hat{\mu}$ is unbiased for all $n$.

7.4. DETERMINATION OF PLAN PARAMETERS

Two different approaches have been provided for the determination of the plan parameters $n$ and $t_0$. Both the approaches specify two values of $R(\mu)$.

APPROACH – I

For $\mu = \mu_0$ and $\mu = \mu_1 (< \mu_0)$, the plan parameters $(n, t_0)$ are obtained in such a way that

$$R(\mu_0) = 1 \quad (7.3)$$

$$R(\mu_1) = d_1 (< 1) \quad (7.4)$$

where $d_1$ is usually specified by the producer to safeguard his interest when the process average falls short of its break-even value. Whenever $R(\mu) < 1$, the producer desires that the average price paid for the lot by the consumer equals a given fraction $d_1$ of its actual price for $\mu = \mu_1 (< \mu_0)$. It is assumed that $\sigma, l$ and $r$ are known.
We have
\[
PT(\mu) = NP_1 P(X > 1) + NP_2 P(X \leq 1)
\]
\[
E(\mu) = NP_1 P(\hat{\mu} > t_0) + NP_2 P(\hat{\mu} > t_0)
\]
where
\[
P_\mu(X \leq 1) = F\left(\frac{l - \mu_0}{\sigma}\right)
\]  
(7.5)

And \(F(x) = \int_{-\infty}^{x} f(y) \, dy\) with \(f(.)\) given as in equation (7.1) with \(\mu = 0\)
and \(\sigma = 1\). Then the pay-off ratio is given by
\[
R(\mu) = E(\mu) / PT(\mu)
\]
\[
= \frac{P(\hat{\mu} > t_0) + r P(\mu \leq t_0)}{P(X > 1) + r P(X \leq 1)}
\]  
(7.6)

In view of the asymptotic equivalence of MML and ML estimators, it follows that, for large \(n\).
\[
\hat{\mu} \sim N(\mu, \frac{\sigma^2 V^*}{n})
\]
\[
\sqrt{n} (\hat{\mu} - \mu) / \sigma \sim N(0, V^*)
\]
Where \(V^* = \frac{(p-3/2)(p+1)}{p(p-1/2)}\)
\]  
(7.7)

\[
P(\hat{\mu} \leq t_0) = P\left[\sqrt{n} \frac{\hat{\mu} - \mu}{\sigma \sqrt{V^*}} < \frac{\sqrt{n}(t_0 - \mu)}{\sigma \sqrt{V^*}}\right]
\]
\[
= P[Z < \frac{\sqrt{n}(t_0 - \mu)}{\sigma \sqrt{V^*}}]
\]  
(7.8)

From (7.6), \(R(\mu_0) = 1\) implies
\[
P(\hat{\mu} \leq t_0) = P(X \leq 1)
\]
\[ P \left( Z < \frac{\sqrt{n} \left( \hat{\mu} - \mu_0 \right)}{\sigma \sqrt{v^*}} \right) = F \left( \frac{\hat{\mu} - \mu_0}{\sigma} \right) = C_2 \text{ (Known)} \]

Then we get the equation

\[ ZC_2 = \frac{\sqrt{n} \left( \hat{\mu} - \mu_0 \right)}{\sigma \sqrt{v^*}} \quad (7.9) \]

where \( ZC_2 \) is the \( C_2 \text{th} \) quantile of standard normal distribution.

Again \( R(\mu_1) = d_1 \) implies

\[
P \left( \hat{\mu} > t_0 \right) + r P \left( \hat{\mu} \leq t_0 \right) = d_1 \left[ P (X > 1) + r P (X \leq 1) \right]
\]

\[
= d_2 \left[ 1 + (r - 1) F \left( \frac{l - \mu_1}{\sigma} \right) \right]
\]

\[
= C_3
\]

Then we get

\[
P \left( \hat{\mu} > t_0 \right) = \frac{1 - C_3}{1 - r} = C_4
\]

\[
P \left( Z < \frac{\sqrt{n} \left( \hat{\mu} - \mu_1 \right)}{\sigma \sqrt{v^*}} \right) = C_4
\]

Then we have

\[ ZC_4 = \frac{\sqrt{n} \left( \hat{\mu} - \mu_1 \right)}{\sigma \sqrt{v^*}} \quad (7.10) \]

where \( ZC_4 \) is the \( C_4 \text{th} \) quantile of the standard normal distribution.

Now, from (7.9) and (7.10) we get

\[
\frac{ZC_2}{ZC_4} = \frac{t_0 - \mu_0}{t_0 - \mu_1}
\]
\[ t_0 = \frac{ZC_4 \mu_0 - ZC_2 \mu_1}{ZC_4 - ZC_2} \]  

(7.11)

From (7.9)

\[ ZC_2^2 = \frac{n(t_0 - \mu_0)^2}{\sigma^2 V^*} \]

\[ n = \frac{\sigma^2 V^*}{(t_0 - \mu_0)^2} ZC_2^2 \]

Using (7.11) we get

\[ n = \frac{\sigma^2 V^* (ZC_4 - ZC_2)^2}{(\mu_0 - \mu_1)^2} \]  

(7.12)

**Example 7.1**

Consider \( \mu_0 = 3, \mu_1 = 2.5, \sigma = 1, l = 2, d_1 = 0.7, p = 5, r = 0.5 \)

Then \( m = 2p - 3 = 7 \) and \( V^* = 0.9333 \)

From (7.1), \( f(x) = 0.44 / (1 + x^2 / 7) \)

Using Simpson's \((1 / 3)^{rd}\) rule we compute

\[ F \left( \frac{l - \mu_0}{\sigma} \right) = F (-1) = 0.1431 \]

Using (7.9), \( ZC_2 = -1.0669 \). Also,

\[ F \left( \frac{l - \mu_1}{\sigma} \right) = F (-0.5) = 0.2923 \]

Substituting the above calculation we get, \( C_3 = 0.5977 \)

Using this value \( C_4 = (1 - C_3) / (1 - r) = 0.8046 \)

Then from normal table we get \( ZC_4 = 0.8596 \)
From (7.11), $t_0 = 2.7231$ and from (7.12), $n = 14.$

Hence the optimal value of $(n, t_0)$ is $(14, 2.7231)$.

**Example 7.2**

Consider $\mu_0 = 5, \mu_1 = 4.8, \sigma = 1, l = 4.5, d_2 = 0.9, p = 2, r = 0.75$

Then $m = 2p - 3 = 1$ and $V^* = 0.5$

From (7.1), $f(x) = 0.64 / (1 + x^2 / 7)^5$

Using Simpson's $(1/3)^{rd}$ rule we compute

$$F\left(\frac{l - \mu_0}{\sigma}\right) = F(-0.5) = 0.2251$$

Using (7.9), $ZC_2 = -0.7554$. Also,

$$F\left(\frac{l - \mu_1}{\sigma}\right) = F(-0.3) = 0.3196$$

Substituting the above calculation we get, $C_3 = .8281$

Using this value $C_4 = (1 - C_3) / (1 - r) = 0.6876$

Then from normal table we get $ZC_4 = 0.4902$

From (7.11), $t_0 = 4.8787$ and from (7.12), $n = 19.$

Hence the optimal value of $(n, t_0)$ is $(19, 4.8787)$.

**APPROACH – II**

For $\mu = \mu_0$ and $\mu = \mu_2 (> \mu_0)$, the plan parameters $(n, t_0)$ are obtained from the conditions

$$R(\mu_0) = 1 \text{ and}$$

$$R(\mu_2) = d_2 (> 1) \quad (7.13)$$

where $d_2$ is generally specified by the consumer to safeguard his
interest when the process average exceeds its break-even value. The quantities $\sigma$, $l$ and $r$ are known. (The first condition in both the approaches is same.) The value of $d_2$ should be chosen in such a way that $R (\mu_2)$ satisfy (7.13). Following the same procedure as those in approach I, it is possible to determine the required values of $n$ and $t_0$.

**Example 7.3**

Consider $\mu_0 = 5$, $\mu_2 = 5.3$, $\sigma = 1$, $l = 5.2$, $d_2 = 1.1$, $p = 2$, $r = 0.75$

Then $m = 2p - 3 = 1$ and $V^* = 0.5$

From (7.1), $f(x) = 0.64 / (1 + x^2 / 7)^5$

Using Simpson’s (1/3)rd rule we compute

$F \left( \frac{l - \mu_0}{\sigma} \right) = F(0.2) = 0.6241$

Using (7.9), $ZC_2 = 0.3160$. Also ,

$F \left( \frac{l - \mu_2}{\sigma} \right) = F(-0.1) = 0.4367$

Substituting the above calculation we get, $C_3 = .9799$

Using this value $C_4 = (1 - C_3) / (1 - r) = 0.0804$

Then from normal table we get $ZC_4 = -1.4051$

From (7.11), $t_0 = 5.055$ and from (7.12), $n = 37$.

Hence the optimal value of $(n, t_0)$ is $(37, 5.055)$. 

111
Example 7.4

Consider $\mu_0 = 3$, $\mu_1 = 3.4$, $\sigma = 1$, $l = 3.3$, $d_2 = 1.1$, $p = 5$, $r = 0.5$

Then $m = 2p - 3 = 7$ and $V^* = 0.9333$

From (7.1), $f(x) = \frac{0.44}{(1 + x^2/7)^5}$

Using Simpson's (1/3)\textsuperscript{rd} rule we compute

$$F\left(\frac{l - \mu_0}{\sigma}\right) = F(0.3) = 0.6804$$

Using (7.9), $ZC_2 = 0.4677$. Also,

$$F\left(\frac{l - \mu_2}{\sigma}\right) = F(-0.1) = 0.4367$$

Substituting the above calculation we get, $C_3 = 0.8598$

Using this value $C_4 = (1 - C_3) / (1 - r) = 0.2804$

Then from normal table we get $ZC_4 = -0.5828$

From (7.11), $t_0 = 3.178$ and from (7.12), $n = 6$.

Hence the optimal value of $(n, t_0)$ is $(6, 3.178)$. 