International Journal of Pure and Applied Mathematics
Volume 13 No. 2 2004, 249-258

S^2 CONTROL CHART BASED ON DOUBLE SAMPLING

Michael B.C. Khoo

School of Mathematical Sciences
University of Science, Malaysia - USM
11800 USM, Penang, MALAYSIA
e-mail: mkbc@usm.my

Abstract: A double sampling (DS) $\overline{X}$ chart is proposed by Daudin [2] as an alternative to the classical Shewhart $\overline{X}$ chart for the monitoring of the process mean. The DS $\overline{X}$ chart gives better efficiency than the Shewhart $\overline{X}$ chart. When using a DS $\overline{X}$ chart, a decision about the state of a process is made based on either the information in the first sample alone or the combined information of samples one and two. In this paper a double sampling (DS) S^2 chart will be proposed for the monitoring of the process variance.

AMS Subject Classification: 49J20, 93C20, 49K20, 35K20
Key Words: double sampling (DS) $\overline{X}$ chart, double sampling (DS) S^2 chart, in-control, out-of-control

1. Introduction

In the monitoring of a process variance, it is more meaningful to use a double sampling (DS) chart for process dispersion as a counterpart to the DS $\overline{X}$ chart. However, all presently available control charts for process dispersion, namely the R, S and S^2 charts [4], [3] and [1] are based on a single sampling approach. Thus, the aim of this paper is to propose a DS S^2 chart for process dispersion. The DS S^2 chart can be used hand in hand with a DS $\overline{X}$ chart so that both the process dispersion and process mean can be monitored simultaneously.

Received: March 25, 2004

2. The Double Sampling (DS) S^2 Chart

Similar to the other control charts for process dispersion, a DS S^2 chart requires the assumption that the sample variances are independent with a common variance. The same procedure which is used in the DS $\overline{X}$ chart is used here. First, take a sample of size n_1 and compute the sample variance, S_1^2 using the formula:

$$S_1^2 = \frac{\sum_{i=1}^{n_1} X_i^2 - n_1 \overline{X_1}^2}{n_1 - 1}\,, \tag{1}$$

where $\overline{X_1} = \frac{\sum_{i=1}^{n_1} X_i}{n_1}$ is the estimated mean of the first sample. Let μ and σ_0^2 denote the process mean and variance from an in-control process respectively. Assume that the process follows a normal distribution. There are three possibilities during the first stage (see Figure 1):

(*i*) If S_1^2 is smaller than L_1, conclude that the process is in-control.

(*ii*) If S_1^2 is greater than L_2, conclude that the process is out-of-control.

(*iii*) If S_1^2 is greater than L_1 but is less than L_2, a second sample is taken.

After stage one, if a second sample is required, then a sample of size n_2 is taken at the second stage. Compute the sample variance, S_2^2, using the formula:

$$S_2^2 = \frac{\sum_{i=1}^{n_2} X_i^2 - n_2 \overline{X_2}^2}{n_2 - 1}\,, \tag{2}$$

where $\overline{X_2}$ is the estimated mean of the second sample. Decisions at the second stage are based on the pooled sample variance computed using the formula:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\,. \tag{3}$$

At this stage, there are only two possibilities:

(*i*) If S_p^2 is smaller than L_3, the process is in-control.

(*ii*) If S_p^2 is greater than L_3, the process is out-of-control.

The construction of a DS S^2 chart depends on the values of five parameters, i.e., n_1, n_2, L_1, L_2 and L_3. The following guidelines are considered in the design of a DS S^2 chart:

(*i*) L_2 must be higher than the upper control limit of a classical S^2 chart, given by $UCL = \frac{\sigma_0^2}{n-1}\chi^2_{n-1,\alpha}$, where n and α are the sample size and Type-I error of a classical S^2 chart respectively.

(*ii*) L_1 must be lower than the classical upper control limit of a S^2 chart given in (i) above.

(*iii*) The ratio between n_2 and n_1 is about $n_2 = 2n_1$ or $n_2 = 3n_1$.

3. Characteristics of the Double Sampling (DS) S^2 Chart

It is assumed that the observations in each sample are independent and normally distributed. When the process shifts, its standard deviation changes from σ_0 to a larger value σ while the process mean remains constant. The following four intervals which are shown in Figure 1 are defined as follows:

(*i*) $I_1 = [0, L_1]$.

(*ii*) $I_2 = (L_1, L_2]$.

(*iii*) $I_3 = (L_2, \infty)$.

(*iv*) $I_4 = [0, L_3]$.

For a stable process, let P_a denotes the probability that the process is in-control while P_{a1} and P_{a2} be the probabilities that the process is in-control at stages one and two respectively. Thus, $P_a = P_{a1} + P_{a2}$. Here,

$$\begin{aligned} P_{a1} &= Pr(S_1^2 \in I_1) \\ &= Pr(S_1^2 < L_1) \\ &= Pr[(\frac{n_1-1}{\sigma_0^2})S_1^2 < (\frac{n_1-1}{\sigma_0^2})L_1] \\ &= H_{n_1-1}[(\frac{n_1-1}{\sigma_0^2})L_1]\,, \end{aligned} \tag{4}$$

where $H_v(\cdot)$ is the chi-square distribution function with v degrees of freedom. Also,

$$\begin{aligned} P_{a2} &= Pr(S_p^2 \in I_4 \text{ and } S_1^2 \in I_2) \\ &= \int_{s\in I_2} Pr(S_p^2 \in I_4 \mid S_1^2 = s) f(s)ds\,, \end{aligned} \tag{5}$$

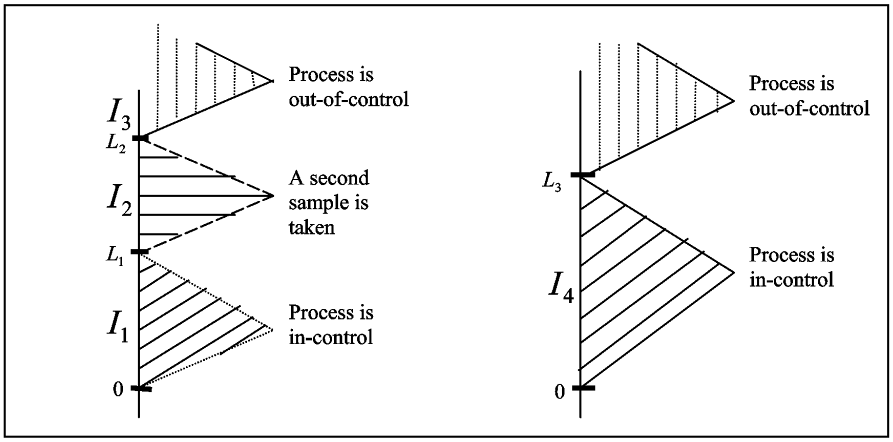

Figure 1: Graphical view of DC S^2 chart

where $f(\cdot)$ is the probability density function of S_1^2. P_{a2} in eq. (5) can be computed using a *Mathematica* 4.0 program based on the integration shown in Appendix. Similar to a DS $\overline{X}$ chart, the run length of a DS S^2 chart is defined as the number of different time points at which samples are taken. Thus, for an in-control process, the average run length is given by:

$$ARL_0 = \frac{1}{1 - P_a}\,. \tag{6}$$

4. The Simulation Study

A simulation study is conducted using *SAS*, version 8, to evaluate the performance of the proposed DS S^2 chart. In-control observations are generated from a normal, $N(\mu, \sigma_0^2)$ distribution, where $\mu = 0$ and $\sigma_0^2 = 1$. For an out-of-control process, the observations are generated from a $N(\mu, \sigma_1^2)$ distribution. Here, $\sigma_1^2 > \sigma_0^2$, where $\sigma_1 = \delta\sigma_0$ for $\delta \in \{1, 1.05, 1.1, 1.25, 1.5, 2, 2.5, 3, 4,\ 5, 6, 10\}$. In-control ARLs ($ARL_0s$) of 200 and 500 are considered. For $ARL_0 = 200$, the upper control limit of a classical S^2 chart having sample size, $n = 5$, is $UCL = \frac{\sigma_0^2}{n-1}\chi^2_{4,\frac{1}{200}} = \frac{14.86}{4} = 3.715$. Similarly, its UCL is $\frac{\chi^2_{4,\frac{1}{500}}}{4} = \frac{16.92}{4} = 4.23$ for an in-control ARL of 500.

A DS S^2 chart will be designed based on the above information. The DS S^2 chart will have a sample of size $n_1 = 3$ at stage one and a sample of size

$n_2 = 6$ at stage two. Since for $ARL_0 = 200$ the UCL of the above classical S^2 chart with a sample of size $n = 5$ is 3.715, then the L_1 limits of the DS S^2 chart considered are $2, 3$ and 3.5 where all these values are lower than 3.715. The same value of $L_3 = 2.7$ is used for the three cases of the DS S^2 chart. Based on the values of L_1 and L_3, the limit L_2 (see Table 1) is determined using a *Mathematica 4.0* program given in Appendix so that $ARL_0 = 200$. Note that other values of L_3 can also be considered. For the case where $ARL_0 = 500$, the limit of L_1 must be set lower than 4.23. Here, the values of L_1 considered are $2.5, 3$ and 4. A value of $L_3 = 4$ is used for the three cases of different L_1 values. The corresponding L_2 limits for these three cases are determined using *Mathematica 4.0* so that $ARL_0 = 500$.

The simulation results are displayed in Table 1 and Table 2 for ARL_0s of 200 and 500 respectively. The simulation results give the ARL profiles and the average sample size at each time point for the three different cases in each Table. The ARL profiles in both Table 1 and Table 2 show that the ARL values decrease as the magnitude of the shifts increase. The average sample size on the other hand increases from $\delta = 1$ onwards until $\delta = 2$ and then decreases thereafter. For a stable process, i.e., $\delta = 1$, the first sample is usually sufficient to conclude that the process is in-control because there is a high probability of $S_1^2 \in I_1$. However, for moderate shifts, a higher probability of $S_1^2 \in I_2$ is the case since the results show that a larger average sample size is required compared to that of the other magnitude of shifts. For very large shifts, say $\delta = 6$ or 10, the average sample size is close to 3 because the probability of $S_1^2 \in I_3$ is high.

	$\delta = \frac{\sigma}{\sigma_0}$	ARL	Average Sample Size
	1	204.56	3.09
	1.05	113.17	3.13
	1.1	69.52	3.16
$n_1 = 3$	1.25	23.33	3.28
$n_2 = 6$	1.5	7.95	3.40
$L_1 = 3.5$	2	3.13	3.43
$L_2 = 5.75$	2.5	2.13	3.36
$L_3 = 2.7$	3	1.74	3.28
	4	1.39	3.19
	5	1.25	3.12
	6	1.17	3.09
	10	1.06	3.03
	1	204.34	3.15
	1.05	112.62	3.21
	1.1	68.17	3.25
$n_1 = 3$	1.25	22.23	3.39
$n_2 = 6$	1.5	7.57	3.53
$L_1 = 3$	2	3.03	3.57
$L_2 = 6$	2.5	2.09	3.47
$L_3 = 2.7$	3	1.73	3.37
	4	1.39	3.24
	5	1.25	3.16
	6	1.18	3.12
	10	1.06	3.04
	1	206.18	3.38
	1.05	108.20	3.45
	1.1	65.42	3.52
$n_1 = 3$	1.25	20.62	3.68
$n_2 = 6$	1.5	6.96	3.83
$L_1 = 2$	2	2.86	3.82
$L_2 = 6.7$	2.5	2.05	3.68
$L_3 = 2.7$	3	1.72	3.56
	4	1.41	3.37
	5	1.27	3.27
	6	1.19	3.19
	10	1.07	3.07

Table 1. ARL profiles for the DS S^2 chart based on $ARL_0 = 200$

	$\delta = \frac{\sigma}{\sigma_0}$	ARL	Average Sample Size
	1	505.28	3.05
	1.05	282.57	3.07
	1.1	171.65	3.09
$n_1 = 3$	1.25	50.37	3.16
$n_2 = 6$	1.5	13.99	3.28
$L_1 = 4$	2	4.04	3.36
$L_2 = 6.239$	2.5	2.40	3.33
$L_3 = 4$	3	1.87	3.27
	4	1.44	3.18
	5	1.27	3.12
	6	1.19	3.09
	10	1.07	3.03
	1	520.52	3.13
	1.05	291.37	3.17
	1.1	175.97	3.21
$n_1 = 3$	1.25	51.04	3.33
$n_2 = 6$	1.5	13.62	3.47
$L_1 = 3$	2	3.88	3.53
$L_2 = 6.25$	2.5	2.32	3.47
$L_3 = 4$	3	1.81	3.39
	4	1.42	3.25
	5	1.26	3.17
	6	1.18	3.13
	10	1.07	3.05
	1	521.40	3.22
	1.05	295.22	3.27
	1.1	179.12	3.32
$n_1 = 3$	1.25	51.13	3.44
$n_2 = 6$	1.5	13.54	3.58
$L_1 = 2.5$	2	3.80	3.62
$L_2 = 6.24$	2.5	2.28	3.54
$L_3 = 4$	3	1.79	3.44
	4	1.41	3.29
	5	1.25	3.20
	6	1.18	3.15
	10	1.06	3.05

Table 2. ARL profiles for the DS S^2 chart based on $ARL_0 = 500$

5. Conclusion

The DS S^2 chart share some of the advantages of a double sampling plan which are discussed in [4]. Since that for a DS S^2 chart the first sample taken at stage one is usually smaller than the sample that would be required using a classical S^2 chart, then if the process is accepted or rejected on the first sample, the cost of sampling will be lower for a DS S^2 chart than it would be for a classical S^2 chart. In some situations a DS S^2 chart also has an advantage of giving a process a second chance if the sample variance, S_1^2, computed from the first sample falls between L_1 and L_2. The design of a DS S^2 chart depends on five parameters, namely, n_1, n_2, L_1, L_2 and L_3. Future research should consider finding an approach to determine the optimal values of these five parameters so that the DS S^2 chart can be made to provide a quick detection for a shift of a desired magnitude.

Acknowledgements

This research is supported by the University Science of Malaysia "Fundamental Research Grant Scheme (FRGS)", No. 304/pmaths/670039.

References

[1] L.C. Alwan, *Statistical Process Analysis*, McGraw-Hill, New York (2000).

[2] J.J. Daudin, Double sampling $\overline{X}$ charts, *Journal of Quality Technology*, **24** (1992), 78-87.

[3] E.L. Grant, R.S. Leavenworth, *Statistical Quality Control*, 5-th Edition, McGraw-Hill, New York (1980).

[4] D.C. Montgomery, *Introduction to Statistical Quality Control*, 4-th Edition, Wiley, New York (2001).

Appendix

It will now be shown how the probability P_{a2} in equation (5) is computed. Since S_1^2 and S_2^2 are independent, from equation (3) we have

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \sim \frac{\sigma_0^2}{n_1+n_2-2}\chi^2_{n_1+n_2-2}\,.$$

First, consider the probability $S_p^2 \in I_4$. Here,

$$\begin{aligned} Pr(S_p^2 \in I_4) &= Pr(S_p^2 \le L_3) \\ &= Pr[\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \le L_3] \\ &= Pr[(n_1-1)S_1^2 + (n_2-1)S_2^2 \le (n_1+n_2-2)L_3]. \end{aligned}$$

It follows that,

$$\begin{aligned} Pr(S_p^2 \in I_4 \mid S_1^2 = s) &= Pr[(n_2-1)S_2^2 \le (n_1+n_2-2)L_3 - (n_1-1)s] \\ &= Pr[\frac{(n_2-1)S_2^2}{\sigma_0^2} \le \frac{(n_1+n_2-2)L_3-(n_1-1)s}{\sigma_0^2}] \\ &= H_{n_2-1}\{\frac{(n_1+n_2-2)L_3-(n_1-1)s}{\sigma_0^2}\}. \end{aligned}$$

Then, from equation (5),

$$P_{a2} = \int_{s \in I_2} [H_{n_2-1}\{\frac{(n_1+n_2-2)L_3-(n_1-1)s}{\sigma_0^2}\}] f(s)ds\,,$$

where $f(\cdot)$ is the probability density function of S_1^2. Thus,

$$P_{a2} = \int_{s \in I_2} [\int_0^{\frac{(n_1+n_2-2)L_3-(n_1-1)s}{\sigma_0^2}} \frac{1}{\Gamma(\frac{n_2-1}{2})}(\frac{1}{2})^{\frac{n_2-1}{2}} t^{\frac{n_2-1}{2}-1} e^{\frac{-t}{2}} dt] \times f(s)ds. \quad (7)$$

To find the density function of S_1^2, let

$$W = (n_1-1)\frac{S_1^2}{\sigma_0^2} \sim \chi^2_{n_1-1}. \quad (8)$$

Then, the density function of W is

$$g(w) = \frac{1}{\Gamma(\frac{n_1-1}{2})}(\frac{1}{2})^{\frac{n_1-1}{2}} w^{\frac{n_1-1}{2}-1} e^{-\frac{w}{2}} I_{(0,\infty)}(w).$$

From (8),

$$S_1^2 = u(w) = \frac{\sigma_0^2}{n_1-1} W.$$

It follows that

$$W = u^{-1}(S_1^2) = (n_1-1)\frac{S_1^2}{\sigma_0^2}.$$

n1=
n2=
L1=
L2=
L3=
V=

$$\mathbf{Pa2}=\left(\frac{\mathbf{n1-1}}{\mathbf{V}}\right)^{\frac{\mathbf{n1-1}}{\mathbf{2}}}\times\frac{\mathbf{1}}{\mathbf{Gamma}\left[\frac{\mathbf{n1-1}}{\mathbf{2}}\right]\times\mathbf{Gamma}\left[\frac{\mathbf{n2-1}}{\mathbf{2}}\right]}$$

$$\mathbf{NIntegrate[Integrate[}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^{\frac{\mathbf{n1+n2-2}}{\mathbf{2}}}\mathbf{t}^{\frac{\mathbf{n2-1}}{\mathbf{2}}-\mathbf{1}}\mathbf{e}^{-\frac{\mathbf{t}}{\mathbf{2}}},$$

$$\mathbf{\{t, 0, }\frac{\mathbf{(n1+n2-2)L3-(n1-1)s}}{\mathbf{V}}\mathbf{\}]s}^{\frac{\mathbf{n1-3}}{\mathbf{2}}}\mathbf{e}^{-\frac{\mathbf{(n1-1)\times s}}{\mathbf{2V}}}\mathbf{, \{s, L1, L2\}]}$$

Figure 2: *Mathematica 4.0* program to compute P_{a2}

Therefore,

$$\begin{aligned} f(s) &= f(S_1^2 = s) \\ &= g[u^{-1}(S_1^2 = s)] \mid \frac{d}{ds} u^{-1}(S_1^2 = s) \mid \\ &= g[\frac{(n_1-1)s}{\sigma_0^2}] \mid \frac{d}{ds}[\frac{(n_1-1)s}{\sigma_0^2}] \mid \\ &= \frac{1}{\Gamma(\frac{n_1-1}{2})}(\frac{1}{2})^{\frac{n_1-1}{2}}(\frac{n_1-1}{\sigma_0^2})^{\frac{n_1-1}{2}-1} s^{\frac{n_1-1}{2}-1} e^{-\frac{(n_1-1)s}{2\sigma_0^2}}(\frac{n_1-1}{\sigma_0^2}) \\ &= \frac{1}{\Gamma(\frac{n_1-1}{2})}(\frac{1}{2})^{\frac{n_1-1}{2}}(\frac{n_1-1}{\sigma_0^2})^{\frac{n_1-1}{2}} s^{\frac{n_1-3}{2}} e^{-\frac{(n_1-1)s}{2\sigma_0^2}} I_{(0,\infty)}(s). \end{aligned}$$

The following *Mathematica 4.0* program in Figure 2 can be used to compute P_{a2} in equation (7): Note that σ_0^2 in equation (7) is represented by V in Figure 2. The limit, L_2, can be determined if the values of n_1, n_2, L_1, L_3 and P_{a2} are known. Here, P_{a2} is computed using the formula $P_{a2} = P_a - P_{a1}$ where P_a is determined based on a desired Type-I error while P_{a1} is obtained using equation (4).