Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol II
IMECS 2009, March 18 - 20, 2009, Hong Kong

# Performance Evaluation of a Tukey's Control Chart in Monitoring Gamma Distribution and Short Run Processes

Chau-Chen Torng, Hong-Neng Liao, Pei-Hsi Lee, and Jhih-Cyuan Wu

*Abstract*—A Tukey's control chart applies a principle of box plot to set control limits. In this paper, we apply a Tukey's control chart to monitor gamma distribution and short run processes and evaluate its performance. Average run length (ARL) obtained by a computer simulation shows its performance. From the evaluation results, we found Tukey's chart is suitable to monitor the short run process, but its detection shift ability will become un-sensitivity when the population is skew away the normal distribution.

*Keywords* - ***Tukey's control chart; short production runs; average run length; box plot; skew probability distribution***

## I. Introduction

Statistical methods are often used in process monitoring, where the technology of control chart is one of the major tools. When an assignable cause occurs, the process will have variation. The control chart can immediately signal the process variation to operators for finding out the cause of variation. In 1942, Dr. Shewhart developed first control chart for monitoring process mean.

A use of control chart can divide setup and monitoring phases. Major tasks in setup phase are the observation collection and setting of the control limits, and this phase is also called phase I. Monitoring phases is also called phase II, and the control chart will be applied to monitor process in this phase. Shewhart individual control chart requires large numbers of observations to set the control limits in phase I [2], otherwise its performance in phase II will be increased. When a control chart is applied to monitor the data of a job-shop, most job-shops are characterized by short production runs, and many of these shops produce parts on production runs of fewer 50 units. In this short run case, we can not collect large numbers of observations in phase I, and Shewhart individual chart can not be constructed for monitoring the short run process.

Alemi [4] and Borckardt et al. [6] have presented the application of John Tukey's control chart and directly called it Tukey's control chart. The Tukey's chart has several characteristics: applied few numbers of observations to set the control limits; only single observation per period; and not affected by the outlier data. Tukey's control chart may be appropriate to monitor short run process. Statistical process control methods are often based on two assumptions: first, the sample observations are statistically independent; second, the process observations follow a normal distribution. Violation of the normality assumption can cause the increasing of the error probability when applying control charts in process monitoring. In the past studies, Kao and Ho [8] examined the performance of an R chart and found the R chart is robust to non-normality; Lin and Chou [10-13] had presented the performances of the adaptive control charts under non-normality, and Lin and Chou [10] found the VP $\overline{X}$ chart detects small shifts in the process mean faster than the standard $\overline{X}$ chart and its performances are not significantly different under normal or non-normal distributions. Borckardt et al. [6] had discussed the principle of Tukey's control chart and shown the performance in monitoring autocorrelation data, but Borckardt et al. [6] had not shown the performance to monitor short run and non-normal process.

Skew probability distributions always occur in the process monitoring of real industries, and a gamma distribution as a skew probability distribution is often selected to examine and compare the performances of control charts. In this study, we selected several gamma distributions to evaluate the performances of Tukey's control chart in monitoring short run process.

## II. Background information

Manuscript received November 30, 2008. This work was supported by the National Science Council of Taiwan, ROC, under the grant NSC 97-2221-E-224-032.

C. C. Torng is now with Graduate School of Industrial Engineering and Management, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C. (e-mail: torngcc@yuntech.edu.tw).

H. N. Liao was with Graduate School of Industrial Engineering and Management, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C. (e-mail: onizukagoods@yahoo.com.tw)

P. H. Lee is with Graduate School of Industrial Engineering and Management, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C. (Tel: 886953462766, Fax: 886-5-5312073, e-mail: g9321801@yuntech.edu.tw).

J. C. Wu was with Graduate School of Industrial Engineering and Management, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C. (e-mail: abcat904826@yahoo.com.tw)

### A. Tukey's control chart

The Tukey's control chart is an individual control chart like the Shewhart individual control chart, thus the period numbers on setting the Tukey's control chart are total observation numbers. A calculation of Box plot is the basic principle for the Tukey's control chart. Tukey's control chart setup has several steps: 1. Sorting the data; 2. calculating the 1st quartile which is denoted Q1 and 3rd quartile which is denoted Q3; and 3. calculating the control limits by using the following formulas:

$$\begin{aligned}\text{Lower control limit} &= \text{Q1} - k \times (\text{Q3} - \text{Q1}) \\ \text{Upper control limit} &= \text{Q3} - k \times (\text{Q3} - \text{Q1})\end{aligned} \quad (1)$$

The parameter $k$ determines the control limit coefficient and its default is 1.5[4, 6].

Here we quote a healthcare example in Alemi [4] to illustrate the setup of Tukey's control chart. There are 7 observed periods, each period only has a single observed data point, and these data in sequence are 0, 25, 30, 30, 32, 35, and 50. We sort these data as 0, 25, 30, 30, 32, 35, and 50, and then the Q1 and Q3 are 27.5 and 33.5, respectively. Q3-Q1=6. The lower control chart limit (LCL) is 18.5(=27.5 - 1.5 × 6), and the upper control chart limit (UCL) is 42.5(= 33.5 + 1.5 × 6). In the monitoring phase, the data in sequence is 45, 31, 20, 40, 60, 45, 60, 45, 32, 50, and 60. Figure 1 presents the Tukey's control chart in monitoring of healthcare data.

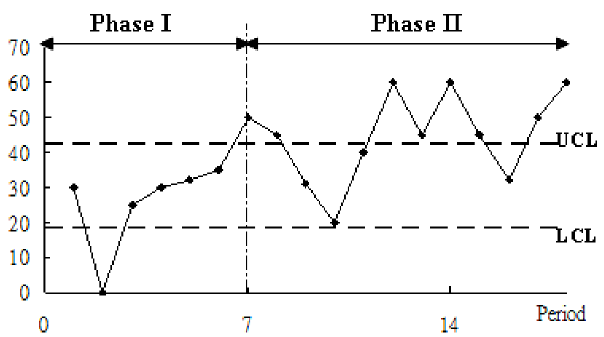


Fig. 1 Tukey's control chart for the healthcare example

In Fig. 1, the two dotted lines are control limits based on phase I procedure, and there are several out-of-control points in phase II. It is possible that the patient's health has changed, and further check-up is required.

### B. Average run length

Average run length (ARL) is the most commonly used tool for evaluating the performance of control chart. $\mu_0$ and $\sigma$ are defined as initial mean and standard deviation of the process, respectively. If the process mean stays initial mean, the process is an in-control state. When an assignable cause occurs, the process mean shifts to $\mu_1$, and $\mu_1 = \mu_0 + \delta\sigma$, where $\delta$ is called the shift size coefficient, and the process state is out-of-control. If the process state is in-control, the probability of an observation falling out the control limits can be expressed by the following formula:

$$\alpha = 1 - P\left(LCL \le x \le UCL \middle| \mu = \mu_0\right) \quad (2)$$

where the $\alpha$ is called the probability of type I error or false alarm rate. The in-control ARL denoted by $\text{ARL}_0$ can be calculated by $1/\alpha$.

Assuming a process mean shift has occurred, and then the probability of an observation falling in the control limits can be expressed by the following formula:

$$\beta = P\left(LCL \le x \le UCL \middle| \mu = \mu_1 = \mu_0 + \delta\sigma\right) \quad (3)$$

where the $\beta$ is called the probability of type II error, and the out-of-control ARL denoted by $\text{ARL}_\delta$ can be calculated by $1/(1-\beta)$.

If the error probability or ARL of the control chart is not easy to calculate, a computer simulation approach can assist to obtain the error probabilities from ideas of some studies [2, 3, 9].

## III. Simulation and Discussion

### A. Gamma distribution

In this section, we will investigate the performance of Tukey's control chart in monitoring short run and gamma distribution processes. The gamma distribution, denoted by $g(a,b)$, has the probability density function

$$f(x|a,b) = \frac{x^{a-1}\exp(-x/b)}{b^a\Gamma(a)}, \quad x > 0,\ a, b > 0 \quad (4)$$

where $x$ is a random variable, the $a$ and $b$ are respectively the shape and scale parameters, and $\Gamma(\bullet)$ is a gamma function. The mean and variance of a gamma distribution are $ab$ and $ab^2$, respectively.

In this paper, we refer to [1, 5, 7, 10], and choose $a$ = 4, 2, 1 and a fixed $b$ = 1 for the gamma distribution. Figure 2 shows the gamma distributions we have selected and their corresponding normal distributions that have the same mean and variance. When $a$ increases, the gamma distribution gets closer to a normal distribution. Through the use of these three types of gamma distributions, we can understand the effect of skewness change on the performances of control charts.

Lin and Chou [10] had used a simulation approach to calculate performances of adaptive control charts under non-normality. We adopt their approach and use a computer simulation to calculate the ARL of Tukey's chart. A computer simulation procedure is as the following:

(1) Generate $n$ observations from a gamma distribution.

(2) Calculate Q1, Q3, and $IQR$, and set the control limits with a given shift value.

(3) Generate 50,000 observations from the same

probability distribution.

(4) Record the number of observations that falls in the outside of the control limits, and then divide them by 50,000 respectively to calculate the probability of the out-of-control regions. If the shift value is equal to 0, the probability of the out-of-control regions is false alarm rate otherwise is 1-$\beta$.

(5) Calculate the run-length values.

(6) Repeat steps (1) through (5) 300,000 times to obtain the simulated run-length values.

(7) Average the 300,000 simulated run-length values to obtain ARL value.

We set respectively numbers of observations in phase I be equal to 10, 20, 30 and 10000 to simulate ARL values of Tukey's control charts. The ARL values of 10000 observations will regard as theoretical value.

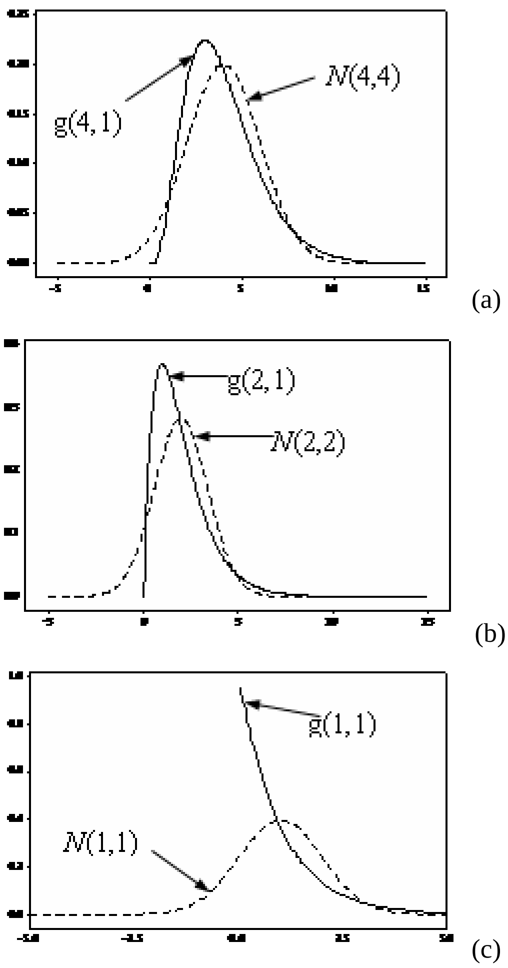


Fig. 2 The probability density function for various gamma and normal distributions: (a) g(4,1) and *N*(4,4); (b) g(2,1) and *N*(2,2); (c) g(1,1) and *N*(1,1)

## *B. Evaluation and discussion*

To compare performance of Tukey's chart in monitoring gamma distribution and short run processes, we need to set the same in-control ARL for each gamma distribution and number of observations. We adjust control limit coefficient *k* of Tukey's charts to set the in-control ARL is approximate 370.4 for each gamma distribution and number of observations. Table I shows the ARL values of Tukey's chart.

In the table I, ARL values of 10000 observations can be regarded as the theoretical values, and all out-of-control ARL values are approximate the theoretical values. That is to say if only fewer numbers of observations can be obtained in phase I, the detecting shift ability of Tukey's chart is not significant change in phase II. If the *a* decreases, the detecting shift ability in phase II will become slow. When only few numbers of observations can be obtained in phase I, and the population is far away the normal distribution, the control limit coefficient *k* of Tukey's chart has to increase for the maintenance of the in-control ARL value.

Table I.
ARL values of Tukey's chart for various numbers of observations and gamma distributions

| | Numbers of observations | $k$ | $\delta$ = 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| g(4,1) | 10 | 2.840 | 372.39 | 85.68 | 22.59 | 6.12 |
| | 20 | 2.710 | 373.20 | 83.66 | 21.01 | 5.96 |
| | 30 | 2.670 | 374.04 | 82.31 | 20.51 | 5.93 |
| | 10000 | 2.594 | 370.40 | 82.80 | 20.45 | 5.84 |
| g(2,1) | 10 | 3.300 | 372.00 | 119.22 | 34.89 | 10.98 |
| | 20 | 3.210 | 371.82 | 107.42 | 32.21 | 10.40 |
| | 30 | 3.180 | 369.52 | 106.10 | 31.72 | 10.13 |
| | 10000 | 3.138 | 370.38 | 106.56 | 31.72 | 9.95 |
| g(1,1) | 10 | 4.500 | 371.92 | 146.26 | 54.06 | 19.73 |
| | 20 | 4.320 | 373.95 | 143.94 | 53.46 | 20.33 |
| | 30 | 4.200 | 372.72 | 137.11 | 51.12 | 18.51 |
| | 10000 | 4.122 | 370.47 | 136.29 | 50.14 | 18.44 |
| N(0,1) | Theoretical value | 1.724 | 370.40 | 43.88 | 6.32 | 2.00 |

## IV. SUMMARY

This paper presented the statistical performance when the Tukey's chart is applied to monitor short run and a skew distribution process. From the evaluation results, we found Tukey's chart is still suitable to monitor the short run processes. If the population is far away the normal distribution, the detection ability of Tukey's chart will become un-sensitivity. We suggest that when the Tukey's chart is applied to monitor short run and a skew distribution process, the control limit coefficient *k* must be increased for reduction of type I error probability.

## REFERENCES

[1] C. M. Borror, D. C. Montgomery, and G. C. Runger, "Robustness of the EWMA control chart to non-normality," *J. qual. technol.*, vol. 31, pp. 309-316, 1999.

[2] C. P. Quesenberry, "The effect of sample size on estimated limits for x-bar and x control chart, " *J. qual. technol.*, vol. 25, no. 4, pp. 237-247, 1993.

[3] D. Montgomery, C. Keats, J. B. Runger, G. Messina, and C. S. William, “Integrating statistical process control and engineering process control, ” *J. qual. technol.*, vol 26, no 2, pp. 79-87, 1994.

[4] F. Alemi, “Tukey’s control chart, ” *Qual. manag. health care*, vol. 13 no. 4, pp. 216-221, 2004.

[5] G. Stoumbos, M. R. Reynolds Jr., “Robustness to non-normality and autocorrelation of individuals control charts,” *J. stat. comput. simul.*, vol. 66, pp. 145-532, 2000.

[6] J. J. Borckardt, M. R. Nash, S. Hardesty, J. Herbert, H. Cooney, and C. Pelic, “An empirical evaluation of Tukey's control chart for use in health care and quality management applications,” *Qual. manag. health care*, vol. 14, no. 2, pp. 112-115, 2005.

[7] M. E. Calzada, and S. M. Scariano, “The robustness of the synthetic control chart to non-normality,” *Commun. stat., Simul. comput.*, vol. 30, pp. 311–326, 2001.

[8] S. C. Kao, and C. Ho, “Robustness of R-chart to non-normality,” *Commun. stat., Simul. comput.*, vol. 36, pp. 1089-1098, 2007.

[9] W. Charles, L. Champ, and A. Jones, “Designing phase I -X charts with small sample sizes,” *Qual. reliab. eng. int.*, vol. 20, pp. 497-510, 2004.

[10] Y. C. Lin, and C.Y. Chou, “Non-normality and the variable parameters $\overline{X}$ control charts,” *Eur. j. oper. res.*, vol. 176, pp. 361-373, 2007.

[11] Y. C. Lin, and C.Y. Chou, “On the design of variable sample size and sampling intervals $\overline{X}$ charts under non-normality,” *Int. j. prod. econ.*, vol. 96, no. 2, pp. 249-261, 2005.

[12] Y. C. Lin, and C.Y. Chou, “On the design of variable sampling intervals $\overline{X}$ charts under non-normality,” *Int. j. ind. eng.*, vol. 12, no. 3, pp. 246-255, 2005.

[13] Y. C. Lin, and C.Y. Chou, “Robustness of the variable sample size and control limit $\overline{X}$ chart to non normality,” *Commun. stat., Theory Methods.*, vol. 34, no. 3 pp. 721-743, 2005.