IMPLEMENTING STSTISTICAL PROCESS CONTROL FOR CUSTOMER WAITING TIME

Xiaofeng Zhao
College of Business, University of Mary Washington
Fredericksburg, VA 22406 Email: xzhao@umw.edu

Hui Zhao
Software Engineering Institute, East China Normal University,
Shanghai, China, 200062
Email: hzhao@sei.ecnu.edu.cn

ABSTRACT

Statistical control charts, which signal an out-of-control condition when a single point falls beyond a three-sigma limit, have been the standard control charts for variables and attributes. In this paper, a statistical control chart is developed for the waiting time in the GI/G/1 queue. A significant feature of the method is that it is mathematically intractable and can be implemented in a spreadsheet format

Key Words: statistical process control, control chart, customer waiting time, queuing theory, approximation methods.

INTRODUCTION

Statistical process control (SPC) is a management philosophy that relies on straightforward statistical tools to identify and solve process problems. Closely linked to the total quality management (TQM) philosophy, SPC helps firms to improve profitability by improving process and product quality. Although initially used in manufacturing, SPC tools and methods work equally well in a service environment.

SPC methods are used extensively by organizations to enable systematic learning. Using methods developed in the 1920s by Walter Shewhart and subsequently enhanced by quality consultants William Edwards Deming and Joseph Juran, organizations are able to use a set of straightforward statistics to find out whether or not their processes conform to expectations. Furthermore, the use of SPC methods can help to identify instances of process variation that may signal a problem in the process. By identifying process variation and potential nonconformance with design expectations early in the production or service environment, managers can proactively make corrections before the process variation negatively impacts quality and customer perceptions.

Customer waiting time is a source of inherent tension between a business and its customers. Waiting experiences are typically considered as having negative influence on customers' overall satisfaction with the products and services (Kumar et al 1997, Davis and Heineke 1998), so the importance of properly managing customer waiting times is of significant interest to most organizations. In this paper we develop Statistical process control for customer waiting time.

Research on customer waiting time has traditionally been the domain of queuing theory. A queuing system is a service system in which customers arrive at random to receive service. Queues occur because of uncertainty in the environment; whenever the demand for service exceeds the ability to provide service, a queue forms. Queues differ according to various characteristics that distinguish them from one another. A major distinction classifies queues according to the number of servers and the distributions that characterize the arrival rates of customers and the service times.

In this research, a Statistical Process Control (SPC) based control chart for the distribution customer waiting time W_q is developed for a general single server queue GI/G/1, where there are no prior assumptions regarding either the inter-arrival time or the service time distributions. We first provide a mathematically tractable exact expression for the standard deviation of waiting time for Markov queues. We then apply this expression to give a two-moment approximation to the standard deviation of waiting time in a general queue. A Shewhart-like general control chart is combined with this approximation for the waiting time, to develop a statistical control chart for the waiting time in a GI/G/1system. The effectiveness of this approach is discussed and demonstrated by applying the proposed control chart to a sample of queues with various inter-arrival time and service time distributions. Thus, any assignable cause that may lead to outlying observations of W_q may indicate lack of control. Examples for such assignable causes are changes in the traffic intensity or in the system service rates, or inflation of variance in the distributions of either or both inter-arrival times or service times. The measurement requires only the mean and standard deviation or the coefficient of variation of the inter-arrival and service time distributions. It is simple enough to be implemented in manual or spreadsheet calculations, but in comparisons to Monte Carlo simulations has proven to give a good statistical control chart.

STANDARD DEVIATION FOR MARKOV QUEUES

To develop the standard deviation of waiting time, we have studied an equivalent problem of finding a mathematically tractable formula of estimating the coefficient of variation of waiting time $c_q = \sigma_q / W_q$, where W_q and σ_q are respectively the mean and standard deviation of the waiting time in queue. Compared with σ_q, c_q is more appropriate measure of variability. As W_q has a good explicate approximation, σ_q can be easily calculated if c_q is known.

Notations

To use queuing theory to describe the performance of a single queue, this research assumes the following basic parameters are known:

M: exponential (Markov) distribution

G: general distribution
λ : Arrival rate of entering customers
μ : Service rate of each server
ρ : Average utilization of a server ($\rho = \lambda / \mu$)
c_a : Coefficient of variation of inter-arrival time
c_s : Coefficient of variation of service time

The performance measures we will focus on are:

L : Average number of customers in system
L_q : Average number of customers waiting in queue
W : Average time a customer spends in system
W_q : Average time a customer spends waiting in queue
σ_q : Standard deviation of waiting time
c_q : Coefficient of variation of waiting time
$P_n(t)$ =probability of n customers in system at time t

Exact coefficient of variation of waiting time for M/M/1 Queue

We first derive exact results of c_q for Markov queue M/M/1. M/M/1 is initially considered because it is tractable and offers valuable insights into more complex and realistic situations, although not an accurate representation of most systems. It yields important intuition and serves as building blocks for more general systems. We present the exact formula for M/M/1 and show how we apply approximation methods to extend it to general GI/G/1 queue.

The key to analyzing the M/M/1 queue is the memory-less property of the exponential distribution. To begin, we require information about the inter-arrival and service times. Since both are assumed to be exponential, all we need to know are the means (because the standard deviation is equal to the mean for the exponential distribution). Beyond that, the only other information we need is how many customers are currently in the system. Because the inter-arrival and process time distributions are memory-less, the time since the last arrival and the time the current customer has been in process are irrelevant to the future behavior of the system. Because of this, the state of the system can be expressed as a single number, representing the number of costumers currently in the system. By computing the long-run probability of being in each state, we can characterize all the long-term (steady state) performance measures, including $L_q, L, W, and\ W_q$.

$$\sigma_q^2 = E[T_q{}^2] - (E[T_q])^2 = \frac{\lambda}{\mu} \cdot \frac{2}{(\mu-\lambda)^2} - \frac{\lambda^2}{\mu^2(\mu-\lambda)} == \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}$$

Hence,

$$c_q = \frac{\sigma_q}{t_q} = \sqrt{\frac{\rho(2-\rho)}{\mu^2(1-\rho^2)}} \bigg/ \frac{\rho}{(1-\rho)\mu} = \sqrt{\frac{2-\rho}{\rho}}$$

$$P\{T_q > t\} = 1 - W_q(t) = \begin{cases} \rho & (t=0) \\ \rho e^{-\mu(1-\rho)t} & (t>0) \end{cases}$$

So the probability of a customer waiting $P(T_q > 0) = \rho$

Therefore, we have

$$c_q = \sqrt{\frac{2 - P(T_q > 0)}{P(T_q > 0)}}$$

Exact coefficient of variation of waiting time for G/M/n Queue

For G/M/n queue, the distribution of waiting time is

$W_q(t) = 1 - \frac{Cr^n}{1-r} e^{-n\mu(1-r)t} \quad (t \geq 0)$ where C is a constant and r is the root of probability generating function (r < 1) (Gross and Harris 2002).

$$\sigma_q^2 = E[T_q^2] - (E[T_q])^2 = \frac{2Cr^n(1-r) - C^2 r^{2n}}{n^2 \mu^2 (1-r)^4}.$$

For G/M/n we want to verify

$\sigma_q^2 = \frac{2 - P(T_q > 0)}{P(T_q > 0)} \cdot W_q$ Equivalently, $\frac{\sigma_q^2}{W_q^{\ 2}} = \frac{2 - P(T_q > 0)}{P(T_q > 0)}$.

LHS= $\frac{\sigma_q^2}{W_q^{\ 2}} = \left(\frac{2Cr^n(1-r) - C^2 r^{2n}}{n^2 \mu^2 (1-r)^4} \right) \bigg/ \left(\frac{Cr^n}{n\mu(1-r)^2} \right)^2 = \frac{2}{Cr^n/(1-r)} - 1 = \frac{2}{P(T_q > 0)} - 1.$

RHS= $\frac{2 - P(T_q > 0)}{P(T_q > 0)} = \frac{2}{P(T_q > 0)} - 1.$

Therefore, LHS=RHS, $\frac{\sigma_q^{\ 2}}{W_q^{\ 2}} = \frac{2 - P(T_q > 0)}{P(T_q > 0)}$;

Hence, for both G/M/n and M/M/1, $c_q = \sqrt{\frac{2 - P(T_q > 0)}{P(T_q > 0)}}$

The above relationship does not depend at all on the inter-arrival time distribution or the number of server n. This implies that for G/M/n and M/M/1 queues, all of the needed information about the inter-arrival time distribution and the number of servers is contained in the probability of waiting $P(T_q > 0)$.

To estimate the standard deviation of waiting time, we know from formula (1) that the key point is to calculate the probability of waiting. From the previous discussion, the probability of waiting for M/M/1 is $P(T_q > 0) = \rho = \lambda/\mu$.

For the G/M/1 queue, we know $W_q(t) = 1 - re^{-\mu(1-r)t} \quad (t > 0)$

So $P(T_q > 0) = r$ and $W_q = \frac{r}{\mu(1-r)}$, only $r = \rho = \lambda/\mu$, $P(T_q > 0) = \rho = \lambda/\mu$

For M/G/1queue, $W_q(t) = (1-\rho)\sum_{n=0}^{\infty} \rho^n [R^n(t)]$, $P_n = (1-\rho)\rho^n$.

So $P(T_q > 0) = \rho = \lambda/\mu$

Hence, for all M/G/1, G/M/1, M/M/1 queues, $P(T_q > 0) = \rho = \lambda/\mu$.

Coefficient of variation of waiting time for the GI/G/1queue

Most practical queuing problems are the GI/G/1 system, which reflects the real world. Unfortunately, without the memory-less property of the exponential distribution to facilitate analysis, we can't compute exact performance measures for the GI/G/1 queue. However, this does not mean that we should give up on modeling queuing systems, only that we need to be concerned with finding good approximations. In contrast, an exact formula may be capable of giving the exact answers to the wrong problem or a mathematically intractable answer to the problem of interest. Consequently, approximations have been studied extensively. This research develops a new model by means of a two-moment approximation for GI/G/1 queue, which makes use of only the mean and standard deviation or coefficient of variation c of the inter-arrival and service time distributions.

The approximation is motivated by the above results of G/M/n and M/M/1 queues. By using the concept of isomorphism and similarity, Whitt (1993) conjectured that the exact formula for the distribution of waiting times of M/G/1 can be used as an approximation for the M/G/n model. Seelan and Tijms (1984) provided additional support for this approximation. This research uses the same concept of isomorphism and conjecture that for the GI/G/1queue these relationships still hold and all queuing systems have the same relationship. In other words, we conjecture that formula (1) can be used as an approximation for the GI/G/1queue since it applies to G/M/n and M/M/1. This formula has the form of the exact variance of waiting times for these queues and hence it can be easily calculated.

For the standard deviation of a general multi-server queue with infinite waiting capacity (GI/G/1), it has the properties of G/M/n and M/M/1 queues as $c_q = \sqrt{\dfrac{2 - P(T_q > 0)}{P(T_q > 0)}}$

AVERAGE WAITING TIME FOR GENERAL QUQUE

We consider the standard steady state GI/G/1queuing system with unlimited waiting room, the first come first served discipline and independent sequence of independent and identically distributed (i.i.d) inter-arrival times and service times. We assume that the general inter-arrival time and service time distributions are each partially specified by their first two moments. All descriptions of these models thus depend only on the basic 4 parameter λ , μ , c_a , c_s . To apply the approximations, the above 4 queue specifications are assumed to be known.

From definition $c_q = \sigma_q / W_q$, we know that in order to measure σ_q, we also need to calculate average waiting time W_q . We proceed by introducing an expression for the waiting time in queue

W_q and then computing the other performance measures. The approximation for W_q, which was first investigated by Kingman (1970), is given by $W_q(G/G/1)=\left(\frac{c_a^{\ 2}+c_s^{\ 2}}{2}\right)\left(\frac{\rho}{1+\rho}\right)\frac{1}{\mu}$. To develop an approximation for this situation, note that for $GI/G/1$, the approximation can be rewritten as $W_q(G/G/1)=\left(\frac{c_a^{\ 2}+c_s^{\ 2}}{2}\right)\cdot W_q(M/M/1)$ Where $W_q=\frac{\rho}{1-\rho}\cdot\frac{1}{\mu}$ is the waiting time in queue for $M/M/1$queue.

This approximation has nice properties (Hopp and Spearman 2000). It is exact for the M/M/1queue. It also happens to be exact for the GI/G/1queue, although this is not evident from our discussion here. It neatly separates into three terms: a dimensionless variability term V, a utilization term U and a time term T, as $W_q(G/G/1)=VUT$. We refer to this as Kingman's equation or as the VUT equation. From it, we see that if the V factor is less than one, then the waiting time, and hence other congestion measures, for the GI/G/1queue will be smaller than those for the M/M/1queue. This formula does not require any type of iterative algorithm to solve and is therefore easily implemented in a spreadsheet program. Because the approximation works well, this approximation is the basis of several commercially available queuing analysis packages (Hopp and Spearman 2000).This formula is used in our research when calculating mean waiting time for the GI/G/1queue.

SIMULATION AND NUMERICAL COMPARISIONS

Due to the characteristics of the input or service mechanism and the nature of the queuing discipline, or combinations of the above, for GI/G/1queue, there are no exact analytical results available. To evaluate the accuracy of our approximations, we conduct simulation experiments using the Extend simulation program. We compare the approximations with the simulation values of the standard deviations of waiting time .These numerical comparisons show that our approximation performs remarkably well.

In this simulation research, we concentrate on a single queue. In each case, we performed independent replications using 54000 minutes of simulation time and estimated 95 % confidence intervals. We characterize the queuing models by the parameters c_a, c_s and ρ. Here c_a is the coefficient of variation of an inter-arrival time; c_s is the coefficient of variation of the service time; ρ is the utilization.

The results of the approximations of the standard deviation of waiting time are compared to the results of Monte Carlo simulations. The errors are calculated by using $\%\,Error=100\left|(\text{spread sheet }\sigma-\text{simulation }\sigma)\right|/\,simulation\ \sigma$ .The comparisons are for different cases with the following combinations of parameters:
Utilization: 0.4, 0.8, 0.9, 0.95, 0.99.
Coefficient of variation of inter-arrival times: 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5.
Coefficient of variation of service times: 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5.

Distribution of inter-arrival times: Normal, Gamma.
Distribution of service times: Normal, Gamma.

Here we have simulation results corresponding to different experiments. A subset of these comparisons displays expected mean and standard deviation of cycle time in specific queuing systems. The difference and relative error analysis are displayed in a separate spreadsheet. For those cases with both $c_a, c_s \leq 1.25$, the approximations appear to be remarkably accurate. In the literature, we have seen Seelan and TIJM (1984) and Whitt (1989) used Erlang and H (hyper-exponential) distribution to represent general distribution. In our simulation experiments, we have used gamma and normal distributions to represent general distribution. Other distributions, such as hyper-exponential and Weibull, can also be tested.

Utilization=0.8		service process					
Arrival process	methods	CVs=0	CVs=0.25	CVs=0.5	CVs=0.75	CVs=1.00	CVs=1.25
CVa=0.00	approx.	0.00	0.05	0.11	0.17	0.25	0.35
	sim.	0.00	0.05	0.11	0.16	0.26	0.35
CVa=0.25	approx.	0.01	0.05	0.11	0.18	0.26	0.36
	sim.	0.00	0.05	0.10	0.17	0.24	0.33
CVa=0.5	approx.	0.04	0.07	0.12	0.19	0.28	0.38
	sim.	0.04	0.07	0.13	0.18	0.28	0.36
CVa=0.75	approx.	0.08	0.10	0.16	0.23	0.32	0.42
	sim.	0.08	0.10	0.15	0.22	0.30	0.40
CVa=1.00	approx.	0.14	0.16	0.22	0.28	0.37	0.48
	sim.	0.15	0.17	0.22	0.27	0.37	0.46
CVa=1.25	approx.	0.22	0.24	0.29	0.36	0.45	0.56
	sim.	0.24	0.24	0.29	0.36	0.47	0.54

CONTROL CAHRT

The basic theory of statistical process control was developed by Shewhart and was popularized worldwide by Deming. They both observed that repeated measurements from a process will exhibit variation—Shewhart originally worked with manufacturing processes but he and Deming quickly realized that their observation could be applied to any sort of process. If a process is stable, its variation will be predictable and can be described by one of several statistical distributions.

A basic principle underlying a of Shewhart control charts is that the distribution of the plotted statistic may be approximated by a normal distribution, with parameters that preserve the true mean and standard deviation. As evidenced by the wide-spread implementation of Shewhart control charts this practice has proven to be very useful. Numerous studies have repeatedly demonstrated the general robustness of the Shewhart control chart to deviation from normality (Schilling and Nelson 1976, Wheeler 1991, Bai and Choi 1995, and Shore 2000).

For GI/G/1 queue, customer waiting time control chart can be easily developed as we can calculate the mean W_q and standard deviation of waiting time σ_q, assuming normal distribution is still hold. Upper Control Limit (UCL), Lower Control Limit (LCL) and mean can be derived as:

$$UCL: W_q + 3\sigma_q$$
$$Mean: W_q$$
$$LCL: W_q - 3\sigma_q$$

However, if the performance of the SPC scheme is adversely affected by a violation of the normal assumption (for different general distribution G, the performance of the SPC could be either normal or non-normal), modification of the methods used is essential to guarantee the required performance. The problem of applying SPC for non-normal populations has been the focus of much research efforts (Yourstone and Zimmer 1992, Bai and Choi 1995, Shore 2000). Recently, control charts for attributes data have been developed, where the monitoring statistic may have a skewed distribution. The effectiveness of these control charts has been demonstrated for the G/G/n system. This section briefly overviews the approach developed in Shore (2000, 2006).The interested reader is advised to consult these resources for details.

Shore (2000) summarized three unique features characterize the modified charts. First, they are based on a modified normal approximation. These charts thus become a natural extension of the traditional Shewhart approach, where a normal approximation is used that preserves only the first two moments. Secondly, as with the Shewhart chart, the control limits are expressed explicitly in terms of the moments of the monitoring statistic as linear combinations of the mean, the standard deviation and the product of the standard deviation and the skewness measure. Consequently, when skewness is assumed to be zero, the proposed control chart reduces naturally to the Shehwart chart. Thirdly, unlike the Shewhart chart, which traditionally employs only standard normal quantiles (reflected in the ±3 coefficients that appear in the control limits), the proposed control charts for attributes may use those of the distribution of any symmetrically distributed standardized random variable. As a result, when the underlying distribution is skewed so as to render invalid application of the traditional Shewhart chart, the resulting control limits for the number of customers in $GI/G/1$ queue are:

$$UCL: W_q + 3\sigma_q + 1.324\sigma_q S_k - 1/2$$
$$CL: W_q$$
$$LCL: W_q - 3\sigma_q + 1.324\sigma_q S_k + 1/2$$

Shore (2006) verified that the first two moments and the skewness of the response are preserved in the control limits. By adding a simple constant term to both control limits of the conventional Shewhart chart, these limits now reflect the actual skewness of the monitoring statistic, with much improved Average Run Length (ARL) properties. Probability limits, which assume ARLs for in-control states different than those of control limits, may similarly be developed, as demonstrated in Shore (2000). For the UCLs the tail-area probability are very near to the nominal value. For the LCLs the tail-area probability are less accurate. We find out that the control limits deliver good performance for a system that was considered to be out-of-range for the classical statistical process control approach.

CONCLUSIONS

Statistical process control based monitoring of performance measures of queuing systems has eluded mainstream practices within the quality management discipline. One major reason is that there are no general theoretical formulae exist that may provide a platform to calculate control limits for a general GI/G/1 queue. To fill the gap, this research provides a mathematically exact expression for the standard deviation of waiting time for Markov queues. It then applies this expression to give a two-moment approximation to the standard deviation of waiting time for a general single-server queue with infinite waiting capacity. The measurement requires only the mean and standard deviation or the coefficient of variation of the inter-arrival and service time distributions, and the number of servers. The quality of the approximations is not the same for all cases, but in comparisons to Monte Carlo simulations has proven to give good approximations (within $\pm$ 10%). A significant feature of the approximation method is that it is mathematically intractable and can be implemented in a spreadsheet format.

REFERENCES

Albers, W., Kallenberg, W., and Otten, G. (1998) 'Accurate Test Limits Under Non-normal Measurement Error'. *Metrika*, Vol.47, No. 1, pp.1-33.

Allon, G. and Federgruen, A. (2008) 'Service competition with general queuing facilities', *Operations Research*, Vol.56, No.4, pp.827-849.

Amin, R.W., Reynolds, Jr., M. R. and Bakir, S. (1995) 'Nonparametric quality control charts based on the sign statistic', Comm. Statistics Theory Methods 24, pp. 1597-1623.

Atkinson, J B. (2008) 'Two new heuristics for the GI/G/n/0 queuing loss system with examples based on the two-phase coaxial distribution', *Journal of Operations Research Society*, May, pp.818-830.

Au, L. et al. (2009) 'Predicting overflow in an emergency department', *IMA J Management Math*, 20(1), pp. 39-49

Babad,Y.Dada, M., and Saharia, A. (1996) 'An appointment-based service center with guaranteed service', *European Journal of Operational Research*, Vol.89, Issue 2, March, pp.246-258.

Bai, D.S. and Choi, I.S. (1995) '*X* bar and *R* control charts for skewed populations', *Journal of Quality Technology* 27, pp.120-131.

Bertsimas, D. (1990) 'An analytic approach to a general class of G/G/s queuing systems', *Operations Research*, Vol. 38, No. 1. pp.139-155.

Borror, C. M., Montgomery, D. C. and Runger, G. C (1999) 'Robustness of the EWMA control Chart to non-normality', *Journal of Quality Technology*, Vol.31, pp.309-316.

Boxma, O. J. Cohen, J.W. and Huffels, M. (1979) 'Approximations of the mean waiting time in M/G/s queuing system', *Operations Research*, pp. 1115-1127

Burr, I. J. (1967) 'The effect of non-normality on constants for x-bar and R charts, *Industrial Quality Control*, 23(11), pp. 563–569.

Calzada M, Scariano S (2001) ‘The robustness of the synthetic control chart to non-normality’, *Communications in Statistics: Simulation and Computation*, Vol.30, pp 311-326.

Davis, M., Heineke, J. (1998) ‘How disconfirmation, perception and actual waiting times impact customer satisfaction’, *International Journal of Service Industry Management*, Vol. 9, No. 1, pp.64 – 73.

Gross, D. and Harris, C.M. (1998) *Fundamentals of Queuing Theory,* NY, John Wiley &Sons.

Holland, W and Griffiths, J.D. (1999) ‘A time-dependent approximation for the queue M/M (1, s)/*c.’, IMA J. Management Math, Vol.* 10(3), pp. 213-223

Hopp, W.J. and Spearman, M.L. (2000) *Factory Physics,* 2nd ed. McGraw-Hill/ Irwin.

Kingman, J.F.C.(1970) ‘Inequalities in the Theory of Queues’, *Journal of the Royal Statistical Society*. Series B. pp.32.

Kimura, T. (1986)‘A two-moment approximation for the mean waiting time in the GI/G/s queue’, *Management Science*, Vol.32, pp.751-763.

Kleinrock, L. (1976) *Queuing Systems, Vol. II: Theory*. New York, NY, John Wiley and Sons.

Larson, R. C. (1987) ‘Perspectives on queues: social justice and the psychology of queuing’, *Operations Research*, 35, 6, November-December, pp.895-904.

Lee, H. and Padmanabhan,V. (2004) ‘Information distortion in a supply chain: the bullwhip effect’, *Management Science*, Vol. 43, No. 4. pp. 546-558.

Metters, K and Pullman, M. (2003). *Successful service operations management*. 2nd edition. Thomson-South Western.

Montgomery, D.C. (2004). *Introduction to statistical quality control*, 5th ed. John Wiley, New York

Petter, J., and Peppiatt, E. (1996) ‘Managing perceptions of waiting times in service queues’, *International Journal of Service Industry Management*. Vol. 7. No.5. pp 47-61

Ryan, T.P. and Schwertman, N.C. (1997) ‘Optimal limits for attributes control charts’, *Journal of Quality Technology*, Vol. 29, pp.86–98.

Sakasegawa, H. (1977) ‘An approximate formula $Lq = \alpha\beta\ \rho\ /\ (1 - \rho)$’, *Ann. Inst. Statist. Mathematics,* Vol.29, pp.67-75.

Schilling, E.G., Nelson, P.R. (1976) ‘The Effect of Non-Normality on the Control Limits of X Charts’, *Journal of Quality Technology,* Vol. 25.

Schoonhoven, M, and Ronald J. M. M. (2010). “The X control chart under non- normality”. *Quality and Reliability Eng. Int*. 26(2): 167-176.

Seelen, L.P. and TIJMS. (1984) ‘Approximations for the conditional waiting times in GI/G/c queue’, *Operations Research Letters,* Vol. 3. pp.183- 190.

Shanthikumar, J. G. Buzacott, J. A. (1980) ‘On the approximations to the single server queue’, *International Journal of Production Research*, Vol. 18, Issue 6, pp.761-773

Shore, H. (1988) ‘Simple approximations for the GI/G/c queue, part I: the steady state Probabilities’, *Journal of the Operational Research Society*, Vol.39(3), pp.279–284.

Shore, H. (2000) ‘General control charts for attributes’, *IIE Transactions*, 32(12), pp.1149–1160.

Shore, H. (2006) ‘Control charts for the queue length in a G/G/S system’, *IIE Transactions*, Vol. 38, Dec. No. 12, pp.1117-1130.

So, K.C., Song J.S. (1998) ‘Price, delivery time guarantees and capacity selection’, *European Journal of Operational Research*, Vol.111, Issue 1. November, 28-49

Taylor, S. (1994) ‘Waiting for service: the relationship between delays and the evaluation of service’, *Journal of Marketing*, 58, April, pp. 56-69.

Whitt, W. (1989) ‘An interpolation approximation for the mean workload in GI/G/1 queue’, Operations Research, Nov/Dec89, Vol.37 6, pp.936-952.

Whitt, W., (1993) ‘Approximations for the GI/G/m queue’, *Production and Operations Management*, 2 (2), pp.114-161.

Whitt, W. (1999) ‘Predicting queuing delays’, *Management Science,* Vol. 45, No. 6.pp. 870-888.

Whitt, W.(2004) ‘A diffusion approximation for the GI/G/n/m queue’, Operations Research, Nov/Dec, Vol. 52, Issue: 6. pp. 922-941.

Yourstone, S. A. and Zimmer, W. J. (1992) ‘Non-Normality and the Design of Control Charts for Averages’, *Decision Sciences*, Vol.23,pp. 1099–1113

Woodall, W. H.; Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004) ‘Using Control Charts to Monitor Process and Product Quality Profiles’, *Journal of Quality Technology,* Vol. 36, pp. 309–320.