

New Skewness Correction *S* Control Chart for Monitoring Dispersion of Skewed Data with Application in Healthcare

A.M.A., Atta^{1,2}, S.S. S. Yahaya¹, Z. Zain^{1,3*}, N. Ahmad¹

¹School of Quantitative Sciences (SQS), College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia.

²Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar

³Centre for Testing Measurement and Appraisal (CeTMA), Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia.

E-mail: aatta@qu.edu.qa

sharipah@uum.edu.my

zac@uum.edu.my

nurzalikhahamad@gmail.com

Article History:

Submitted: 28.01.2020

Revised: 23.03.2020

Accepted: 09.04.2020

ABSTRACT

Control chart is as a highly reputable statistical quality control tool in monitoring process stability. The classical control charts are designed on the basis of normality assumption, which is often not valid for real situations in industry; such violation of normality render the charts less accurate. This paper presents an alternative skewness correction *S* (SC-*S*) chart specifically developed to monitor process dispersion (e.g. *S* for standard deviation) for skewed distributions. Its false alarm rate (Type I error) is compared with those of various heuristics charts and standard *S* chart, while the probability of out-of-control (OOC) detection is evaluated along with the exact *S* chart. The proposed charts designed for process variables following Weibull and gamma distributions are assessed alongside the normal distribution. An extensive simulation study affirms that the SC-*S* chart performs well in regard to false alarm rate at wide range of skewness levels and sample sizes. Meanwhile, its probability of OOC detection is closer to

that of the exact *S* chart in comparison to the established charts. In conclusion, the new SC-*S* chart outperforms the established ones in monitoring process dispersion for skewed distributions.

Keywords: Probability of Out-of-Control Detection; Control Chart; Skewness Correction; False Alarm Rate, Quality Control.

Correspondence:

Z. Zain

School of Quantitative Sciences (SQS), College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok Kedah, Malaysia

Centre for Testing Measurement and Appraisal (CeTMA), Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia

E-mail: zac@uum.edu.my

DOI: [10.31838/srp.2020.4.31](https://doi.org/10.31838/srp.2020.4.31)

©Advanced Scientific Research. All rights reserved

INTRODUCTION

Control chart is among the highly reputable statistical quality control tool in monitoring a process to ensure stability in producing consistent quality output. When used effectively, it can lead to increased process efficiency resulting in quality improvement. A control chart essentially comprises three decision lines: a center line (CL) which is process average, a lower control limit (LCL) and an upper control limit (UCL) - each spaced at 3 sigma away from the CL [1]. Essentially, the key difference among control charts lies in the estimation of the standard deviation. The main purpose is to assess whether or not a process is stable by monitoring a specific characteristic of random samples from the process. Hence, an ideal control chart should produce the smallest possible Type I error in stable situation and be capable of detecting any shift in an OOC situation almost instantaneously [1].

The current control charts for monitoring process variability such as the standard deviation *S*, range *R*, and variance *S*² charts are entirely based on the assumption of normality. In cases where the data distribution is not normal, heuristic methods are adopted, such as the weighted variance (WV) *R* chart [2], weighted variance (WV) *S* chart [3], scaled weighted variance (SWV) *S* chart [4], and skewness correction (SC) *R* chart [5]. Numerous studies on heuristic methods and skewed distributions for process monitoring are published in [6], [7], [8] and [9].

The present study employs the skewness correction (SC) method in formulating the control limits of *S* chart. Theoretically, good performance of this new chart is

anticipated when the process quality characteristic data is skewed.

HEURISTIC CONTROL CHARTS

In general, the UCL and LCL are computed respectively by adding and subtracting 3 standard deviations away from the mean. What distinguishes one control chart from another is the way of estimating the standard deviation. In this section, four types of the established heuristic control charts are reviewed with their formulation for comparison purposes.

Weighted Variance *R* (WV-*R*) Chart

The essence of the weighted variance (WV) method [7], is the slicing of skewed data into two portions around its mean and using each to build a symmetric distribution. These two symmetric distributions possess the same mean, but their standard deviations differ. According to [2], the control limits are given by:

$$UCL_{WV-R} = \mu_R + 3\sigma_R \sqrt{2P_X} \quad (1)$$

and

$$LCL_{WV-R} = \mu_R - 3\sigma_R \sqrt{2(1 - P_X)} \quad (2)$$

where μ_R and σ_R are respectively the process mean (of range) and the standard deviation, and $P_X = P(X \leq \mu_X)$. It is to be noted that P_X is the probability that the quality characteristic *X* does not exceed

its mean. In cases of unknown process parameters, the limits of WV-R chart can be computed as:

$$UCL_{WV-R} = \left[1 + 3 \frac{d_3^*}{d_2^*} \sqrt{2\hat{P}_X} \right] \bar{R} \quad (3)$$

and

$$LCL_{WV-R} = \left[1 - 3 \frac{d_3^*}{d_2^*} \sqrt{2(1 - \hat{P}_X)} \right]^+ \bar{R}, \quad (4)$$

where $d_2^* = \frac{\bar{R}}{\sigma_x}$ and $d_3^* = \frac{\sigma_R}{\sigma_x}$ are the control limit constants.

Here, P_X can be estimated by counting the observations with values not larger than the mean of the sample mean \bar{X} ,

$$\hat{P}_X = \frac{\sum_{i=1}^m \sum_{j=1}^n I(\bar{X} - X_{ij})}{m \times n} \quad (5)$$

Weighted Variance S (WV-S) Chart

Similar to those of WV-R chart, except the subscript S is used to denote standard deviation instead of R for range, the WV-S chart's control limits are expressed as [3]:

$$UCL_{WV-S} = \mu_S + 3\sigma_S \sqrt{2P_X} \quad (6)$$

and

$$LCL_{WV-S} = \mu_S - 3\sigma_S \sqrt{2(1 - P_X)}. \quad (7)$$

Using the estimated parameters, the limits become:

$$UCL_{WV-S} = \left[1 + 3 \frac{\sqrt{1-(C'_4)^2}}{C'_4} \sqrt{2\hat{P}_X} \right] \bar{S} \quad (8)$$

and

$$LCL_{WV-S} = \left[1 - 3 \frac{\sqrt{1-(C'_4)^2}}{C'_4} \sqrt{2(1 - \hat{P}_X)} \right] \bar{S}. \quad (9)$$

where, $C'_4 = \frac{E(S)}{\sigma_x}$ is a constant for a given population and

$\bar{S} = \frac{\sum_{i=1}^r S_i}{r}$ is the average of the sample standard deviations, S_i estimated from r samples.

Skewness Correction (SC-R) Chart

An alternative approach using Cornish-Fisher expansion was adopted to construct the skewness correction SC-R chart, with the limits [5]:

$$UCL_{SC-R} = \mu_R + \left(3 + \frac{4/3 \alpha_3(R)}{1 + 0.2 \alpha_3^2(R)} \right) \sigma_R \quad (10)$$

and

$$LCL_{SC-R} = \mu_R + \left(-3 + \frac{4/3 \alpha_3(R)}{1 + 0.2 \alpha_3^2(R)} \right) \sigma_R, \quad (11)$$

where, $\alpha_3(R)$ is the skewness (of range). For unknown process parameters, the limits are:

$$UCL_{SC-R} = \left[1 + (3 + d_4^*) \frac{d_3^*}{d_2^*} \right] \bar{R} \quad (12)$$

and

$$LCL_{SC-R} = \left[1 + (-3 + d_4^*) \frac{d_3^*}{d_2^*} \right]^+ \bar{R}, \quad (13)$$

with the constant, $d_4^* = \frac{4/3 \hat{\alpha}_3(R)}{1 + 0.2 \hat{\alpha}_3^2(R)}$.

Scaled Weighted Variance S (SWV-S) Chart

To further enhance the performance of WV method, the scaled weighted variance (SWV) [10] method was introduced. Specifically, a scaled weighted variance S (SWV-S) chart was developed by [4], with the limits expressed as:

$$UCL_{SWV-S} = \mu_S + \Phi^{-1} \left(1 - \frac{\alpha}{4(1-P_X)} \right) \sqrt{\frac{P_X}{1-P_X}} \sigma_S \quad (14)$$

and

$$LCL_{SWV-S} = \mu_S - \Phi^{-1} \left(1 - \frac{\alpha}{4P_X} \right) \sqrt{\frac{(1-P_X)}{P_X}} \sigma_S. \quad (15)$$

where μ_S is the mean and σ_S is the standard deviation (of S) respectively, while α is the false alarm rate. For unknown process parameters, the limits are computed as:

$$UCL_{SWV-S} = \left[1 + \Phi^{-1} \left(1 - \frac{\alpha}{4(1-\hat{P}_X)} \right) \frac{\sqrt{1-(C'_4)^2}}{C'_4} \sqrt{\frac{\hat{P}_X}{(1-\hat{P}_X)}} \right] \bar{S} \quad (16)$$

and

$$LCL_{SWV-S} = \left[1 - \Phi^{-1} \left(1 - \frac{\alpha}{4\hat{P}_X} \right) \frac{\sqrt{1-(C'_4)^2}}{C'_4} \sqrt{\frac{(1-\hat{P}_X)}{\hat{P}_X}} \right]^+ \bar{S}. \quad (17)$$

SKEWNESS CORRECTION S (SC-S) CONTROL CHART

Adopting skewness correction method for the SC-R chart [5], a new SC-S chart is proposed in this study. Suppose S_i for $i = 1, 2, \dots, r$ be a sequence of sample standard deviation of r samples, then the new SC-S chart is constructed by plotting the S_i values using the calculated limits:

$$UCL_{SC-S} = \mu_S + \left(3 + \frac{4/3 \alpha_3(S)}{1 + 0.2 \alpha_3^2(S)} \right) \sigma_S \quad (18)$$

and

$$LCL_{SC-S} = \mu_S + \left(-3 + \frac{4/3 \alpha_3(S)}{1 + 0.2 \alpha_3^2(S)} \right) \sigma_S, \quad (19)$$

where, μ_S , $\alpha_3(S)$ and σ_S are the mean, skewness (of S) and standard deviation respectively. In situations of unknown process parameters, the limits are:

$$UCL_{SC-S} = \left[1 + (3 + k_4^*) \frac{\sqrt{1-(C'_4)^2}}{C'_4} \right] \bar{S} \quad (20)$$

and

$$LCL_{SC-S} = \left[1 + (-3 + k_4^*) \frac{\sqrt{1 - (c_4')^2}}{c_4'} \right] \bar{S}, \quad (21)$$

where $K_4^* = \frac{4/3 \hat{a}_3(S)}{1 + 0.2 \hat{a}_2^2(S)}$

Illustration of Application to Healthcare

The data in Table 1 consist of 85 observations on the length of hospital stay due to birth problems and delivery complications. In this study, only 17 subgroups (17 weeks) are available for sample size, $n = 5$ taken for each week.

Table 1: Data of Birth Problems and Delivery Complications

Sample, patients							
Sample number, i	P1	P2	P3	P4	P5	S_i	R_i
Week 1	0	5	3	6	2	2.387	6
Week 2	5	5	9	3	4	2.280	6
Week 3	2	3	3	8	3	2.387	6
Week 4	2	4	4	4	9	2.608	7
Week 5	2	1	3	4	11	3.962	10
Week 6	4	3	2	5	5	1.304	3
Week 7	4	3	2	5	5	1.304	3
Week 8	3	1	1	3	5	1.673	4
Week 9	3	2	3	36	5	14.687	34
Week 10	3	4	2	1	5	1.581	4
Week 11	4	3	3	3	4	0.548	1
Week 12	3	5	2	5	7	1.949	5
Week 13	3	1	1	3	2	1	2
Week 14	6	4	3	1	4	1.817	5
Week 15	3	4	3	3	4	0.548	1
Week 16	10	2	5	7	5	2.95	8
Week 17	3	3	2	4	3	0.707	2
						$\bar{S} = 2.570$	$\bar{R} = 6.2$
							94

From these data, the statistics obtained are $\hat{\mu}=4.09$, $\hat{\sigma}_x=4.034$, $\bar{S}=2.570$, $\bar{R}=6.294$ and $c_4=0.637$. It is observed that 61 observations are less than $\hat{\mu}$, thus by using equation 5, $\hat{P}_x=0.72$. The factors of skewness correction, $K_4^*=1.335$ and $d_4^*=1.367$ are computed to calculate the limits of SC-S and SC-R charts. The limits of the SC-S computed using equations (21) and (22) are $UCL_{SC-S}=16.052$ and $LCL_{SC-S}=-2.608 \Rightarrow 0$; the limits of SC-R computed using equations (12) and (13) are $UCL_{SC-R}=38.934$ and $LCL_{SC-R}=-5.911 \Rightarrow 0$; the limits of the SWV-S chart are $UCL_{SWV-S}=16.628$ and $LCL_{SWV-S}=-3.463$, $UCL_{WV-S}=10.487$ and for the standard S chart the limits are $UCL_{STD-S}=11.899$ and $LCL_{STD-S}=-6.7591 \Rightarrow 0$.

It is to be noted that the distance between UCL and CL for the SC and SWV methods are farther than those for the WV and STD methods. Figure 1 illustrates the resulting S control charts for the SC, SWV, WV and STD methods.

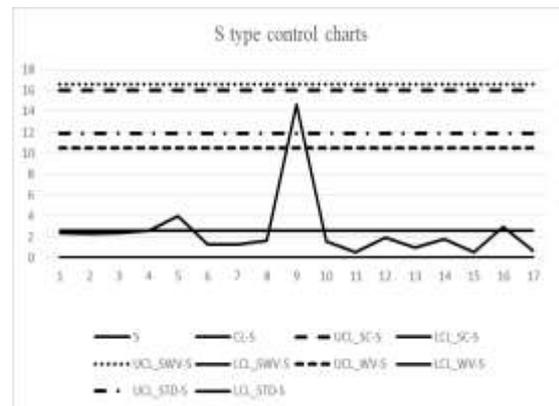


Figure 1: S type control charts

Focusing on the second dotted line from top (UCL_{SC-S}), it is evident that all the data values are indeed within the control limits of the SC-S chart, thus indicating a statistically stable process. Nevertheless, it is obvious that a point (sample 9) is beyond the UCL of the WV-S and STD charts, potentially signaling a 'false alarm'. To assess the performance of each of these charts, a simulation study is undertaken, and the findings are further discussed.

PERFORMANCE EVALUATION

The performance of the SC-S chart is evaluated along with the established heuristic, standard S and exact S charts. Monte Carlo simulation study using SAS 9.4 is

accomplished for all the charts under evaluation. The established heuristic charts considered here are the SWV-S [4], SC-R [5], WV-R [2], and WV-S [3]. Adhering to the standard, all these control charts are developed based false alarm rate $\alpha = 0.0027$, which corresponds to in-control average run length (ARL) of 370. Meanwhile, the shifted process dispersion can be defined as $\sigma_1 = \delta\sigma_X$, where $\delta \in \{1.5, 2.5, 3.0, 3.5, 4.0\}$ is the magnitude of the shift from the process standard deviation, σ_X .

For comparison purposes in this study, Weibull, gamma and normal distributions applied. The scale parameters of the skewed distributions are fixed at one ($\beta = 1$), while the skewness coefficients are set as $\alpha_3 = (0.5, 1.0, 1.5, 2.0, 2.5, 3)$ for evaluation of false alarm rate, and $\alpha_3 = 2$ is considered for probability of OOC detection. Simulated 10,000 runs, each using sample sizes of $n = (6, 9)$ are accomplished for each of the cases. Tables 2 and 3 show the results of the false alarm rate, while the subsequent Tables 4 and 5 display those of the probability of OOC detection for various simulation conditions.

Table 2: False Alarm Rates of the SC-S and Established Charts, n = 6

Distribution	α_3	SC-S	SWV-S	SC-R	WV-R	WV-S	STD-S
normal	0.0	0.0038	0.0039	0.0045	0.0045	0.0038	0.0038
β	Weibull 3.63	0.0019	0.0025	0.0013	0.0027	0.0024	0.0015
	2.23	0.5	0.0020	0.0024	0.0017	0.0032	0.0030
	1.57	1.0	0.0024	0.0034	0.0037	0.0051	0.0048
	1.21	1.5	0.0034	0.0044	0.0046	0.0072	0.0070
	0.99	2.0	0.0039	0.0051	0.0044	0.0086	0.0087
	0.86	2.5	0.0044	0.0055	0.0124	0.0098	0.0098
	0.76	3.0	0.0047	0.0056	0.0336	0.0106	0.0105
	gamma 38000	0.0	0.0031	0.0034	0.0020	0.0044	0.0035
α	15.4	0.5	0.0036	0.0041	0.0043	0.0053	0.0049
	3.91	1.0	0.0038	0.0052	0.0039	0.0070	0.0070
	1.79	1.5	0.0045	0.0053	0.0048	0.0080	0.0080
	0.98	2.0	0.0041	0.0047	0.0035	0.0083	0.0079
	0.65	2.5	0.0044	0.0046	0.0045	0.0087	0.0087
	0.44	3.0	0.0041	0.0042	0.0044	0.0090	0.0091

Table 3: False Alarm Rates of the SC-S and Established Charts, n = 9

Distribution	α_3	SC-S	SWV-S	SC-R	WV-R	WV-S	STD-S
normal	0.0	0.0034	0.0035	0.0045	0.0044	0.0034	0.0034
β	Weibull 3.63	0.0	0.0020	0.0022	0.0015	0.0026	0.0022
	2.23	0.5	0.0021	0.0023	0.0038	0.0033	0.0031
	1.57	1.0	0.0025	0.0028	0.0044	0.0051	0.0040
	1.21	1.5	0.0033	0.0037	0.0056	0.0070	0.0061
	0.99	2.0	0.0039	0.0042	0.0051	0.0082	0.0074
	0.86	2.5	0.0041	0.0044	0.0043	0.0090	0.0083
	0.76	3.0	0.0042	0.0046	0.0045	0.0096	0.0091
	gamma 38000	0.0	0.0027	0.0031	0.0027	0.0042	0.0031
α	15.4	0.5	0.0035	0.0037	0.0040	0.0054	0.0043
	3.91	1.0	0.0042	0.0043	0.0038	0.0070	0.0058
	1.79	1.5	0.0044	0.0044	0.0049	0.0078	0.0063
	0.98	2.0	0.0040	0.0042	0.0046	0.0079	0.0071
	0.65	2.5	0.0042	0.0040	0.0045	0.0084	0.0079
	0.44	3.0	0.0038	0.0038	0.0039	0.0086	0.0081

Tables 2 and 3 clearly show that the proposed SC-S chart produces the smallest false alarm rate for almost all skewness levels and sample sizes, n tested, thus indicating its superior performance to others. As anticipated, the proposed chart is reduced to the standard S chart when the data is normally distributed. Meanwhile, Tables 4 and 5

present the probability of OOC detection for data following Weibull distribution when $\alpha_3 = 2$, $\beta = 1$ and $n = 6$ and 9. The values closest to those of the exact S control chart are considered good.

Table 4: Probability of OOC Detection of Exact, SC-S, SC-R, SWV-S, WV-S, WV-R and STD-S when $\alpha_3 = 2$, $\beta = 1$, $n = 6$.

Type of Distributions	Shift, δ	Exact	SC-S	SC-R	SWV-S	WV-S	WV-R	STD-S
Weibull	1	0.9973	0.9961	0.9952	0.9950	0.9915	0.9914	0.9618
	1.5	0.9786	0.9565	0.9676	0.9494	0.9298	0.9292	0.8167
	2	0.9193	0.8628	0.8685	0.8480	0.8076	0.8058	0.6264
	2.5	0.8232	0.7363	0.7787	0.7161	0.6626	0.6614	0.4581
	3	0.7113	0.6087	0.6167	0.5833	0.5246	0.5261	0.3280
	3.5	0.5992	0.4878	0.5049	0.4659	0.4095	0.4064	0.2354
	4	0.4982	0.3913	0.3850	0.3697	0.3180	0.3177	0.1704

Table 5: Probability of OOC Detection of Exact, SC-S, SC-R, SWV-S, WV-S, WV-R and STD-S when $\alpha_3 = 2$, $\beta = 1$, $n = 9$.

Type of Distributions	Shift, δ	Exact	SC-S	SC-R	SWV-S	WV-S	WV-R	STD-S
Weibull	1	0.9973	0.9962	0.9956	0.9958	0.9926	0.9924	0.9531
	1.5	0.9726	0.9423	0.9619	0.9413	0.9176	0.9208	0.7572
	2	0.8839	0.8563	0.8644	0.8027	0.7513	0.7719	0.5049
	2.5	0.7411	0.6276	0.7203	0.6226	0.5575	0.5898	0.3091
	3	0.5828	0.4585	0.5322	0.4538	0.3899	0.4267	0.1831
	3.5	0.4400	0.3228	0.4145	0.3196	0.2647	0.2987	0.1080
	4	0.3241	0.2259	0.3133	0.2222	0.1783	0.2058	0.0647

The SC-S chart (Table 4) and SC-R chart (Table 5) produce the closest values of probability of OOC detection to those of the exact S control chart for all the values of shift magnitude, δ and sample sizes, n . In general, the proposed SC-S control chart exhibits the most desirable performance among the charts considered in the cases of false alarm and probability of OOC detection for almost all skewness levels, α_3 and sample sizes, n .

CONCLUSIONS

The skewness correction method is theoretically sound and proven as practicable. When applied to the real data, the new SC-S chart indicates the process in statistical control, while the established charts issued some potential false alarm. The finding is substantiated in the extensive simulation study which further suggests that the SC-S chart outperforms the established heuristic charts for the skewed distributions in regard to the false alarm rates (Type I error). In terms of the probability of OOC detection, the SC-S chart performs better compared to the established heuristic charts. Application to real healthcare data demonstrates its usefulness and practicability. In general, the SC-S control chart is worthy as a better or equivalent alternative for monitoring process variability when the quality characteristic data is skewed.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Ministry of Higher Education, Malaysia and Universiti Utara Malaysia for the financial support under the Fundamental Research Grant Scheme (FRGS/2/2013: S/O 12904).

REFERENCES

1. T.P. Ryan, "Statistical Method for Quality Improvement", 2nd ed. John Wiley & Sons, Inc., 2000.
2. D.S. Bai, and I.S. Choi, " \bar{X} and R control charts for skewed populations," J. Quality Technology, pp. 120-131, 1995.
3. M.B.C. Khoo, A.M.A. Atta, and C-H. Chen, "Proposed \bar{X} and S control charts for skewed distributions" Proceedings of the 2009 IEEE, IEEM, pp. 389-393, 2009.
4. A.M.A. Atta, M.H.A. Shoraim, and S.S.S. Yahaya, "A multivariate EWMA control chart for skewed population using weighted variance method," Int. Res. J. of Science & Engineering, vol. 2, no. 6, pp.191-202, 2014.
5. L.K.Chan, and H.J. Choi, "Skewness correction \bar{X} and R charts for skewed distributions," Naval Research Logistics, vol. 50, no. 6, pp. 555-573, 2003.

6. M.E. Calzada, and S.M. Scariano, "The robustness of the synthetic control chart to non-normality," *Communications in Statistics-Simulation and Computation*, vol. 30, no. 2, pp. 311-326, 2001.
7. F. Choobineh and J.L. Ballard, "Control-limits of QC charts for skewed distribution using weighted variance," *IEEE Transactions on Reliability*, Volume R-36, No. 4, pp. 473-477, 1987.
8. Chen, K. Athero-protective actions of two oral antidiabetic drugs: Suppression of inflammation and oxidative stress (2012) *Journal of Cardiovascular Disease Research*, 3 (1), pp. 3-4. DOI: 10.1016/S0975-3583(12)31002-4
9. Oyediji, A.T., Akintunde, A.A., Ajayi, E.A., Akinwusi, P.O. Coexistence of Cor triatriatum sinistrum and a prominent Eustachian valve mimicking a Cor triatriatum dextrum (2012) *Journal of Cardiovascular Disease Research*, 3 (2), pp. 170-172. DOI: 10.4103/0975-3583.95378
10. M.B.C. Khoo, Z. Wu, and A.M.A. Atta, "A synthetic control chart for monitoring the process mean of skewed populations based on the weighted variance method," *Int. J. of Reliability, Quality and Safety Engineering*, vol. 15, no. 3, pp.217-245, 2008.
11. M.D. Nichols and W.J. Padgett, "A Bootstrap Control Chart for Weibull percentiles," *Quality and Reliability Engineering International*, pp. 141-151, 2005.
12. P. Castagliola, " \bar{X} control chart for skewed populations using a scaled weighted variance method," *J. of Reliability, Quality Safety Engineering*", vol. 7, no. 3, pp. 237-252, 2000.