ISSN 2229-5518

A new method of implementing Single Sampling Plan by variables using interval estimates of process mean and variance

Reddi Rani K [1], Sai Sarada V [2], Sarma KVS [3]
[1]Department of Statistics, SV University, Tirupati. 517 502, India
[1]reddiranikolle@yahoo.in, [2]saisaradav@gmail.com, [3]smskvs@gmail.com

Abstract: In this paper we propose a new method of implementing a Single Sampling Plan by Variables (SSPV) when a lower specification limit is given. The idea is to utilize the information contained in the (1-α)% confidence intervals (CI) of a) the process mean and b) process variance in the place of their point estimates and a new test statistic Z is proposed assuming that the data follows normal distribution. Taking the lower, middle and upper values of the CI for each of the two parameters, we get 9 possible combinations to derive the Z statistic. We have proposed a method of weighted average of these 9 estimates and examined the quality of decision about sentencing the lot. It is shown by simulation that the new method gives better results than the method based on the point estimates.

Key words: Confidence Intervals, Specification Limit, Lot Acceptance Probability

1. Introduction

Single sampling plan for variable type of inspection (SSPV) is a classical concept in the field of acceptance sampling. The SSPV is specified as (*n*, *k*) where *n* denotes the sample size to be drawn from the lot and the lot is accepted or rejected basing on the outcome of the sample inspection. Let X be the quality characteristic following N(μ, σ^2) and we assume that the production process is in a steady state capable of meeting the speciation limits.

Define L as the lower specification limit such that any item for which the observed measurement x ≥ L is accepted.

Let $\bar{x} = \sum_{i=1}^{n} x_i$ and $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{(n-1)}$ be the unbiased point estimates of μ and σ^2 respectively based on a sample data of *n* observations. The classical method of implementing the sampling plan is to compute a statistic

$$Z_L = \frac{(\bar{x} - L)}{s} \tag{1}$$

and the lot is accepted if $Z_L \geq k$ and rejected otherwise. The Operating Characteristic (OC) of the plan is the proportion of lots accepted at a given level of lot quality, expressed as μ.

The basics of this plan can be found in Montgomery [1], Duncan [2] and Grant and Leavenworth [3]. In general we will not be knowing the true mean or variance of the process but proceed with prior knowledge or from process history by computing the sample mean and variance as point estimates. The value of Z_L is sensitive to the values of $\bar{x}$ and s and as such the lot sentencing depends on how good these values are known.

In this paper we propose a new method of estimating the lot quality taking into account the interval estimates of μ and σ^2 and generate a new expression in the place of Z_L.

2. The confidence intervals of μ and σ^2

Let f(x,θ) be the density function of X and θ be estimated from a sample {x_1, x_2, …,x_n} drawn from f(x,θ). Then the 100(1-α)% interval estimate of θ will be [$\hat{\theta}_L, \hat{\theta}_U$] where $\hat{\theta}_L$ and $\hat{\theta}_U$ denote the lower and upper bounds of the CI. (see Rohatgi [4]). We denote these bounds by θ_1 and θ_3, which can be considered as *extreme estimates* of θ and define θ_2 as the value at the middle of the CI. Then from the confidence interval one gets three possible estimates of θ viz., θ_1, θ_2 and θ_3, out of which θ_2 is likely to be closer to θ than the other two. In the light of this approach we get three estimates for μ and three for σ so that there will be 9 possible combinations at which (1) can be evaluated. We call this the *triple estimate method.* Further, each combination is a candidate for lot sentencing. We wish to combine them into a single value and examine its properties.

Since $\bar{x}$ follows N(μ,σ^2/*n*) the (1-α)% *Confidence Interval* (CI) for μ is given by

$$\{\bar{x} - t_{(n-1),\alpha/2}\frac{s}{\sqrt{n}}\ ,\ \bar{x} + t_{(n-1),\alpha/2}\frac{s}{\sqrt{n}}\ \} \tag{2}$$

where $t_{(n-1),\alpha/2}$ denote the upper α/2[th] percentile of the Student's t-distribution with (*n*-1) degree of freedom.

The value of μ is unknown but on an average it is equal to $\bar{x}$. However, the true value of μ lies in the interval given in (1) with probability (1-α).

Similarly let $\chi^2_{\alpha/2}$ and $\chi^2_{1-\alpha/2}$denote the lower and upper α/2 percentiles on the Chi-square distribution. These values can be found by using simple Excel functions. Then for the process σ^2, the 100(1-α)% CI, based on the Chi Square distribution is given by

$$\{\ \frac{(n-1)}{\chi^2_{\alpha/2}}s^2\ ,\frac{(n-1)}{\chi^2_{1-\alpha/2}}s^2\ \}. \tag{3}$$

[1] *Sarma KVS is a professor (retired) from Sri Venkateswra University, Tirupati, India.* [2,3] *Reddy Rani and Sai Sarada are research scholars working in the area of Statistical Quality Control methods, in Sri Venkateswara University, Tirupati, India.*

Again the true value of σ^2 lies in (3) with probability (1-α). The width of the CIs in both (2) and (3) depends on the values of $\bar{x}$ and s. The following are the lower, middle and upper values of the estimates for μ.

$$m_1 = \bar{x} - t_{(n-1),\alpha/2}\frac{s}{\sqrt{n}},\ m_2 = \bar{x},\ m_3 = \bar{x} + t_{(n-1),\alpha/2}\frac{s}{\sqrt{n}} \qquad (4)$$

and the estimates for σ are

$$s_1 = \sqrt{\frac{(n-1)}{\chi^2_{\alpha/2}}s^2},\ s_2 = \sqrt{s^2},\ s_3 = \sqrt{\frac{(n-1)}{\chi^2_{1-\alpha/2}}s^2}. \qquad (5)$$

This method of summarizing confidence intervals to generate the *criterion* was earlier used by Vishnu Vardhan *et al* [6],[7] while trying to estimate the area under the ROC curve. Sai Sarada *et al* [5] have used this method and obtained a new expression for estimating the process capability index of a process following normal distribution. In the following section, we derive a pooled estimate of Z_L by utilizing the information from the 9-combinations of estimates of μ and σ.

3. New method of implementing the SSPV

Consider a characteristic X for which the lower specification limit is given as L. For any item, if the observation x on X satisfies the condition $x \leq L$, the item is considered as defective. Let α denotes the producer's risk and define $k = Z_\alpha = \phi^{-1}(\alpha)$ be the value of standard normal distribution at α where ϕ denotes the cumulative standard normal distribution. The general procedure for decision making (Duncan[2]) is as follows.

1. Find $Z_L = \frac{(\bar{x} - L)}{s}$
2. Calculate $k = \phi^{-1}(\alpha)$
3. Accept the lot if $Z_L \geq k$; else reject.

Consider the following propositions.

Propositon-1: Define $Z_{ij} = \frac{(\bar{x}_i - L)}{s_j}$ as the standard normal deviate defined at the combination ($\bar{x}_i$, s_j) for i, j = 1,2,3. Then the new test statistic is given as

$$Z_{pooled} = \sum_{i=1}^{3}\sum_{j=1}^{3} w_{ij} Z_{ij} \qquad (6)$$

where w_{ij} is the weight given to Z_{ij}. The decision rule is to accept the lot if $Z_{pooled} > k$; else reject. ■

Propositon-2: We can alternatively define in the one dimensional space, {Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8, Z_9} as the vector of the 9 estimates and take its scalar product with the weight vector {W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9}. When the sum of weights is unity, then Z_{pooled} will be a convex combination of the 9 individual values. Further, it can be shown that $Z_{pooled} \leq \text{Max}\{Z_{ij}\}$.

Propositon-3: One way of allotting the weights is to allot 0.5 to Z_5 where both the point estimates are used and the remaining 8 weights can be taken as 0.5/8 each. The vector of weights in this case will be

W = {0.0625, 0.0625, 0.0625, 0.0625, 0.5, 0.0625, 0.0625, 0.0625, 0.0625}. (7)

Propositon-4: Define Z_0 = target value of Z obtainable when the process operates at μ and σ. Then an objective way of allotting the weights is to take

$$W_j = (Z_j - Z_0)^{-2} \quad \text{for } j = 1 \text{ to } 9. \qquad (8)$$

This means that the weights are proportional to the inverse of the *squared error* in Z_j from the target Z_0. These weights are in a way *adaptive* and depend on the sample data.

In the following section, we conduct simulated trials and study the properties of the new decision rule based on Z_{pooled}. We also find the operating characteristic (OC) by empirically finding $P(Z_{pooled} > k)$.

Let us consider a hypothetical sampling plan with n = 20, k = 2.17009 with producer's risk as α = 0.015 assuming that the variance is unknown.

Algorithm

The following steps can be followed to implement the new method.

1. Generate *m* random samples each of size *n* = 20 from $N(\mu, \sigma^2)$.
2. Evaluate the sample $\bar{x}$ and s for each sample.
3. Find the three estimates of μ as m_1, m_2 and m_3 and three estimates of σ as s_1, s_2 and s_3 using (4) and (5)
4. Find Z_{pooled} using (6) and (8).
5. For each sample (lot), mark 'A' if $Z_{pooled} >$ k, to indicate acceptance of the lot and mark R (rejected) otherwise and find the proportion of A's out of *m*.

If *m* lots are tested with this plan, the proportion of lots accepted is an estimate of the lot acceptance probability $\hat{P}(A) = r/m$ where *r* denotes the number of lots accepted by the plan and *m* is the total number of lots inspected. Since each lot has a binary outcome, the standard error of this estimate is $\sqrt{\hat{P}(A)(1-\hat{P}(A))/m}$. This $\hat{P}(A)$ produces different values at the 9-estimates of Z and also at Z_{pooled}.

In the following section a simulated experiment is reported along with a template in Excel to demonstrate the working of the new method.

4. Illustration

Consider a situation where X follows normal with μ =10.8 and σ =1.5. Let the lower specification limit (L) is given as 7.0. We have generated 30 samples each of size 20 from the above normal distribution using the Random Number Generation tool of the Data Analysis Pak in Excel. The confidence intervals are generated by taking α = 0.05. The Excel function CONFIDENCE(0.05,1.5,20) has been used to derive the lower and upper limits.

The three estimates of means and standard deviations each based on the 95% confidence intervals are worked out using Excel template. Assuming that the process *truly* operates at μ = 10.8 and σ = 1.5, the Z statistics in relation to L will be the target Z (denoted by Z_0) obtainable when samples are randomly drawn from this process. In this case Z_0 becomes 2.5333.

Since we have fixed the producer's risk at α = 0.015 we get k = 2.17009. Table-1 shows the adaptive weights for each of the 9 values of Z and Z_{pooled} is computed. The rule is to accept the lot if $Z_{pooled} \geq$ k. For each lot we have computed the decision to accept or reject (marked as A or R in the sheet with a condition “ =IF(AN11>AP1,"A","R")”.

We have also computed the decision outcome using Z_{22} as done conventionally and compared the decision from these two methods. The proportion of accepted lots is found and its standard error is also reported.

In the Figure-1, the heading *des1* refers to Probability of accepting the lot basing on Z_{pooled} and *des2* refers to Probability of acceptance basing on Z_{22} (which is the same as Z_5).

We have repeated the above experiment with 500 samples each and the results are summarized in Table-1.

Figure-1: Comparative performance of Z_{pooled} and Z_{point} methods

BC13

	A	B	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY
1	L	7.00							K	2.17009										
2		LOT NO.	z21	z22	z23	z31	z32	z33	w1	w2	w3	w4	w5	w6	w7	w8	w9	Zpooled	des1	des2
3		1	3.4239	2.6038	1.7827	3.9096	2.9732	2.0357	6.10256	11.1924	0.99303	1.2609	201.198	1.77501	0.52792	5.16744	4.03765	2.5897	A	A
4	Target	2	3.1950	2.4298	1.6636	3.7397	2.8440	1.9472	73.0842	3.72971	0.75173	2.28394	93.2688	1.32195	0.68708	10.3584	2.91088	2.5269	A	A
5	2.5333333	3	2.8959	2.2023	1.5078	3.3901	2.5782	1.7652	57.6665	2.00115	0.60765	7.60738	9.12554	0.95089	1.36217	497.4	1.69474	2.5526	A	A
6		4	2.8035	2.1320	1.4597	3.3174	2.5229	1.7273	16.817	1.5934	0.55588	13.7046	6.20867	0.86755	1.62649	9150.11	1.53928	2.5223	A	R
7		5	3.4659	2.6357	1.8046	4.0094	3.0491	2.0876	6.61088	10.3403	0.97687	1.14997	95.3434	1.88305	0.45895	3.7586	5.03405	2.5987	A	A
8		6	3.4142	2.5965	1.7777	3.9762	3.0238	2.0703	9.82778	7.54029	0.91017	1.28865	250.616	1.75151	0.48036	4.15632	4.66445	2.5890	A	A
9		7	3.5086	2.6683	1.8269	4.1386	3.1474	2.1549	8.39024	8.44016	0.93437	1.05137	54.9323	2.00357	0.38805	2.65206	6.98278	2.5993	A	A
10		8	2.5191	1.9158	1.3117	3.0383	2.3106	1.5820	3.5153	0.97571	0.44923	4954.55	2.62204	0.67002	3.92205	20.1537	1.10486	2.5173	A	R
11	alpha	9	2.9948	2.2775	1.5593	3.5002	2.6619	1.8225	517.868	2.44006	0.65336	4.69597	15.2804	1.0541	1.0697	60.5253	1.97901	2.4995	A	A
12	(Prod. Risk)	10	2.8194	2.1441	1.4680	3.3369	2.5377	1.7374	18.6675	1.63203	0.56127	12.2215	6.60094	0.8811	1.54875	53354.6	1.57869	2.5375	A	R
13	0.015	11	3.0616	2.3283	1.5941	3.5900	2.7301	1.8692	6.9E+07	2.71543	0.67814	3.5834	23.7915	1.13361	0.89565	25.8155	2.26742	2.5332	A	A
14		12	3.2664	2.4841	1.7008	3.8387	2.9193	1.9988	38.6684	4.26072	0.78239	1.86062	412.575	1.44268	0.58682	6.71171	3.49931	2.5006	A	A
15		13	3.5541	2.7028	1.8505	4.0629	3.0898	2.1155	3.81687	21.1394	1.11326	0.95981	34.8103	2.14488	0.42742	3.22925	5.72737	2.5575	A	A
16		14	3.3939	2.5811	1.7672	3.9508	3.0045	2.0571	10.8399	7.08183	0.89653	1.3502	439.12	1.70349	0.4977	4.50359	4.4093	2.5785	A	A
17		15	2.9168	2.2182	1.5187	3.4588	2.6304	1.8009	39.8206	1.89063	0.59465	6.79992	10.0704	0.97142	1.16763	106.215	1.86415	2.5334	A	A
18		16	3.5178	2.6753	1.8317	4.1230	3.1355	2.1468	6.95016	9.8708	0.96728	1.03175	49.6352	2.03111	0.39572	2.75779	6.69187	2.5937	A	A
19		17	4.4160	3.3583	2.2993	5.2414	3.9861	2.7291	0.89468	25.7016	2.26946	0.28214	1.46929	18.2605	0.13636	0.47384	26.0916	2.6411	A	A
20		18	3.6174	2.7510	1.8835	4.3737	3.3262	2.2773	9.30724	7.82522	0.91818	0.85089	21.1038	2.36821	0.29524	1.59084	15.256	2.5530	A	A
21		19	3.1403	2.3882	1.6351	3.7176	2.8272	1.9357	1127.4	2.93098	0.69587	2.71394	47.4788	1.23948	0.71304	11.5808	2.79955	2.5561	A	A
22		20	3.3255	2.5290	1.7315	3.8964	2.9632	2.0288	20.4318	5.20037	0.82782	1.59364	53104.9	1.5554	0.53823	5.41249	3.92802	2.5290	A	A
23		21	2.0129	1.5308	1.0481	2.3919	1.8190	1.2454	1.23618	0.60022	0.35322	3.69187	0.99491	0.45331	49.9578	1.95963	0.60285	2.2774	A	R
24		22	3.6839	2.8016	1.9181	4.3563	3.3129	2.2682	4.37332	16.9206	1.07319	0.75539	13.8981	2.64222	0.30092	1.64541	14.2291	2.4970	A	A
25		23	4.2254	3.2134	2.2001	4.8804	3.7115	2.5411	0.92992	30.2334	2.19912	0.34929	2.16251	9.00349	0.18153	0.72042	16450.7	2.5414	A	A
26		24	3.4786	2.6455	1.8112	4.0090	3.0488	2.0874	5.80877	11.7912	1.00353	1.11911	79.5452	1.91788	0.45923	3.76359	5.02876	2.5962	A	A
27		25	3.8806	2.9511	2.0205	4.5716	3.4766	2.3803	2.32232	86.1795	1.3133	0.55096	5.72892	3.80276	0.24071	1.1238	42.7177	2.4451	A	A
28		26	3.2287	2.4554	1.6811	3.7881	2.8808	1.9724	54.1862	3.94543	0.76471	2.06832	164.511	1.37684	0.6351	8.28043	3.1783	2.5075	A	A
29		27	3.8230	2.9074	1.9906	4.4330	3.3712	2.3082	2.16443	123.898	1.35094	0.60122	7.14787	3.39451	0.27711	1.42432	19.7237	2.4587	A	A
30		28	2.4619	1.8722	1.2819	2.9284	2.2270	1.5247	3.45576	0.96904	0.4478	195.901	2.28811	0.63849	6.40794	10.6564	0.98304	2.4362	A	R
31		29	3.3404	2.5404	1.7393	3.9373	2.9943	2.0501	22.6183	5.0081	0.81929	1.53509	20158.6	1.58608	0.50733	4.70668	4.28183	2.5405	A	A
32		30	3.7216	2.8303	1.9378	4.3418	3.3019	2.2607	3.09743	32.7843	1.18552	0.70818	11.3413	2.81946	0.30577	1.69309	13.4502	2.4690	A	A
33																		P(A)	1.00	0.83
34																		SE		

From the above experiments we observe the following.

1. The new estimate Z_{pooled} provides Z values closer to the target when compared with the conventional point estimate method given by Z_{22}.
2. At different values of the process mean, Z_{pooled} is less than Z_{22}

Table-1: Performance of Z_{pooled} method in comparison with the classical Z method from 500 simulations (n = 20, k = 2.170)

Criterion value	Mean = 10.2, SD = 1.5 Z_0 = 2.1333			Mean = 10.5, SD = 1.5 Z_0 = 2.3333			Mean = 10.8, SD = 1.5 Z_0 = 2.5333		
	Mean	$\hat{P}$ (A)	SE	Mean	$\hat{P}$ (A)	SE	Mean	$\hat{P}$ (A)	SE
Z_{pooled}	2.1269	0.1960	0.0132	2.3306	0.9920	0.0039	2.5218	0.9980	0.0020
Z_{22}	2.2298	0.4800	0.0167	2.5095	0.7380	0.0196	2.6393	0.8260	0.0169

4. Analysis of simulations

We have performed a detailed analysis of the output from 500 simulations. Each trial represents a lot from which 20 samples were drawn at random from N(10.5, 1.5) with lower specification limit L = 7.00. The target value which we expect to get out of random samples from the process will be Z_0 = 2.3333.

A run chart was also constructed for the 500 individual values of Z_{22} and Z_{pooled} which shows that Z_{pooled} is more consistent than Z_{22}. The mean of Z_{22} is 2.5094 with 3 sigma control limits as (1.0607, 3.9583) and 4 samples (lots) out of 500 have shown out of control.

In the case of Z_{pooled} the mean is 2.3306 with 3 sigma control limits as (2.1593, 2.5019) and only one out of 500 samples fell out of control. We find that 3 samples (lots) out of 500 have shown out of control. Similar results are found when the process mean has shifted to left (10.2) and to the right (10.8).

Kolmogorov- Smirnov test for normality of Z_{22} and Z_{pooled} has shown that Z_{pooled} is closer to normality (p = 0.546) than Z_{22} (p = 0.078) confirming that the new statistic has a good normal distribution.

Figure-2 gives the Box plot of the distribution of Z_{22} and Z_{pooled} which shows that the distribution of Z_{pooled} has lower dispersion when compared to that of Z_{22} though both statistics have few extreme values. We therefore conclude that the new estimated Z_{pooled} offers a better estimate of the test criterion instead of Z_{pooled}.

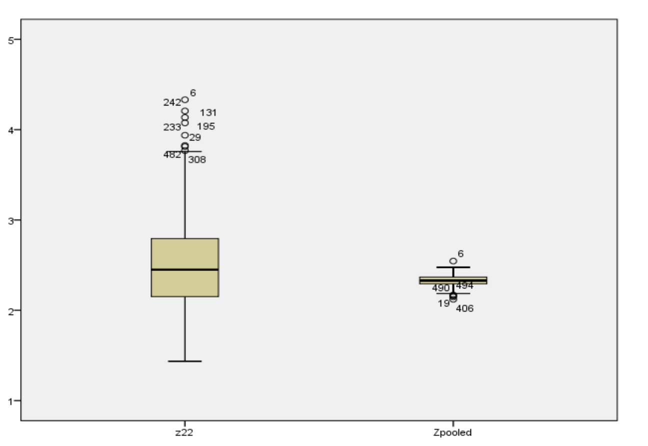
Conclusions

It is observed that the new criterion (Z_{pooled}) which is the weighted of Z values obtained from the 9 possible combinations of the estimates of mean and variance, has lower standard error than single Z value based on the point estimates. Hence we recommend the use of Confidence Interval based decision rule in place of the conventional method based on point estimates.

Acknowledgements

The authors wish to thank Prof. Rao. S. Jammalamadaka, University of California, for his valuable suggestions and directions on a draft version of the method proposed in this paper.

Figure-2: Box plot for Z_{22} and Z_{pooled}



References

[1] Douglas C. Montgomery (1985), Introduction to Statistical Quality control By John Wily & sons, Inc.

[2] Duncan A.J (1969), Quality Control and Industrial Statistics, Tarapoorvala Sons & Co. Pvt., Bombay.

[3] Grant E.L and R.S. Leavenworth (2000), Statistical Quality Control, Tata McGraw- Hill

[4] Rohatgi V.K (1976), An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Limited.

[5] Sai Sarada V, M. Subbarayudu and K.V.S.Sarma (2017), Estimation of Process Capability Index using confidence intervals of process parameters, Working paper-I, submitted for publication.

[6] Suresh Babu N, R.V. Vardhan and Sarma K.V.S. (2011), On the Comparison of Binormal ROC Curves using Confidence Interval of Means, *International Journal of Statistics and Analysis*, Volume 1(3), pp. 213-221.

[7] Vishnu Vardhan R and K.V.S. Sarma (2010), Estimation of the Area Under ROC Curve using Confidence Interval of Mean, *J. Phy. Sci.* 2(1), 29-39