ON THE MOST ECONOMICAL SAMPLE SIZE FOR CONTROLLING THE MEAN OF A POPULATION

By H. Weiler

New South Wales University of Technology, Sydney, Australia

Summary. For quality control charts controlling the mean of a population either small samples may be taken out at frequent intervals or larger samples at less frequent intervals. In this paper, a simple formula is derived by which the most suitable sample size can be determined, leading to the detection of any given change of the population mean with a minimum of inspection.

1. Introduction. Consider a normal variate x representing some measure of a mass-produced article, and suppose that control limits similar to those used in the preparation of control charts [1], [2] are to be determined to detect changes of the population mean of x. Such control limits may be placed on a chart similar to that used in control chart analysis.

After the mean and standard deviation of a population have been estimated by means of an initial large sample, smaller samples of fixed size N are taken during the production, and their arithmetic means $\bar{x} = \sum x/N$ are calculated. A chart is then constructed with control limits $m \pm 3\sigma/\sqrt{N}$, where m and σ are the estimates of the population mean and S.D. obtained from the original large sample. The various values of $\bar{x}$ are then entered in the chart in chronological order, and as soon as one such value falls outside the control limits, production is stopped to allow investigation.

The aim of this paper is to determine the most economical sample size, that is, that value of N which would indicate a change of the population mean after a minimum amount of inspection. It will be found that the most economical sample size depends on the amount by which the population mean has changed. Thus, if the population mean changes from m to $m + k\sigma$, while σ remains constant, the most economical sample size $N = n$ will be a function of k. In particular, it will be shown that this function is

$$n = \frac{12.0}{k^2} \text{ when the control limits are } m \pm \frac{3.09\sigma}{\sqrt{n}}$$

$$n = \frac{11.1}{k^2} \text{ when the control limits are } m \pm \frac{3\sigma}{\sqrt{n}}$$

$$n = \frac{6.65}{k^2} \text{ when the control limits are } m \pm \frac{2.58\sigma}{\sqrt{n}}$$

$$n = \frac{4.4}{k^2} \text{ when the control limits are } m \pm \frac{2.33\sigma}{\sqrt{n}}.$$

Tables are calculated, giving the value of n for other control limits and giving also the average amount of inspection in each case.

Finally, a new chart, based on two sets of control limits, is discussed briefly.

2. The average amount of inspection for a given N. Let m be the original mean and σ the S.D. of the population, and let $m \pm B\sigma/\sqrt{N}$ be the control limits adopted for the arithmetic mean $\bar{x} = \sum x/N$ of a sample of given size N.

TABLE 1

N	$3.09 - 0.4\sqrt{N}$	P	S(N)	A(N)
1	2.690	0.0036	278	278
2	2.524	0.0058	173	345
3	2.396	0.0083	121	362
4	2.290	0.0110	91	364 Max
5	2.195	0.0141	71	355
9	1.890	0.0294	34	306
16	1.490	0.0681	14.7	235
25	1.090	0.1379	7.26	182
36	0.690	0.2451	4.08	147
49	0.290	0.3859	2.59	127
64	−0.110	0.5438	1.84	118
75	−0.375	0.6460	1.55	116 Min
81	−0.51	0.6950	1.44	117

If the population mean changes from $\mu = m$ to $\mu = m + k\sigma$ $(k > 0)$, the probability that $\bar{x}$ exceeds the upper control limit $m + B\sigma/\sqrt{N}$ is (assuming that σ remains unchanged)

$$P = P\left(\bar{x} \geqq m + \frac{B\sigma}{\sqrt{N}} \,\Big|\, \mu = m + k\sigma\right)$$

$$= P\left(\bar{x} - m - k\sigma \geqq \frac{B\sigma}{\sqrt{N}} - k\sigma \,\Big|\, \mu = m + k\sigma\right) \tag{1}$$

$$= P\left(\frac{\bar{x} - m - k\sigma}{\sigma/\sqrt{N}} \geqq B - k\sqrt{N} \,\Big|\, \mu = m + k\sigma\right) = P(z \geqq B - k\sqrt{N}),$$

where z is the standardized normal variate (mean zero and S.D. one).

Thus, when μ becomes equal to $m + k\sigma$, about $100P$ samples in every 100 samples, or one in every $1/P$ samples will give a mean $\bar{x}$ above the upper control limit. It follows that on the average $S(N) = 1/P$ samples, or $A(N) = N/P$ articles have to be tested before a change of the mean from m to $m + k\sigma$ can be expected to be revealed.

3. Example. For illustration, take the example $k = 0.4$, $B = 3.09$. We obtain Table 1 for various values of N (using normal probability tables).

The example shows the following interesting points:

(a). *Suppose that the population mean changes by* $+0.4$ *standard deviations.* If the chart is based on the customary sample size used for control charts [3], namely, $N = 4$ or $N = 5$, about 360 articles must be tested before detection of the change can be expected. On the other hand, if the control chart is based on a sample size between 50 and 80 (say), about 120 items only are required to indicate the change. The usual engineering practice of using charts for small samples requires thus about three times as much inspection as would be required with a chart for a suitably large sample size.

(b). *Suppose that the population mean does not change.* In that case $\bar{x}$ will fall above the upper control limit about once in 1000 samples. This means that with sample size 4 a "false alarm" will be raised about once for every 4000 articles tested. With a sample size 75, on the other hand, a "false alarm" will be raised only once for about 75,000 articles tested.

The two points raised suggest that in certain cases it may be of advantage to deviate from the usual practice of using small sample control charts. A third argument in favor of large samples is that the control limits are based on the assumption that $\bar{x}$ is normally distributed. This assumption is usually satisfied with great accuracy when the sample is large, but may not be justified when the sample is small.

The above arguments hold also for other values of k and B, and we may state that, unless special reasons exist for making the samples small, the sample size N should be chosen such that the average amount of inspection $A(N)$ becomes a minimum.

4. The minimum amount of inspection. We define the most economical sample size n as that value of N for which the average amount of inspection $A(N)$ required to detect a given change of the population mean becomes a minimum.

If the standardized normal probability density is denoted by $\varphi(z) = e^{-\frac{1}{2}z^2}/\sqrt{2\pi}$, we have

$$A(N) = N/P = N \Big/ \int_{B-k\sqrt{N}}^{\infty} \varphi(z)\, dz. \tag{2}$$

Differentiating with regard to N, we have

$$\frac{dA}{dN} = \frac{\int_{B-k\sqrt{N}}^{\infty} \varphi(z)\, dz - \frac{1}{2}k\sqrt{N}\,\varphi(B - k\sqrt{N})}{\left[\int_{B-k\sqrt{N}}^{\infty} \varphi(z)\, dz\right]^2}. \tag{3}$$

The condition for $A(N)$ to be a minimum is $(dA/dN)_{N=n} = 0$, which reduces to

$$P(u) = \int_{u}^{\infty} \varphi(z)\, dz = \tfrac{1}{2}(B - u)\varphi(u), \tag{4}$$

where $u = B - k\sqrt{n}$.

Equation (4) is easily solved, using tables of ordinates and integrals of the normal distribution. To do this, we put the minimum condition (4) in the form

$$Q = \frac{2P(u)}{\varphi(u)} = B - u. \tag{5}$$

The left side of this equation can then be calculated for any value of u and the values n, $S(n)$, and $A(n)$, can be deduced. We have

$$n = \frac{(B - u)^2}{k^2} = \frac{Q^2}{k^2}, \tag{6}$$

$$S(n) = \frac{1}{P}, \tag{7}$$

$$A(n) = \frac{n}{P} = \frac{Q^2}{Pk^2}. \tag{8}$$

The values of $S(n)$, $k^2A(n)$, and k^2n are shown in Table 2 for various values of u.

It can be seen from this table that for every value $B > 2.24$ two values N exist which satisfy the condition $dA/dN = 0$. Only the larger value of N, however, corresponds to a minimum amount of inspection. (The smaller value of N corresponds to a maximum of $A(N)$). Thus, for $B = 3$ we find $n_1 = 11.1/k^2$ and $n_2 = 0.606/k^2$. The amounts of inspection corresponding to n_1 and n_2 are $A(n_1) = 17.65/k^2$ and $A(n_2) = 46/k^2$ respectively, which shows that samples of size n_2 would lead to a much larger amount of inspection than samples of size n_1. The lower part of Table 2, corresponding to values of u greater than 0.6, can therefore be ignored.

Besides this, we notice that no values of u exist for which B is less than 2.24. This means that for such values of B no value of N exists which would make dA/dN equal to zero. This case will be discussed later (Section 6).

5. Discussion of the special case B = 3.09. The values $B = 3.09$, $B = 3$, $B = 2.58$, $B = 2.33$, are of special interest because they correspond to the most frequently used control limits. In particular, we have for $B = 3.09$: $n = 12.0/k^2$, $S(n) = 1.55$, $A(n) = 18.6/k^2$. The values of n and $A(n)$ for various values of k are tabulated below.

k	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4	1.6	1.8
n	300	133	75	54	37	28	21	17	12	8	6	5	4
$A(n)$	464	206	116	74	52	38	29	23	19	13	9.5	7.3	5.7

Alternatively, we may plot n and $A(n)$ as functions of k; they become straight lines when plotted on log log paper.

When the most economical sample size is taken, the average number of samples required for the detection of a change in the population mean is the same for all values of k; it is equal to $S(n) = 1.55$ when $B = 3.09$. This form of

TABLE 2

u	$P = \int_u^\infty \varphi(z)dz$	$\varphi(u)$	$Q = \frac{2P}{\varphi}$	$B = Q + u$	$nk^2 = Q^2$	$S(n) = \frac{1}{P}$	$A(n)k^2 = \frac{Q^2}{P}$
−0.5	0.692	0.352	3.92	3.42	15.4	1.45	22.3
−0.4	0.655	0.368	3.56	3.16	12.6	1.53	19.2
−0.37	0.6443	0.3725	3.46	3.09	12.0	1.55	18.6
−0.33	0.6293	0.3778	3.33	3.00	11.1	1.59	17.65
−0.3	0.618	0.381	3.24	2.94	10.3	1.62	17.0
−0.2	0.579	0.391	2.96	2.76	8.8	1.73	15.2
−0.1	0.540	0.397	2.72	2.62	7.4	1.85	13.7
−0.07	0.5279	0.3980	2.655	2.585	7.04	1.89	12.2
0.0	0.500	0.399	2.51	2.51	6.3	2.00	12.6
0.1	0.460	0.397	2.32	2.42	5.4	2.17	11.7
0.2	0.421	0.391	2.15	2.35	4.6	2.38	10.9
0.24	0.4052	0.3876	2.095	2.335	4.36	2.47	10.75
0.3	0.382	0.381	2.00	2.30	4.0	2.62	10.5
0.4	0.345	0.368	1.88	2.28	3.5	2.90	10.1
0.5	0.308	0.352	1.75	2.25	3.06	3.24	9.94
0.6	0.274	0.333	1.64	2.24	2.70	3.65	9.86
0.7	0.242	0.312	1.55	2.25	2.40	4.13	9.94
0.8	0.212	0.290	1.46	2.26	2.13	4.72	10.1
1.0	0.159	0.242	1.31	2.31	1.73	6.28	10.9
1.2	0.115	0.194	1.19	2.39	1.41	8.69	12.2
1.4	0.081	0.150	1.08	2.48	1.16	12.3	14.3
1.6	0.055	0.111	0.99	2.59	0.98	18.2	17.8
2.0	0.0228	0.054	0.844	2.844	0.71	43.8	31
2.22	0.0132	0.0339	0.78	3.00	0.606	76	46

control chart is thus very efficient, for it will indicate a change in about 2 out of 3 samples, whereas it will raise a false alarm (or type I error [4]) only in about one out of 1000 samples.

It appears from the above table that a control chart for small samples, say N between 4 and 10, is adequate only for the detection of changes of the mean greater than one standard deviation.

There is, of course, no need to adhere rigidly to the sample size given by the table, for in most cases the exact change (if any) of the population mean would not be known beforehand, but the table will give valuable information regarding the approximate size of the sample required.

Thus, in the case when $B = 3.09$, we would recommend the following sample sizes:

Change of Mean μ in S.D.'s	*Sample Size N*
0.2–0.3	100–300
0.3–0.4	70–150
0.4–0.5	50– 80
0.5–0.6	30– 60
0.6–0.7	25– 40
0.7–0.8	20– 30
0.8–1.0	10– 25

To detect changes larger than one standard deviation, any convenient size up to 10 could be taken.

6. The case when B < 2.24. When $B < 2.24$, the derivative dA/dN is different from zero for all values of N. This means that $A(N)$ has no relative minimum but increases with N for all values of N. The average amount of inspection is then smallest when $N = 1$, but, for reasons stated in Section 3, it is usually not advisable to take N smaller than, say, 4.

The only case of this type that might be of interest for the purpose of quality control is the case $B = 1.96$, because the control limits $m \pm 1.96\ \sigma/\sqrt{N}$ are often used as so-called inner limits [5]. The average amount of inspection is then (Section 2) $A(N) = N/P = N/P(z \geqq u)$, where $u = B - k\sqrt{N}$. This gives $N = (B - u)^2/k^2$ and $A(N) = (B - u)^2/(k^2P)$. Taking $B = 1.96$ and substituting different values for u, we obtain Table 3.

The table shows clearly that the average amount of inspection $A(N)$ decreases with the size N of the sample, that is, the smaller we make N the more economical will be the test.

7. A chart with two sets of control limits. When a chart with two sets of control limits is used, it is usually set up as follows.

After the mean m and S.D. σ of the population have been reliably estimated, a suitable sample size N is chosen, and inner limits $(m \pm 1.96\ \sigma/\sqrt{N})$ and outer limits $(m \pm 3.09\ \sigma/\sqrt{N})$ are calculated and entered in the chart. Production is stopped as soon as one $\bar{x}$ value falls outside the outer control limits. The main purpose of the inner limits is to provide a first warning when a point falls outside these limits. Production is then not interrupted, but samples are taken more frequently in order to reach a decision without delay.

Now we have seen that small samples are most suitable for the inner limits while larger samples should be taken for the outer limits. It seems therefore indicated to construct a chart involving two sample sizes. (If a smaller sample size is used for the inner control limits, the terms "inner" and "outer" become misleading because the inner limits will then actually be wider than the outer limits.) To detect a change of the population mean from m to $m + k\sigma$, a chart may be prepared as follows. (1) Take samples of size $N = 4$ or 5 (or any other

TABLE 3

$(B = 1.96)$

u	P	$B - u$	$Nk^2 = (B - u)$	$A(N)k^2$	$S(N) = \frac{1}{P}$
1.76	.0392	0.2	0.04	1.0	25.5
1.66	.0485	0.3	0.09	1.9	20.6
1.56	.0594	0.4	0.16	2.7	16.8
1.46	.0721	0.5	0.25	3.5	13.9
1.36	.0869	0.6	0.36	4.1	11.5
1.26	.104	0.7	0.49	4.7	9.6
1.16	.123	0.8	0.64	5.2	8.1
1.06	.145	0.9	0.81	5.6	6.9
0.96	.168	1.0	1.00	6.0	6.0
0.76	.224	1.2	1.44	6.4	4.5
0.56	.288	1.4	1.96	6.8	3.5
0.36	.359	1.6	2.56	7.1	2.8
0.16	.436	1.8	3.24	7.4	2.3
−0.04	.516	2.0	4.00	7.8	1.9
−0.54	.705	2.5	6.25	8.9	1.4

convenient small sample size) and construct a chart with control limits $m \pm 1.96\sigma/\sqrt{N}$. (2) Calculate a second set of control limits $m \pm 3.09\sigma/\sqrt{n}$, based on a sample size $n = \lambda N$, which is a multiple of N as close as possible to the most economical sample size $12.0/k^2$. (3) Calculate the means $\bar{x} = \sum x/N$ and $\bar{X} = \sum \bar{x}/\lambda$. (4a) If a value $\bar{X}$ falls outside the limits $m \pm 3.09\sigma/\sqrt{n}$, stop production and investigate; (4b) if a value $\bar{x}$ falls outside the limits $m \pm 1.96\sigma/\sqrt{N}$, do not interrupt production but take out samples frequently to reach a decision.

While it is true that the "inner" limits serve mainly to provide a first warning for a possible change of the population mean, they may be used also to reach a definite decision. If, for instance, two successive values of $\bar{x}$ fall above the upper inner limit, we may regard this as a significant indication for a change in the population mean, because the probability for this to happen when the population mean is unchanged is only $(0.025)^2 = 0.0006$, which is less than the probability that a single value $\bar{X}$ falls above the upper outer limit $m + 3.09\sigma/\sqrt{n}$.

Again, it is easily shown that the probability that 4 out of 16 successive $\bar{X}$ values fall above the upper inner limit is about the same as the probability that a single $\bar{x}$ value falls above the upper outer limit. We are therefore justified in regarding such an occurrence as a significant indication for a change of the population mean.

8. Conclusion. When a single set of control limits is used for a chart controlling the mean of a population, the above theory leads to much larger samples than those usually taken in industry. However, even with a chart of this type, small samples are not always uneconomical, for there are other factors to be considered which are not covered in this paper.

For instance, it may be of advantage to divide samples into subgroups in order to detect changes due to definite anticipated causes, necessitating the use of smaller samples.

Again, if a change of the population mean from m to $m + k\sigma$ is anticipated and the sample size is determined accordingly, any unsuspected larger change would in the average be detected later than if a smaller sample size had been used.

On the other hand, if small changes of the population mean of a given order are anticipated and if it is unlikely that larger changes occur, the sample size should be calculated according to the above theory.

When it is convenient to use a more elaborate chart, containing two sets of control limits, the theory leads to the customary small samples for the one set $(m \pm 1.96\sigma/\sqrt{N})$ and to the above large samples for the other set $(m \pm 3.09\sigma/\sqrt{n})$. Any unexpected larger change of the population mean is then likely to be detected by means of the small samples.

REFERENCES

[1] W. A. Shewhart, *The Economic Control of Quality of a Manufactured Product*, D. Van Nostrand, 1931.

[2] W. A. Shewhart, *Contribution of Statistics to the Science of Engineering*, University of Pennsylvania Press, 1940.

[3] W. B. Rice, *Control Charts in Factory Management*, John Wiley and Sons, 1947.

[4] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses. I. Unbiassed critical regions of Type A and Type A_1," *Stat. Res. Memoirs*, Vol. 1 (1936), pp. 1–37.

[5] B. P. Dudding and W. J. Jennett, *Quality Control Charts*, No. 600R, British Standards Institution, 1942.

[6] American Society for Testing Materials, *ASTM Manual on Quality Control of Materials*, Philadelphia, 1951.

[7] S. S. Wilks, "Determination of sample size for setting tolerance limits," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 91–96.