NATIONAL CONFERENCE ON QUALITY, RELIABILITY & MANAGEMENT - 2000
EDITOR : DR. BRIJENDRA SINGH, PP. 246-254

ROBUST PROCEDURES IN STATISTICAL PROCESS CONTROL

R. P. Suresh
Indian Institute of Management Kozhikode
CEDT Campus, Calicut REC PO.,
Calicut - 673 601

ABSTRACT

Many Statistical Process Control (SPC) techniques such as Control charts, Acceptance Sampling plans and Process Capability, require that the quality characteristic of interest follow a normal distribution. However, in many applications of Acceptable Sampling Plans, the distribution of the quality characteristic is not known. Recently, Suresh and Ramanathan (1997) proposed a sampling plan when the distribution of the quality characteristic belongs to a class of distributions, which include a wide a range of symmetric distributions. In this paper, we review this method and examine whether the procedure is robust in the class of symmetric distributions.

1. INTRODUCTION

Most of the SPC techniques such as Control Charts, Process Capability and Acceptance Sampling Plans are based on the assumption that the quality characteristic of interest is normally distributed. For example, in establishing variable acceptance sampling plan, we must assume normality in order to obtain correct operating characteristic curve. In many applications of acceptance sampling plans, however, the distribution of the quality characteristic is not known as the incoming lot may be from a mixture of two or more production processes, or the production process may not be normal.

Deviations from normality can increase the risks of accepting bad-quality lots and rejecting good-quality lots (see e.g., ANSI / ASQC Z1.9 (1993) and Montgomery (1996)p.653). Moreover, if the estimator of the parameter of interest is not normally distributed, estimates of the fraction defective based on the sample mean and sample standard deviation will not be the same as if the estimator of the parameter were normally distributed. The difference between these estimated fraction defectives may be large when we are dealing with very small fraction defectives. In a highly competitive market wherein defectives are measured in terms of parts per million (p.p.m.), these differences are quite crucial.

Thus, the normal-based Acceptance Sampling Procedures can not be used when the quality characteristic has a non-normal distribution and there is a need for sampling plans, which are robust for small departures from normal distribution.

There are situations where the quality characteristic has a non-normal distribution. For example, consider the study of the amount of impurity (in ppm) in plastic to be recycled at a plant reported in Albin(1990). Here it is desired to check the amount of impurity in the incoming plastic. An item in which the impurities exceed 400 ppm is considered "defectives". It is desired to use the given data to decide whether the incoming raw material (plastic) is of acceptable quality by using Acceptance Sampling Procedure for the variable "impurity" with $p_1 = .01$, $\alpha = .05$, $p_2 = .07$, $\beta = .06$. Thus, we need to apply Acceptance Sampling with USL = 400 ppm.

A frequency table of the data (ignoring the time sequence) is given in Table 1.1.

TABLE 1.1

Range	*Frequency*
0 - 50	**
50 - 100	******
100 - 150	**********
150 - 200	****
200 - 250	**
250 - 300	*
300 - 350	
350 - 400	
400 - 450	
450 - 500	
500 - 550	*

First, we apply Shapiro-Wilk test to test for normality of the data. The value of the Shapiro-Wilk statistic is $W=0.7997$ with a p-value of 0.00009, hence normality is rejected and hence the normal sampling plans cannot be applied for this data.

In section 2 of this paper, we review the method of Suresh and Ramanathan (1997), and apply this method to the above data. In Section 3, we study robustness of the method by Suresh and Ramanathan(1997) in the class of symmetric distributions considered by them, using a simulation study.

2. SURESH-RAMANATHAN METHOD

This paper is based on the assumption that the distribution of the quality characteristic is symmetric and is a member of the family of symmetric distributions with probability density function (pdf) given by

$$f(x / \mu, \sigma, m) = [\sigma m^{1/2} \beta(1 / 2, p - 1 / 2)]^{-1} \quad [1 + (x - \mu)^2 / (m\sigma^2)]^{-p} \qquad - \infty < x < \infty \qquad (2 - 1)$$

where β (.,.) denotes the beta function, $m = 2p\text{-}3$, m and s are the location and scale parameters respectively. Note that this family (with minor modification of parameters) includes a wide range of symmetric distributions including Student's t, Cauchy and Normal distributions. It may also be noted that the distribution given in (2.1) with p = 5 is almost indistinguishable from a logistic distribution. The authors derive sampling plans for single and double specification limits using the Modified Maximum Likelihood (MML) estimates of the location and scale parameters for the above family. (see Tiku and Suresh (1992)).

2.1 Estimation of μ and σ

Let $X1, X2, \ldots, Xn$ be a random sample from a distribution with pdf f given in (2.1) with known m, and let $X(1), X(2), \ldots, X(n)$ be the corresponding order statistics.

Following Tiku and Suresh (1992), the MML estimators of μ and σ are given by

$$\mu^* = \sum_{1=1}^{n} \beta_I x_{(i)} / \sum_{1}^{n} \beta_i$$

and

$$\sigma^* = \{B + \sqrt{B^2 + 4nC}\} / 2\sqrt{n(n-1)}$$

where

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

$$B = \frac{2p}{m} \sum_{i=1}^{n} \alpha_i x_{(i)}$$

$$C = (2p / m) \sum_{i=1}^{n} \beta_i \{x_{(i)} - \mu^*\}^2$$

$$\alpha_i = (2 / m) t^3_{(i)} / [1 + t^2_{(i)} / m]^2$$

and

$$t(i) = E[X_{(i)} - \mu) / \sigma]$$

The values of $t_{(i)}$ can be obtained using the approximation of moments of order statistics obtained by David and Johnson (1954), see also David (1981).

The MML estimators μ^* and σ^* given above are asymptotically equivalent to the maximum likelihood (ML) estimators and thus are asymptotically unbiased and efficient. Note, however, that μ is unbiased for all n. This follows from the symmetry of (2.1). In view of the asymptotic equivalence of the MML and ML estimators, it follows that, for large n,

$$\sqrt{n}(\theta^*) \sim N_2(0, V), \tag{2.2}$$

where

$\theta^* = (\mu^* - \mu, \sigma^* - \sigma)$, V is a diagonal matrix with diagonals V_1 and V_2, with

$$V_1 = \sigma^2(V_1^*) \text{ and } V_2 = \sigma^2 (V_2^*), \text{ where}$$

$$V_1^* = (p - 3/2)(p+1)/[p(p - 1/2)]$$

and $$V_2^* = (p + 1)/[2p - 1]$$

From relation (2.2), using invariance property of consistent and asymptotic normal estimators, it follows that

$$\sqrt{n}\{(\mu^* - k\sigma^*) - (\mu - k\sigma)\} \sim N(0, V_1 + k^2V_2)$$

and

$$\sqrt{n}\{(\mu^*+k\sigma^*)-(\mu+k\sigma)\} \sim N(0, V_1 + k^2V_2) \tag{2.3}$$

for any k and large n.

2.2 Sampling Plans

2.2.1 Single Specification limit

We adopt a producer's and consumer's risk approach to design the sampling plan. For a single lower specification limit (LSL), an appropriate plan based on μ^* and σ^* is to accept the lot if

$$\mu^* - k\sigma^* \geq LSL \tag{2.4}$$

for a proper choice of n and k. The producer's and consumer's risk gives rise to the following:

$$Pr\{\text{Accepting the lot / proportion defective} = p_1\} = 1 - \alpha$$

and

$$Pr\{\text{Accepting the lot / proportion defective} = p_2\} = \beta$$

That is,

$$Pr\{\mu^* - k\sigma^* \geq LSL / p_1\} = 1 - \alpha$$

and

$$Pr\{\mu^* - k\sigma^* \geq LSL / p_2\} = \beta.$$

Using the asymptotic distribution of $\mu^*\text{-}k\sigma^*$, given in equation (2.3), these become

$$\Pr\left(\sigma_1\sqrt{V_1^* + k^2V_2^*}\ Z \geq \sqrt{n}\,[LSL - (\mu_1 - k\sigma_1)]/p_1\right) = 1-\alpha \tag{2.5}$$

and

$$Pr\left(\sigma_2\sqrt{V_1^* + k^2V_2^*}\ Z \geq \sqrt{n}\,[LSL - (\mu_2 - k\sigma_2)]/p_2\right) = \beta \tag{2.6}$$

where Z is a standard normal random variable,

$$p_1 = Pr[X \leq LSL] = F((LSL - \mu_1)/\sigma_1) \tag{2.7}$$

$$p_2 = Pr[X \leq LSL] = F(LSL - \mu_2)/\sigma_2) \tag{2.8}$$

and $F(x)$ is the cumulative distribution function of the random variable pdf $f(.)$ given in equation (2.1) with $\mu = 0$ and $\sigma = 1$.

Using equations (2.7) and (2.8) we get

$$[LSL - \mu_1]/\sigma_1 = F^{-1}(p_1) = t_{p1}$$

and

$$[LSL - \mu_2]/\sigma_2 = F^{-1}(p_2) = t_{p2}$$

where $F(t_{p1}) = p_1$. Substituting this in (2.5) and (2.6), one obtains

$$\Phi\left(\sqrt{n}(t_{p1} + k)/\sqrt{V_1^* + k^2V_2^*}\right) = 1 - \alpha \tag{2.9}$$

and

$$\Phi\left(\sqrt{n}(t_{p2} + k)/\sqrt{V_1^* + k^2V_2^*}\right) = \beta \tag{2.10}$$

where F is the cumulative distribution function of standard normal random variable. Let Z1-α be such that $\Phi(Z_1 - \alpha) = 1 - \alpha$. Now, equations (2.9) and (2.10) can be written as

$$\sqrt{n}(t_{p1} + k) = Z_{1\ \alpha}\sqrt{V_1^* + k^2V_2^*}$$

and

$$\sqrt{n}(t_{p2} + k) = Z_\beta\sqrt{V_1^* + k^2V_2^*}$$

Solving these two equations, we get

$$k = (t_{p2}\, Z_{1-\alpha} - t_{p1}Z_\beta)/(Z_\beta - Z_{1-\alpha}) \tag{2.11}$$

and

$$n = Z^2_{1-\alpha}\,(V_1^* + k^2V_2^*)/(t_{p1} + k)^2 \tag{2.12}$$

Thus, the proposed sampling plan with a specified LSL is to accept the lot if $\mu^*\text{-}k\sigma^* \geq LSL$, reject otherwise with n and k given by equations (2.11) and (2.12).

Remark 2.1 : We would like to bring to the notice that the expressions given in equation (11) and (12) of Suresh and Ramanathan were wrong, as they involved σ in V_1 and V_2. The correct expressions are indicated in (2.11) and (2.12) respectively.

Proceeding in the same way, one can derive the acceptance sampling plan with upper specification limit (USL) and is given by

Accept the lot if $\mu^*+k\sigma^* \leq USL$, otherwise reject with the same n and κ as above.

Note that the fraction defectives with the given LSL (USL) is the area under the frequency curve given in equation (2.1) to the left of LSL (right of USL). Hence, the fraction defectives is given by

$$PL = Pr\{X \leq LSL\}$$
$$= F((LSL - \mu) / \sigma) \quad \text{and}$$
$$PU = 1 - F((USL - \mu) / \sigma) = F((\mu - USL) / \sigma)$$

The sampling plans derived above can also be described using these fraction defectives as follows:

Accept the lot if PL^* $(PU^*) \leq p^*$, reject otherwise, where

$$PL^* = F((LSL - \mu^*) / \sigma^*)$$
$$PU^* = F((\mu^* - USL) / \sigma^*)$$

and

$$p^* = 1 - F(k)$$

with k and n as in equations (2.11) and (2.12) respectively.

2.2.2 Double Specification

Following Wetherill and Brown (1991,p. 289), we propose the sampling plan given below, when both USL and LSL are given.

Accept the lot if $PL^* + PU^* < p^*$, reject otherwise, where PL^*, PU^* and p^* and n are as given in the case of single specification sampling plan.

In Table 2.1, simulated values of a using proposed sampling plan and normal sampling plan (in brackets) are given for different distributions. The simulation size is taken to be 10000 with nominal value of $\alpha = .01$ and keeping $p_1 = .01$, $p_2 = .07$ fixed through out. The distributions considered for this purpose are Logistic, Student's t and Normal. The acceptance sampling plans for the Logistic distribution is derived by taking $p = 5$.

TABLE 2.1: SIMULATED VALUES OF α

	β = .05	.07	.1
Logistic	.010 (.002)	.011 (.002)	.013 (.003)
t_5	.013 (.005)	.013 (.003)	.012 (.005)
t_{10}	.009 (.001)	.009 (.001)	.010 (.002)
t_{15}	.009 (.002)	.010 (.002)	.009 (.002)
t_{20}	.009 (.002)	.010 (.003)	.010 (.002)
Normal	.025 (.010)	.002 (.013)	.021 (.011)

Notice that except when the characteristic has a normal distribution, the proposed sampling plan performs remarkably well as compared to the normal sampling plan for various choices of β.

The authors recommend that these sampling plans be used with an appropriate choice of *m*. If the distribution of the quality characteristic is identified and if it belongs to a symmetric family, a proper choice of *m* can be made. Otherwise, *m* may be taken to be in the range 10 to 15.

2.3 Albin's Data

Now, we apply the above method to the data described in Section 1.

Since the data is thick-tailed, we take m = 5 in the density (2.1) and obtain estimates as described in section 2.1 and section 2.2, and are given below:

$\mu^* = 131.665, \ \sigma^* = 68.4, \ PU^* = .005, \ p^* = .026.$

Since $PU^* < p^*$, we accept the lot and conclude that the impurities are within the allowable limits.

3. ROBUSTNESS OF THE SAMPLING PLAN

The assumption that the parameter m is known is very crucial for implementation of this method. Thus, it is necessary to study the robustness of the method for departures from the assumed value of the parameter m. Here, we study the robustness of the procedure by using a simulation study of 10000 simulations. The following points on the OC curve are considered, viz. $\alpha = .01, \ \beta = .01, \ p_1 = .01, \ p_2 = .07$. The tables 3.1-3.4 give the simulated values of α.

SIMULATED VALUES OF α

TABLE 3.1 (TRUE VALUE OF M = 5)

Assumed Value of m	Simulated value of α
m = 3	0.0018
m = 4	0.0077
m = 5	0.0009
m = 6	0.0114
m = 7	0.0119
m = 8	0.0138
m = 9	0.0085
m = 10	0.0066

TABLE 3.2 (TRUE VALUE OF M = 10)

Assumed Value of m	Simulated value of α
m = 5	0.0422
m = 6	0.0349
m = 7	0.0245
m = 8	0.0179
m = 9	0.0136
m = 10	0.0114
m = 11	0.0098
m = 12	0.0099
m = 13	0.0086
m = 14	0.0046
m = 15	0.0056

TABLE 3.3 (TRUE VALUE OF M = 15)

Assumed Value of m	Simulated value of α
m = 10	0.0195
m = 11	0.0167
m = 12	0.0139
m = 13	0.0135
m = 14	0.0121
m = 15	0.0106
m = 16	0.0110
m = 17	0.0098
m = 18	0.009
m = 19	0.0088
m = 20	0.0069

TABLE 3.4 (TRUE VALUE OF M = 20)

Assumed Value of m	Simulated value of α
m = 15	0.0172
m = 16	0.0145
m = 17	0.0131
m = 18	0.0131
m = 19	0.0121
m = 20	0.0119
m = 21	0.0125
m = 22	0.0098
m = 23	0.0101
m = 24	0.0101
m = 25	0.0084

From the above tables, it is clear that the acceptance sampling plans proposed by Suresh-Ramanthan are, in fact, robust to small departures from the assumed parametric value viz. m. For example, if on the basis of the data, one assumes a value of m = 6, and if the true value is m = 5, the producer's risk will be equal to .0114, which is only slightly above the desired value of $\alpha = .01$, with which the sampling plan is designed.

Thus, the sampling procedure can be used for non-normal symmetric distributions.

REFERENCES

1. **Albin S.C.** (1990), "The Lognormal distribution for Modeling Quality Data when the mean is near Zero", *Journal of Quality Technology*, **22**, 105-110.
2. ANSI / ASQC Z 1.9 (1993), "Sampling Procedures and Tables for Inspection by Variables by Per cent Non- confirming". *American Society for Quality Control*, Milwaukee, WI.
3. **David, F. N.** and **Johnson** (1954). Statistical treatment of censored data I, Fundamental formulae. *Biometrika*, **41**, 228-240.
4. **David, H. A.** (1981). *Order Statistics*, John Wiley, New York.
5. **Montgomery, D.C.** (1996) *Introduction to Statistical Quality Control*, John Wiley, New York.
6. **Suresh, R.P.** and **Ramanathan, T.V.** (1997) "Acceptance Sampling Plans by Variables for a class of Symmetric Distributions", *Communication in Statistics - Simulation and Computation*, **27**.
7. **Tiku, M.L.** and **Suresh, R.P.** (1992). A New Method of Estimation for Location and Scale Parameters. *Journal of Statistical Planning and Inference*, **30**, 281-292.
8. **Wetherill, G.R.** and **Brown, D.W.** (1991) *Statistical Process Control - Theory and Practice*. Chapman and Hall, London.