An Introduction to Statistical Process Control

F. F. Gan
Department of Mathematics
National University of Singapore

Introduction

The basic idea in statistical process control (SPC) is to take random samples of products from a manufacturing line and examine the products to ensure that certain criteria of quality are satisfied. If the products sampled are found to be of inferior quality, then the manufacturing process is checked to seek out assignable causes of inferior quality to bring the process back to control.

In the past when products were hand-made one at a time, SPC was a redundant concept as all the products made were checked for conformity to certain criteria of quality. An example of early manufacturing processes is coin molding whereby melted metal is poured into a mold. When the metal hardens, the coins are removed and polished. Any defect found would immediately be traced back to these assignable causes: (i) the mold is defective, (ii) the metal used is of inferior quality, (iii) the temperature of the melted metal is too high or too low, (iv) the worker is not skillful enough, or (v) other causes that could be found easily.

The Industrial Revolution brought along machines that can manufacture products at breakneck speeds, so that 100% inspection is virtually impossible (at least in the past) or very costly. The procedure of taking random samples of products from a manufacturing line at intervals and checking the quality is a reasonable alternative. This procedure must have been practised for a long time before Shewhart (1931) formally proposed a graphical procedure for monitoring quality. The proposed procedure is now commonly known as the Shewhart control chart.

Consider a manufacturing process producing tennis balls where the weight of a ball is a key measure of quality. According to official tennis rules, the weight of a tennis ball should be 2 ounces. In order to monitor the quality, random samples of five balls each, for example, are taken from a manufacturing line at regular intervals. The weight of each ball is measured and the sample mean of the measurements is plotted against the time or the sample number. An example of such a graph is displayed in Figure 1. Three horizontal lines representing lower, upper control limits (LCL and UCL) and the target weight are also displayed. Each plotted point gives an indication of whether the process is in control or not. The process is considered to be in control if a plotted point is within the two chart limits. Any point plotted below the LCL or above the UCL is considered an out-of-control point and gives an indication that the process could be out-of-control. Any unusual pattern like 6 or 7 consecutive points above the target value or a cyclical pattern of points also indicates that the process could be out of control. The 3rd to 9th, 51th to 57th and 63rd to 68th points are all above the target value providing evidence that the process could be out of control at the early and also the latter stages.

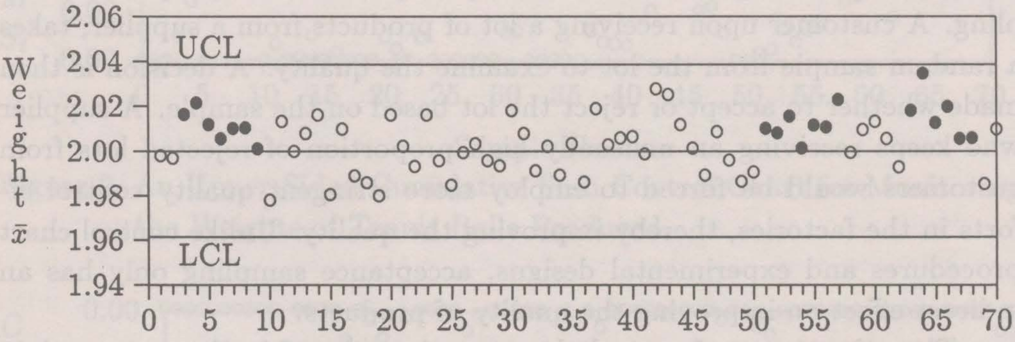

Figure 1. A Shewhart Control Chart for Monitoring the Weights of Tennis Balls Produced.

If out-of-control points are found, the process should be checked immediately for assignable causes and if found, should be removed to bring the process back to control. Sometimes, it is also likely that no assignable cause can be found and that the out-of-control points could be a result of some inherent or natural variations of the process. The quality control engineers should continue to be on alert until the next few points plotted are all within control or no unusual pattern is found. It is prudent to increase the frequency of sampling at this alert state.

It is important to understand that a control chart is merely a procedure for monitoring the quality of products from a manufacturing line, thereby keeping the manufacturing process in control. In other words, a control chart is only capable of removing assignable causes of variations in the quality but not the causes of random variations inherent of a manufacturing process. For example, consider coin molding, the coins produced can only be as good as the mold. If coins with sharper images are required, then the entire mold has to be redesigned. Statistical analyses performed on manufacturing processes to seek out optimal operating conditions, hence reducing the variance in the quality are generally known as experimental designs. The well known Taguchi method which is credited for much of the successes in Japanese industries falls under the category of experimental designs.

The third category of statistical quality control is acceptance sampling. A customer upon receiving a lot of products from a supplier, takes a random sample from the lot to examine the quality. A decision is then made whether to accept or reject the lot based on the sample. A supplier who keeps receiving an unusually high proportion of rejected lots from customers would be forced to employ more stringent quality control efforts in the factories, thereby improving the quality. Unlike control chart procedures and experimental designs, acceptance sampling only has an indirect effect on improving the quality of products.

The other types of control charts are introduced in the next section. An application of control chart procedures to the certificate of entitlement data set is then considered.

Statistical Control Charts

There are four main types of control charts commonly used in manufacturing lines. According to the popularity of usage, they are (1) Shewhart, (2) cumulative sum (CUSUM), (3) exponentially weighted moving average (EWMA) and (4) straight moving average (SMA) charts. The Shewhart chart has already been introduced in the previous section.

Let $\bar{x}_1, \bar{x}_2, \ldots$ be a sequence of independent and identically distributed sample means. If μ_0 is the in-control process mean, then the expected value of $\bar{x}_t - \mu_0$ is zero. Thus, without loss of generality, the target mean is assumed to be zero. The upper-sided and lower-sided CUSUM charts proposed by Page (1954) are obtained by plotting

$$S_t = \max\{0, S_{t-1} + (\bar{x}_t - k)\}$$

and

$$T_t = \min\{0, T_{t-1} + (\bar{x}_t + k)\}$$

against the sample number t for $t = 1, 2, \ldots$, respectively where the chart parameter k is a suitably chosen positive constant. The initial starting values S_0 and T_0 are usually chosen to be zero.

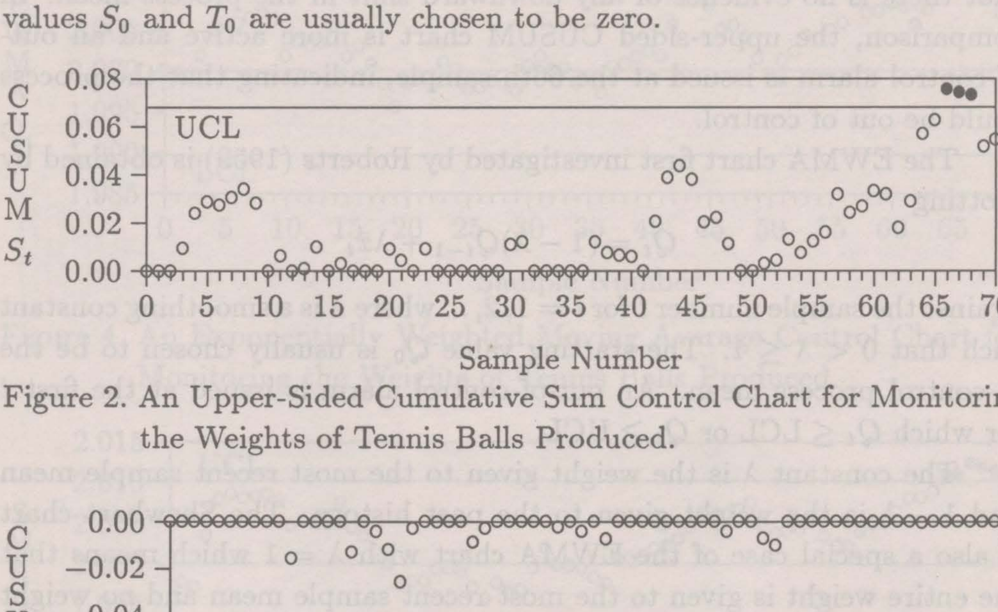

Figure 2. An Upper-Sided Cumulative Sum Control Chart for Monitoring the Weights of Tennis Balls Produced.

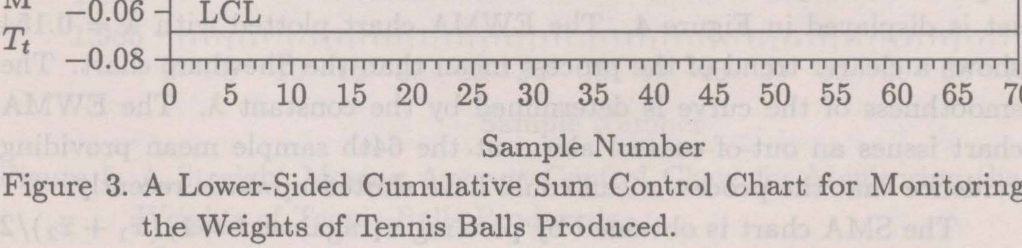

Figure 3. A Lower-Sided Cumulative Sum Control Chart for Monitoring the Weights of Tennis Balls Produced.

The upper-sided CUSUM chart is intended to detect an upward shift in the mean and it issues an out-of-control alarm at the first t for which $S_t \geq$ UCL. Similarly, the lower-sided CUSUM chart is intended to detect a downward shift in the mean and it issues an out-of-control alarm at the first t for which $T_t \leq$ LCL. The upper-sided CUSUM chart remains inactive (that is, $S_t = 0$) as long as $\bar{x}_t < k$ which means that a CUSUM chart with a large value of k would not be sensitive to small shifts in the process mean. Thus, the value of k can be chosen such that it is optimal in detecting a particular shift in the mean which is usually taken to be the smallest shift that is the least tolerable. The Shewhart chart is a special case of the CUSUM chart obtained by setting both LCL and UCL to zero.

The upper-sided and lower-sided CUSUM charts constructed for the tennis balls' data set are displayed in Figures 2 and 3 respectively. The lower-sided CUSUM chart remains inactive most of the time indicating that there is no evidence of any downward shift in the process mean. In comparison, the upper-sided CUSUM chart is more active and an out-of-control alarm is issued at the 66th sample, indicating that the process could be out of control.

The EWMA chart first investigated by Roberts (1959) is obtained by plotting

$$Q_t = (1 - \lambda)Q_{t-1} + \lambda \bar{x}_t$$

against the sample number t for $t = 1, 2, \ldots$ where λ is a smoothing constant such that $0 < \lambda \leq 1$. The starting value Q_0 is usually chosen to be the in-control process mean. An out-of-control alarm is issued at the first t for which $Q_t \leq$ LCL or $Q_t \geq$ UCL.

The constant λ is the weight given to the most recent sample mean and $1 - \lambda$ is the weight given to the past history. The Shewhart chart is also a special case of the EWMA chart with $\lambda = 1$ which means that the entire weight is given to the most recent sample mean and no weight is given to the past history. The EWMA chart for the tennis balls' data set is displayed in Figure 4. The EWMA chart plotted with $\lambda = 0.154$ shows a clearer trend of the process mean than the Shewhart chart. The smoothness of the curve is determined by the constant λ. The EWMA chart issues an out-of-control alarm at the 64th sample mean providing evidence that the process mean could have shifted upwards recently.

The SMA chart is obtained by plotting $\bar{x}_1$ against $t = 1$, $(\bar{x}_1 + \bar{x}_2)/2$

against $t = 2$, ..., $(\bar{x}_1 + \bar{x}_2 + ... + \bar{x}_{m-1})/(m-1)$ against $t = m - 1$ and the average of the m most recent sample means

$$M_t = (\bar{x}_{t-(m-1)} + \bar{x}_{t-(m-2)} + ... + \bar{x}_{t-1} + \bar{x}_t)/m$$

against t for $t = m, m+1, \ldots$. An out-of-control alarm is issued at the first t for which $M_t \leq$ LCL or $M_t \geq$ UCL. The Shewhart chart is also a special case of the SMA chart with $m = 1$. The SMA chart with $m = 12$ for the tennis balls' data set is displayed in Figure 5. The smoothness of the curve of a SMA chart is determined by the constant m. The SMA chart also issues an out-of-control alarm at the 64th sample. A comparison of the EWMA and SMA charts shows that both are very similar in showing the trend of the mean.

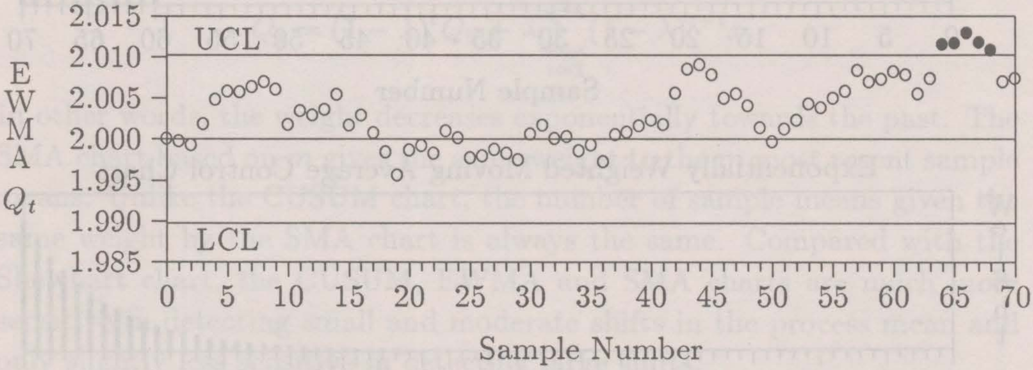

Figure 4. An Exponentially Weighted Moving Average Control Chart for Monitoring the Weights of Tennis Balls Produced.

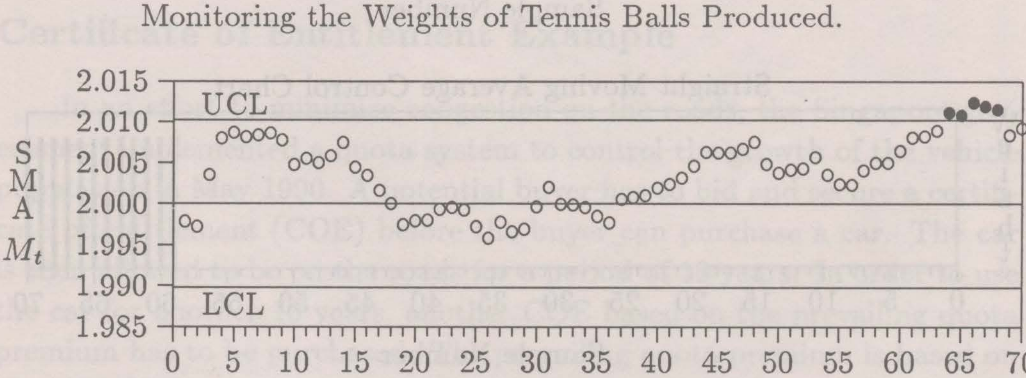

Figure 5. A Straight Moving Average Control Chart for Monitoring the Weights of Tennis Balls Produceds

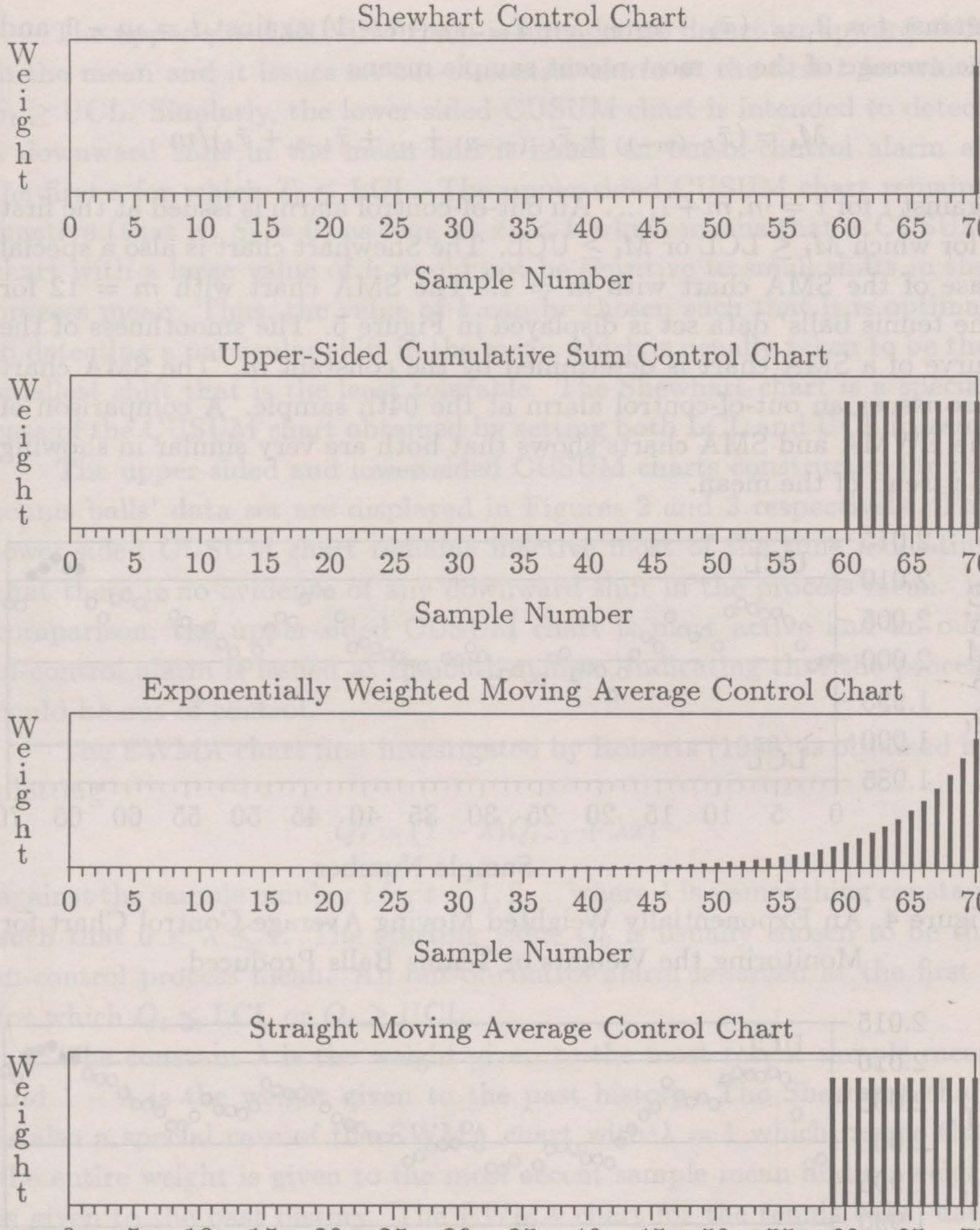

Figure 6. Allocations of Weights to the Sample Means by the Various Control Charts for the Tennis Balls' Data Set.

The weights allocated to the sample means by the various control charts for the tennis balls' data set are displayed in Figure 6. Graphs similar to Figure 6 were first given by Hunter (1986). A Shewhart chart puts all its weight on the most recent sample mean and completely ignores the past history and it is thus sensitive to large shifts in the process mean. The CUSUM chart is the most 'intelligent' in the sense that when it becomes active on encountering a sample mean greater than k, it begins to put equal weights on the present and all future sample means until it becomes inactive again or issues an out-of-control alarm. The EWMA chart gives the largest weight λ to the most recent sample mean, a weight of $\lambda(1-\lambda)$ to the next most recent sample mean and so on according to the coefficient of $\bar{x}$ in the equation

$$Q_t = (1-\lambda)^t Q_0 + \lambda \sum_{i=1}^{t} (1-\lambda)^{t-i} \bar{x}_i.$$

In other words, the weight decreases exponentially towards the past. The SMA chart based on m gives the same weight to the m most recent sample means. Unlike the CUSUM chart, the number of sample means given the same weight by the SMA chart is always the same. Compared with the Shewhart chart, the CUSUM, EWMA and SMA charts are much more sensitive in detecting small and moderate shifts in the process mean and only slightly less sensitive in detecting large shifts.

Certificate of Entitlement Example

In an effort to minimize congestion on the roads, the Singapore government implemented a quota system to control the growth of the vehicle population in May 1990. A potential buyer has to bid and secure a certificate of entitlement (COE) before the buyer can purchase a car. The car is then allowed to be on the roads for a period of 10 years. In order to use the car for another 10 years, another COE based on the prevailing quota premium has to be purchased. The prevailing quota premium is based on the straight moving average of the quota premiums of the most recent 12 months. The quota premiums for vehicles with engine capacities in the range 1001–1600 cc since May 1990 till August 1993 are listed in Table 1. A plot of the quota premium against the month is displayed in Figure 7.

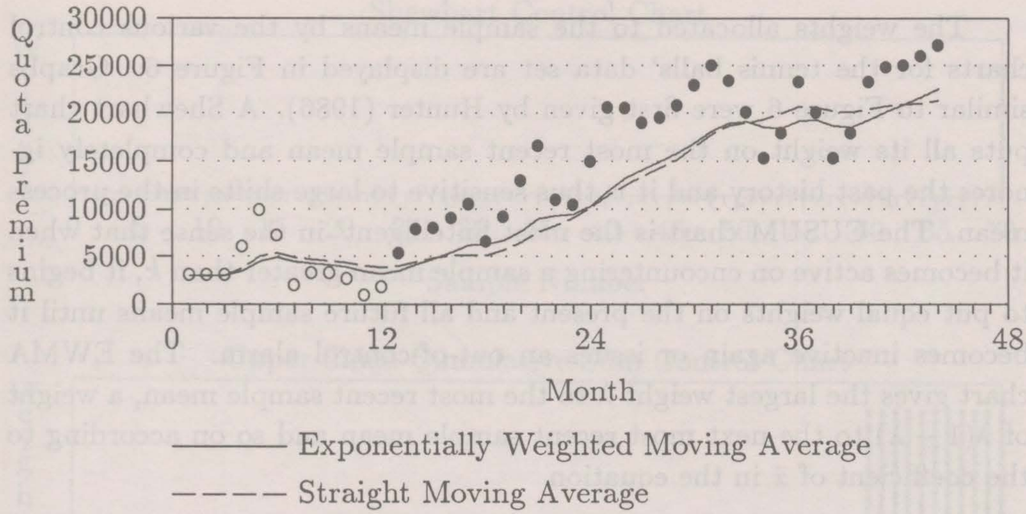

Figure 7. The COE Quota Premiums for Vehicles with Engine Capacities in the Range 1001–1600 cc from May 1990 to August 1993.

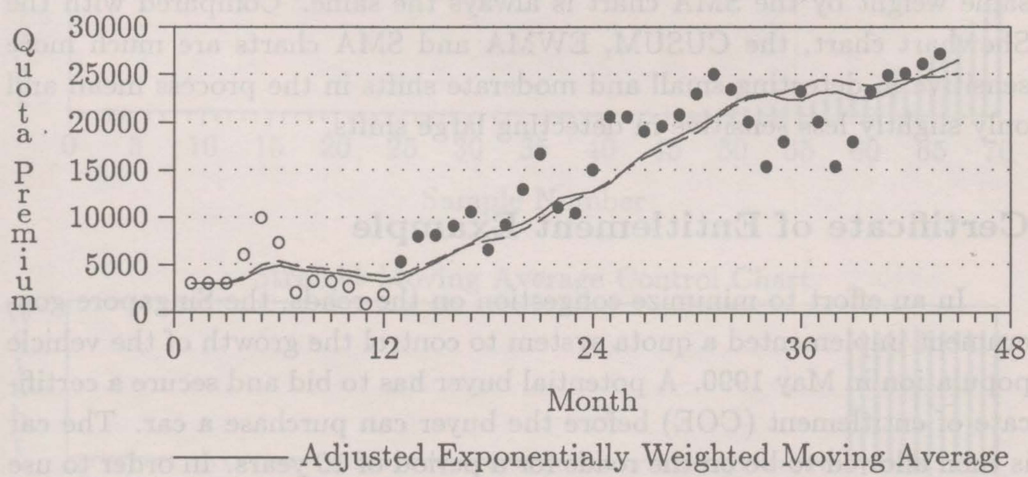

Figure 8. The COE Quota Premiums for Vehicles with Engine Capacities in the Range 1001–1600 cc from May 1990 to August 1993.

Table 1. COE Quota Premiums for Vehicles with Engine Capacities in the Range 1001–1600 from May 1990 to August 1993

3022	3022	3022	6012	9888	7220
2004	3202	3224	2649	909	1804
5258	7875	8002	9040	10520	6528
9188	12958	16602	11000	10406	14958
20542	20500	18994	19510	20741	22888
24982	23360	20010	15280	17890	23180
20010	15280	17890	23180	24800	25000
26000	27080				

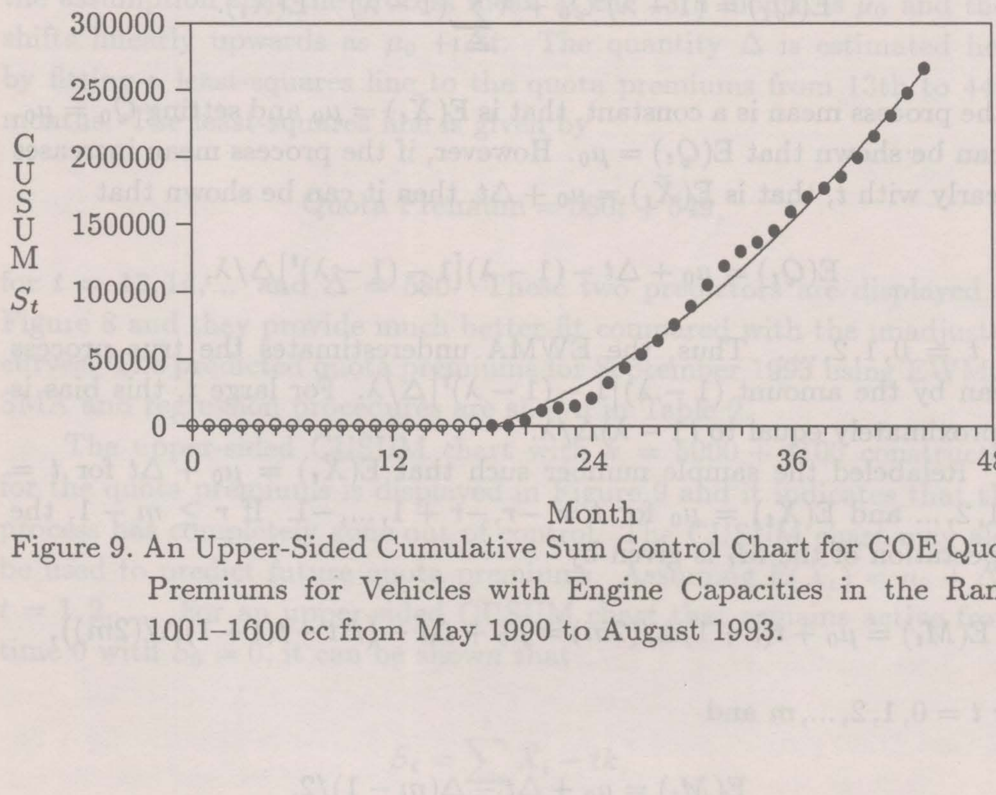

Figure 9. An Upper-Sided Cumulative Sum Control Chart for COE Quota Premiums for Vehicles with Engine Capacities in the Range 1001–1600 cc from May 1990 to August 1993.

The variances of Q_t and M_t for large t are given by $\lambda\sigma_{\bar{X}}^2/(2-\lambda)$ and $\sigma_{\bar{X}}^2/m$ respectively. Equating these two variances yields $m = (2-\lambda)/\lambda$ which provides a reasonable way of finding an EWMA chart that is 'compatible' to an SMA chart based on m. The EWMA (based on $\lambda = 0.154$) and SMA (based on $m = 12$) for the quota premiums are also displayed in Figure 7. In the beginning, the process was in control for approximately 19 months and then went completely out-of-control after that. Since an increase in the number of COE released each month has not curbed the rise in the quota premium, the main assignable cause for this out-of-control state has to be the strong economy prevailing locally. As long as this assignable cause stays, the quota premium will continue to spiral up until it stabilizes at a new value. Both the EWMA and SMA show a very similar trend of the quota premium and both consistently lag behind the quota premium from the 13th month onwards.

The expectation of Q_t is given by

$$\mathrm{E}(Q_t) = (1-\lambda)^t Q_0 + \lambda \sum_{i=1}^{t} (1-\lambda)^{t-i}\mathrm{E}(\bar{X}_i).$$

If the process mean is a constant, that is $\mathrm{E}(\bar{X}_t) = \mu_0$ and setting $Q_0 = \mu_0$, it can be shown that $\mathrm{E}(Q_t) = \mu_0$. However, if the process mean increases linearly with t, that is $\mathrm{E}(\bar{X}_t) = \mu_0 + \Delta t$, then it can be shown that

$$\mathrm{E}(Q_t) = \mu_0 + \Delta t - (1-\lambda)\left[1-(1-\lambda)^t\right]\Delta/\lambda,$$

for $t = 0, 1, 2, \ldots$. Thus, the EWMA underestimates the true process mean by the amount $(1-\lambda)\left[1-(1-\lambda)^t\right]\Delta/\lambda$. For large t, this bias is approximately equal to $(1-\lambda)\Delta/\lambda$.

Relabeled the sample number such that $\mathrm{E}(\bar{X}_t) = \mu_0 + \Delta t$ for $t = 0, 1, 2, \ldots$ and $\mathrm{E}(\bar{X}_t) = \mu_0$ for $t = -r, -r+1, \ldots, -1$. If $r \geq m-1$, the expectation of the M_t is given by

$$\mathrm{E}(M_t) = \mu_0 + t(t+1)\Delta/(2m) = \mu_0 + \Delta t - \big(\Delta t - t(t+1)\Delta/(2m)\big),$$

for $t = 0, 1, 2, \ldots, m$ and

$$\mathrm{E}(M_t) = \mu_0 + \Delta t - \Delta(m-1)/2,$$

for $t = m+1, m+2, \ldots$. For $t \geq m+1$, the bias of M_t is thus given by $\Delta(m-1)/2$ which is exactly the same as the asymptotic bias of the EWMA, $(1-\lambda)\Delta/\lambda$.

The EWMA and SMA may be adjusted such that they are unbiased estimators of a process mean that is increasing linearly:

$$\mathrm{E}(Q_t + (1-\lambda)[1-(1-\lambda)^t]\Delta/\lambda) = \mu_0 + \Delta t, \quad t = 0, 1, 2, \ldots,$$

$$\mathrm{E}(M_t + \Delta t - t(t+1)\Delta/(2m)) = \mu_0 + \Delta t, \quad t = 0, 1, \ldots, m,$$

and

$$\mathrm{E}(M_t - \Delta(m-1)/2) = \mu_0 + \Delta t, \quad t = m+1, m+2, \ldots$$

The quantity Δ can be added to these unbiased estimators so that they can be used to predict the future process mean $\mu_0 + \Delta(t+1)$. The quantity Δ which is the increase in the quota premium per month has to be estimated. Based on the plot of quota premiums in Figure 7, it is reasonable to make the assumption that the process mean at the 13th month is μ_0 and then shifts linearly upwards as $\mu_0 + \Delta t$. The quantity Δ is estimated here by fitting a least-squares line to the quota premiums from 13th to 44th months. The least-squares line is given by

$$\text{Quota Premium} = 580t + 649,$$

for $t = 13, 14, \ldots$ and $\hat{\Delta} = 580$. These two predictors are displayed in Figure 8 and they provide much better fit compared with the unadjusted curves. The predicted quota premiums for September 1993 using EWMA, SMA and regression procedures are shown in Table 2.

The upper-sided CUSUM chart with $k = 5000 + 4106$ constructed for the quota premiums is displayed in Figure 9 and it indicates that the process has completely gone out of control. The CUSUM chart may also be used to predict future quota premiums. Assuming $\mathrm{E}(\bar{X}_t) = \mu_0 + \Delta t$, $t = 1, 2, \ldots$. For an upper-sided CUSUM chart that remains active from time 0 with $S_0 = 0$, it can be shown that

$$S_t = \sum_{i=1}^{t} \bar{X}_i - tk,$$

and

$$E(S_t) = (\mu_0 - k + \Delta/2)t + \Delta t^2/2 = \alpha t + \beta t^2.$$

By fitting a quadratic curve with no intercept on the vertical axis using the values of S_t for $t = 18, 19, ..., 44$, the least-squares estimates of α and β are given by $\hat{\alpha} = 4499$ and $\hat{\beta} = 219$. Estimates of μ_0 and Δ are given by $\hat{\mu}_0 = 13385$ and $\hat{\Delta} = 439$ from which the future mean can be estimated. The predicted quota premium for September 1993 using the CUSUM procedure is also displayed in Table 2. As long as the mean continues to increase linearly with the main assignable causes remaining unchanged or varying little, these predicted values will be accurate. If the mean quota premium stabilizes, then these models must be adjusted accordingly. The predicted values in Table 2 provide idea for realistic bid amounts for very serious car buyers who must get their COEs by September 1993.

Table 2. Predicted Quota Premiums for September 1993 Obtained Using EWMA, SMA, CUSUM and Regression Procedures

EWMA	SMA	CUSUM	Regression
\$26,472	\$25,068	\$25,237	\$26,735

Conclusions

The main objective of SPC is on-line monitoring of the quality of products from a manufacturing line. The Shewhart chart is the simplest type but it is only sensitive in detecting large shifts in the process mean. The CUSUM, EWMA and SMA charts are much more sensitive in detecting small and moderate shifts in the process mean and only slightly less sensitive in detecting large shifts. A more thorough treatment of SPC can be found in Duncan (1986) who has provided tables for determining the chart parameters of a Shewhart chart. Design procedures for determining optimal chart parameters of CUSUM and EWMA charts are given by Gan (1991) and Crowder (1989) respectively. Although the SMA has the advantage that it is a quantity that is easily understood but the run length (defined as the number of samples taken until an out-of-control alarm is issued) properties of the SMA chart remains intractable. Unlike the CUSUM and EWMA charts, the SMA chart cannot be approximated as a Markov Chain.

In recognition of Deming's effort in promoting quality control in Japan, the Deming Prize was instituted in 1951 by the Union of Japanese Scientists and Engineers. The prestigous award and the benefits derived from systematic quality control effort in a company have encouraged many companies to compete for the award. The use of control charts is listed as one of the criteria for the Deming Prize. Finally, it is important to understand that in order to reduce the variance of quality of products, experimental designs would have to be vigorously conducted to seek out optimal operating conditions. Control chart procedures should then be implemented to maintain such optimal operating conditions.

Acknowledgment

I would like to thank Ms Shrlinda See of the Automobile Association of Singapore for providing me with the COE data set.

References

[1] S. V. Crowder, Design of Exponentially Weighted Moving Average Schemes, *Journal of Quality Technology,* 21 (1989), 155-162.

[2] F. F. Gan, An Optimal Design of CUSUM Quality Control Control Charts, *Journal of Quality Technology,* 23 (1991), 279-286.

[3] S. J. Hunter, The Exponentially Weighted Moving Average, *Journal of Quality Technology,* 18 (1986), 203-210.

[4] E. S. Page, Continuous Inspection Schemes, *Biometrika,* 41 (1954), 100-115.

[5] S. W. Roberts, Control Charts Based on Geometric Moving Averages, *Technometrics,* 1 (1959), 239-250.

[6] W. A. Shewhart, *Economic Control of Quality of Manufactured Products,* Van Nostrand, New York (1931).