Available online: February 21, 2019

Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat.
Volume 68, Number 2, Pages 1377–1386 (2019)
DOI: 10.31801/cfsuasmas.530462
ISSN 1303–5991 E-ISSN 2618-6470

http://communications.science.ankara.edu.tr/index.php?series=A1

COMMUNICATIONS SERIES A1

WEIGHTED STANDARD DEVIATION METHOD FOR $\bar{X}$ AND S CONTROL CHARTS

DERYA KARAGÖZ

Abstract. A Shewhart chart based on normality assumption is not appropriate for skewed distributions since its Type-I error rate is inflated. This study presents $\bar{X}$ and S control charts for monitoring the process variability for skewed distributions. We propose Weighted Standard Deviation (WSD) $\bar{X}$ and S control charts. Standard deviation estimator is applied to monitor the process variability for estimating the process standard deviation, in the case of the WSD $\bar{X}$ and S control charts as this estimator is simple and easy to compute. Unlike the Shewhart control chart, the proposed charts provide asymmetric limits in accordance with the direction and degree of skewness to construct the upper and lower limits. The performances of the proposed charts are compared with other heuristic charts for skewed distributions by using Simulation study. The Simulation studies show that the proposed control charts have good properties for skewed distributions and large sample sizes.

1. Introduction

Shewhart $\bar{X}$ and R charts are widely applied technique to monitor the statistical process. However, the normality assumption of process population is not valid in many situations. In the case of non-normality, three methods using asymmetric control limits were proposed as alternatives to the Shewhart method. The Weighted Variance (WV) method proposed by [4], the Weighted Standard Deviations (WSD) proposed by [3] and the Skewness Correction (SC) method proposed by [2] take into consideration the skewness of the process distribution for constructing $\bar{X}$ and R charts. [5] studied on the $\bar{X}$ and R charts for skewed distributions. In these charts the standard deviation is estimated by using the sample range. Since the range method in estimating σ loses statistical efficiency as the sample size, n increases [7], the use of the Shewhart charts is not desirable when n is moderate or large. As n increases, the sample standard deviation is more efficient than the sample

Received by the editors: August 08, 2018; Accepted: September 17, 2018.
2010 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.
Key words and phrases. Weighted standard deviation, lognormal distributions, $\bar{X}$ and S control charts.

range as an estimator of standard deviation. Thus, heuristic charts whose standard deviation are estimated from the sample standard deviation must be constructed for moderate or large n. [6] proposed a weighted to compute the limits of the $\bar{X}$ and S control charts for skewed distributions.

In this paper, we propose a WSD method to construct the limits of $\bar{X}$ and S control charts for monitoring skewed process. Unlike the Shewhart $\bar{X}$ and R charts, the proposed chart, based on the standard deviation estimated from the sample standard deviation and not the sample range, provides asymmetric limits in accordance with the direction and degree of skewness by using different variances in computing the upper and lower limits. Control chart constants are obtained and simulated for lognormal distribution. The probability density function of the lognormal distribution is defined as

$$f(x|\sigma,\mu) = \frac{1}{x\sigma\sqrt{2\pi}} \exp(-\frac{(ln(x)-\mu)^2}{2\sigma^2})$$

for $x > 0$, where σ is a scale parameter and μ is a location parameter. To evaluate the performance of $\bar{X}$ and S control charts we obtain the Type I risk probabilities (p). The performance characteristics in the in-control situation can be derived as follows: The desired type I error probability p is $p = 0.0027$. By using Monte Carlo simulation, the p values of $\bar{X}$ and S control charts are compared with the Shewhart method.

This paper is organized as follows. The proposed WSD method for $\bar{X}$ and S charts are presented and the control chart constants for each method are obtained in Section 2. The next Section 3 presents the simulation study that is given to compare the p and the ARL of $\bar{X}$ and S charts with respect to different subgroup sizes for lognormal skewed distribution. The study ends up with a conclusion in Section 4.

2. Proposed $\bar{X}$ and S Charts based on WSD Method

In this section, we construct the control limits of $\bar{X}$ and S control charts for skewed populations under the WSD method. We estimate μ_x, μ_S and P_X. The μ_x is estimated using the mean of the subgroup means and μ_S is estimated using the mean of the subgroup standard deviations. The control limits are derived by assuming that the parameters of the process are unknown.

The control limits of Shewhart $\bar{X}$ chart based on the standard deviation are given by:

$$UCL_{Shew-\bar{X}_S} = \bar{\bar{X}} + 3\frac{\bar{S}}{c_4\sqrt{n}} \quad LCL_{Shew-\bar{X}_S} = \bar{\bar{X}} - 3\frac{\bar{S}}{c_4\sqrt{n}}. \tag{1}$$

The control limits of Shewhart S chart are given by:

$$UCL_{Shew-S} = \bar{S} + \frac{3\bar{S}\sqrt{(1-c_4^2)}}{c_4} \quad LCL_{Shew-S} = \bar{S} - \frac{3\bar{S}\sqrt{(1-c_4^2)}}{c_4}. \tag{2}$$

In Eq. 1 and 2 $\bar{\bar{X}}$ is the grand mean of the subgroup means and $\bar{S}$ is the mean of the subgroup standard deviations, $c_4 = \frac{E(S)}{\sigma_X}$ is the control chart constant and n is the sample size.

The control limits of weighted variance WV $\bar{X}$ chart based on the standard deviation are given by:

$$UCL_{WV-\bar{X}_S} = \bar{\bar{X}} + 3\frac{\bar{S}}{c_4\sqrt{n}}\sqrt{2\hat{P}_x} \quad LCL_{WV-\bar{X}_S} = \bar{\bar{X}} - 3\frac{\bar{S}}{c_4\sqrt{n}}\sqrt{2(1-\hat{P}_x)}. \quad (3)$$

The control limits of WV S chart are given by:

$$UCL_{WV-S} = \bar{S} + \frac{3\bar{S}\sqrt{(1-c_4^2)}}{c_4}\sqrt{2\hat{P}_X} \quad LCL_{WV-S} = \bar{S} - \frac{3\bar{S}\sqrt{(1-c_4^2)}}{c_4}\sqrt{2(1-\hat{P}_X)}. \quad (4)$$

In Eq. 3 and 4 $P_X = P(X \leq \bar{X})$ is the probability that the quality variable X will be less than or equal to its mean $\bar{X}$.

The WSD method was proposed for $\bar{X}$ and R control charts by [3]. The WSD method decomposes the skewed distribution into two parts at its mean and both parts are considered symmetric distributions which have the same mean and different standard deviation. In this method μ_x and μ_R are normally estimated using the grand mean of the subgroup means $\bar{\bar{X}}$ and the mean of the subgroup ranges $\bar{R}$, respectively. The control limits of $\bar{X}$ chart for WSD method are defined by [3] as follows:

$$\begin{aligned} UCL_{WSD-\bar{X}_R} &= \bar{\bar{X}} + 3\tfrac{\bar{R}}{d_2^{WSD}\sqrt{n}}2\hat{P}_x \\ LCL_{WSD-\bar{X}_R} &= \bar{\bar{X}} - 3\tfrac{\bar{R}}{d_2^{WSD}\sqrt{n}}2(1-\hat{P}_x) \end{aligned} \quad (5)$$

where $d_2^{WSD} = P_X d_2(2n(1-P_X)) + (1-P_X)d_2(2nP_X)$ is the control chart constant for $\bar{X}$ chart based on WSD and $d_2(n)$ is d_2 for the normal distribution corresponding to n. $P_X = P(X \leq \bar{X})$ is the probability that the quality variable X will be less than or equal to its mean $\bar{X}$. In this study, we aim to construct control limits of WSD $\bar{X}$ and S charts by using the grand mean of the subgroup means $\bar{\bar{X}}$ and the mean of the subgroup standard deviations $\bar{S}$, respectively. The control limits of WSD $\bar{X}$ chart based on the standard deviation are defined as follows:

$$UCL_{WSD-\bar{X}_S} = \bar{\bar{X}} + 3\tfrac{\bar{S}}{c_4^{WSD}\sqrt{n}}2\hat{P}_x \quad LCL_{WSD-\bar{X}_S} = \bar{\bar{X}} - 3\tfrac{\bar{S}}{c_4^{WSD}\sqrt{n}}2(1-\hat{P}_x). \quad (6)$$

where $c_4^{WSD} = \frac{E(S_n)}{\sigma_X}$ is the control chart constant for $\bar{X}$ chart based on WSD method.

The control limits of WSD S chart are defined as follows:

$$\begin{aligned} UCL_{WSD-S} &= \bar{S} + \tfrac{3\bar{S}\sqrt{(1-(c_4^{WSD})^2)}}{c_4^{WSD}}2\hat{P}_x \\ LCL_{WSD-S} &= \bar{S} - \tfrac{3\bar{S}\sqrt{(1-(c_4^{WSD})^2)}}{c_4^{WSD}}2(1-\hat{P}_x). \end{aligned} \quad (7)$$

where $c_4^{WSD} = \frac{E(S_n)}{\sigma_X}$ is the control chart constant for $\bar{X}$ chart based on WSD method Eq. 1 allows the probability to be estimated from

$$\hat{P}_X = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} \delta\left(\bar{\bar{X}} - X_{ij}\right)}{nk} \tag{8}$$

where k and n are the number of samples and the number of observations in a subgroup, and $\delta(X) = 1$ for $X \geq 0$, 0 otherwise.

$$c_4^{WSD} = \frac{E(S_n)}{\sigma} = \frac{E(S_{2nP}^{L})}{\sigma} + \frac{E(S_{2n(1-P)}^{U})}{\sigma} \tag{9}$$

where

$$\begin{aligned} E({S^2}_{2nP}^{L}) &= \int_{-\infty}^{\mu} (y-\mu)^2 f(y) dy \\ E({S^2}_{2n(1-P)}^{U}) &= \int_{\mu}^{\infty} (y-\mu)^2 f(y) dy \end{aligned}$$

and $\sigma^2 = \frac{{\sigma^2}_L^W}{(1-P)^2}$ $\sigma^2 = \frac{{\sigma^2}_U^W}{P^2}$ then c_4^{WSD} is derived as follow:

$$\begin{aligned} c_4^{WSD} &= (1-P)\frac{E(S_{2nP}^{L})}{\sigma_L^W} + P\frac{E(S_{2n(1-P)}^{U})}{\sigma_U^W} \\ &= (1-P)c_4'(2nP) + Pc_4'(2n(1-P)) \end{aligned} \tag{10}$$

where $c_4^{'} = \frac{E(S)}{\sigma_X}$ is the control chart constant. The constant c_4^{WSD} which is defined as the mean of relative range $E\left(\frac{S_n}{\sigma}\right)$ has been obtained under the non-normality assumption. This value can be computed via numerical integration once the distribution is specified [1].

An assumption of non-normality is incorporated into the constants c_4 and c_4^{WSD} to correct the control chart limits. Therefore, the constants are corrected under this conditions. The corrected constants are determined such that the expected value of the statistic divided by the constant is equal to the true value of σ.

These two constants are obtained for lognormal distribution via simulation. We obtain $E(\bar{S})$ by simulation: we generate 100.000 times k samples of size n, compute S for each instance and take the average of the values. The results for c_4 and c_4^{WSD} constants for $k = 30$ are presented in Table 1 for $n = 3 - 14$ and Table 2 for $n = 3 - 10$ based on skewness, respectively. The constant c_4^{WSD} is computed via numerical integration under lognormal distribution.

TABLE 1. The values of the constant c_4 based on skewness

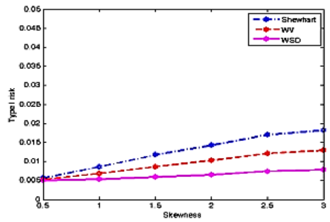
n/k_3	3	4	5	6	7	8
0.5	0.8793	0.9151	0.9340	0.9464	0.9545	0.9604
1	0.8591	0.8960	0.9170	0.9305	0.9400	0.9471
1.5	0.8349	0.8738	0.8961	0.9113	0.9222	0.9307
2.0	0.8102	0.8495	0.8736	0.8904	0.9028	0.9122
2.5	0.7731	0.8149	0.8409	0.8593	0.8732	0.8847
3.0	0.7529	0.7952	0.8220	0.8415	0.8560	0.8682
k_3/n	9	10	11	12	13	14
0.5	0.9652	0.9687	0.9718	0.9742	0.9763	0.9780
1	0.9524	0.9569	0.9608	0.9638	0.9664	0.9688
1.5	0.9371	0.9426	0.9469	0.9508	0.9539	0.9569
2.0	0.9200	0.9261	0.9315	0.9362	0.9400	0.9434
2.5	0.8937	0.9011	0.9074	0.9131	0.9177	0.9221
3.0	0.8781	0.8862	0.8931	0.8991	0.9046	0.9093

TABLE 2. The values of the constant c_4^{WSD} based on skewness

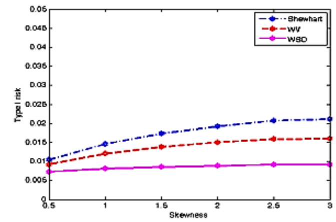
k_3/n	0.5	1.0	1.5	2.0	2.5	3.0
3	0.8739	0.8471	0.8155	0.7851	0.7431	0.7202
4	0.9127	0.8905	0.8639	0.8367	0.7963	0.7721
5	0.9327	0.9135	0.8899	0.8635	0.8250	0.8031
6	0.9454	0.9281	0.9060	0.8815	0.8461	0.8256
7	0.9539	0.9382	0.9176	0.8950	0.8619	0.8426
8	0.9600	0.9455	0.9265	0.9056	0.8740	0.8552
9	0.9647	0.9512	0.9336	0.9141	0.8836	0.8657
10	0.9685	0.9558	0.9394	0.9207	0.8919	0.8747

3. SIMULATION STUDY

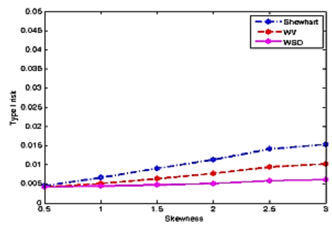
We propose to construct control limits for the $\bar{X}$ and S charts based on WSD method under lognormal distribution. The Monte Carlo simulation study is considered in this section to compare the performance of methods. The Type I risk probabilities and ARL of Shewhart, WV and WSD control charts are obtained in the simulation study 3.1.



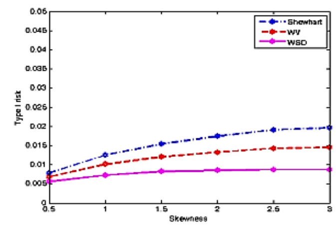
(a) $\bar{X}$ chart n=3



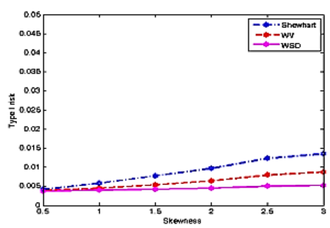
(b) S chart n=3



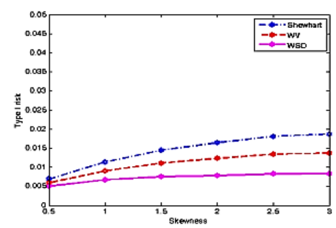
(c) $\bar{X}$ chart n=5



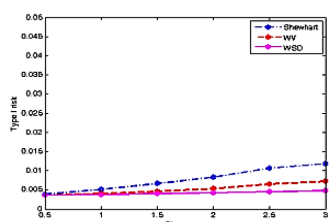
(d) S chart n=5



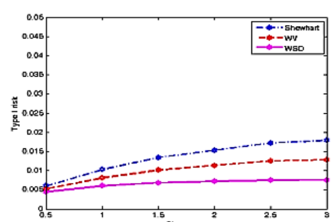
(e) $\bar{X}$ chart n=7



(f) S chart n=7



(g) $\bar{X}$ chart n=10



(h) S chart n=10

FIGURE 1. Type I risks of $\bar{X}$ and S charts under lognormal distribution, n=3,5,7,10

3.1. **Performance of $\bar{X}$ and S control charts.** When the parameters of the process are unknown, control charts can be applied in a two-phase procedure. In Phase I, control charts are used to define the in-control state of the process and to assess process stability for ensuring that the reference sample is representative of the process. The parameters of the process are estimated from Phase I sample and control limits are estimated for using in Phase II. In Phase II, samples from the process are prospectively monitored for departures from the in-control state. The Type I risk indicates the probability of a subgroup $\bar{X}$ falling outside the ± 3 sigma control limits. When the process is in-control, the Type I risks are 0.27%. In the statistical process control, the desired average run length (ARL) value of 370 indicates that the control limits are chosen to provide a p of 0.0027. If p is to be as low as possible and ARL is to be as high as possible, it means that the process in control. Under non-normality, when skewness increases, ARL decreases and therefore p increases.

To evaluate the control chart performance we obtain p and ARL for moderate sample size (30 subgroups of 3-10) for skewed distribution such as lognormal. The simulation study consists of two Phases. The steps of each Phase are described as following.

TABLE 3. The values p of the $\bar{X}$ chart based on the skewness

n=3						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0056	0.0086	0.0117	0.0142	0.0170	0.0182
WV	0.0052	0.0068	0.0086	0.0103	0.0121	0.0129
WSD	**0.0050**	**0.0053**	**0.0059**	**0.0065**	**0.0074**	**0.0078**
n=5						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0045	0.0066	0.0090	0.0113	0.0141	0.0153
WV	0.0042	0.0051	0.0063	0.0077	0.0094	0.0102
WSD	**0.0042**	**0.0044**	**0.0047**	**0.0051**	**0.0058**	**0.0061**
n=7						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0041	0.0058	0.0077	0.0097	0.0123	0.0135
WV	0.0038	0.0045	0.0054	0.0064	0.0079	0.0087
WSD	**0.0038**	**0.0040**	**0.0042**	**0.0045**	**0.0050**	**0.0052**
n=10						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0038	0.0051	0.0066	0.0083	0.0106	0.0118
WV	0.0036	0.0040	0.0046	0.0053	0.0065	0.0072
WSD	**0.0036**	**0.0038**	**0.0040**	**0.0042**	**0.0045**	**0.0047**

Phase I:

1.a. Generate n *i.i.d.* lognormal$(0, \sigma)$ varieties for $n = 3, 5, 7, 10$.
1.b. Repeat step 1.a 30 times $(k = 30)$.
1.c. The control limits of $\bar{X}$ and S charts for the Shewhart, WV and WSD methods are obtained.

TABLE 4. The values p of the S chart based on the skewness

n=3						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0104	0.0146	0.0173	0.0192	0.0207	0.0211
WV	0.0092	0.0120	0.0138	0.0150	0.0158	0.0160
WSD	**0.0073**	**0.0081**	**0.0085**	**0.0088**	**0.0092**	**0.0092**
n=5						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0078	0.0125	0.0154	0.0174	0.0191	0.0196
WV	0.0068	0.0101	0.0120	0.0133	0.0143	0.0145
WSD	**0.0056**	**0.0073**	**0.0082**	**0.0085**	**0.0087**	**0.0087**
n=7						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0069	0.0114	0.0144	0.0164	0.0181	0.0187
WV	0.0059	0.0090	0.0111	0.0123	0.0134	0.0137
WSD	**0.0050**	**0.0067**	**0.0075**	**0.0078**	**0.0082**	**0.0083**
n=10						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	0.0060	0.0103	0.0134	0.0153	0.0172	0.0179
WV	0.0052	0.0081	0.0101	0.0114	0.0125	0.0128
WSD	**0.0044**	**0.0060**	**0.0068**	**0.0072**	**0.0075**	**0.0076**

Phase II:

2.a. Generate n *i.i.d.* lognormal $(0, \sigma)$ varieties using the procedure of step 1.a.
2.b. Repeat step 2.a 100 times $(k = 100)$.
2.c. Compute the sample statistics for $\bar{X}$ and S charts for the Shewhart, WV and WSD methods.
2.d. Record whether or not the sample statistics calculated in step 2.c are within the control limits of step 1.c. for all methods.
2.e. Repeat steps 1.a through 2.d, 100.000 times and obtain p and ARL values for each method.

The results of the p and ARL values for the $\bar{X}$ and S control charts for the lognormal distribution are presented in Tables **??**,**??**, **??** and **??** present. The results of this study can be sum up as following:

Table 5. The values ARL of the $\bar{X}$ chart based on the skewness

n=3						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	178.5045	115.7957	85.6458	70.1873	58.7385	54.9734
WV	192.8380	147.0869	115.9313	97.1676	82.4266	77.4060
WSD	198.9495	188.6081	170.1664	152.8328	135.7773	128.6521
n=5						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	221.8131	150.9161	110.7285	88.4205	71.0712	65.2614
WV	240.5407	195.3659	157.4803	130.2423	106.3309	97.7641
WSD	240.7840	229.4367	213.0379	196.1400	173.5087	164.2818
n=7						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	245.0680	172.8429	129.2040	102.6937	81.4127	73.9568
WV	263.8174	223.1545	186.0569	156.0208	126.1639	115.2990
WSD	261.8144	248.7686	238.1860	222.6031	201.0859	190.7996
n=10						
Method/k_3	0.5	1.0	1.5	2.0	2.5	3.0
Shewhart	261.2467	197.6402	151.6024	120.6258	94.6692	84.8464
WV	278.3964	251.4079	219.5438	186.9823	153.5862	138.5233
WSD	274.4764	262.2469	250.2127	237.9989	222.5140	213.0879

- For the symmetric distributions, the $\bar{X}$ and S control charts give the similar results for all methods, when the WSD $\bar{X}$ chart has the best performance.
- When the skewness increases, the p values of the Shewhart and WV methods increase too much and are higher than the WSD method.
- The WSD $\bar{X}$ control chart gives better results than S chart.
- The Shewhart method does not work well in the case of skewness. So it is not recommended to use this method.
- For the large sample size n=10; the WSD $\bar{X}$ and S control charts has the smallest p values when skewness increases.

Type I risks of $\bar{X}$ and S charts based on Shewhart, WV and WSD methods under lognormal distribution, n=3,5,7,10 are given in Figure 1.

4. Conclusion

This article has presented the WSD $\bar{X}$ and S control chart for monitoring the process variability and obtained the control chart constants for computing the control limits and the central line. Standard deviation estimator is applied to monitor the process variability for estimating the process standard deviation, in the case of the WSD $\bar{X}$ and S control charts as this estimator is simple and easy to compute.

Unlike the Shewhart $\bar{X}$ and S control charts, the proposed charts provide asymmetric limits in accordance with the direction and degree of skewness to construct the upper and lower limits. The simulation studies show that the proposed WSD $\bar{X}$ and S method has good performance for heavy tailed distribution. For large sample sizes, where it leads to better performance than the Shewhart and WV methods, especially when the skewness increases. Finally, alternative to the Shewhart and WV control charts based on standard deviation, WSD $\bar{X}$ and S control charts can be applied to monitor the process variability in the case of skewness, especially in the case of large samples.

References

[1] Bai, D.S., and Choi, I.S., 1995, $\bar{X}$ and R Control Charts For Skewed Populations, *Journal Of Quality Technology*, 27:120-131.

[2] L.K. Chan, and H.J. Cui, 2003, *Skewness Correction $\bar{X}$ and R Charts for Skewed Distributions*, Naval Research Logistics, 50, 1-19.

[3] Chang, Y.S., and Bai, D.S.,2001, Control Charts for Positively Skewed Populations with Weighted Standard Deviations, *Quality and Reliability Engineering International*, 17: 397-406.

[4] Choobineh, F., and Ballard, J.L., 1987, Control-Limits of QC Charts For Skewed Distributions Using Weighted Variance, *IEEE Transactions on Reliability*, 473-477.

[5] Karagöz, D., and Hamurkaroğlu C., 2012, Control Charts for Skewed Distributions: Weibull, Gamma, and Lognormal, *Metodoloski zvezki - Advances in Methodology and Statistics*, 9.2: 95-106.

[6] Koo, M.B., Atta, A.M.A. and Chen C.H., 2009, *Proposed $\bar{X}$ and S Control Charts for Skewed Distributions*, Proceedings of the 2009 IEEE IEEM.

[7] Montgomery, D.C., 1997, *Introduction to Statistical Quality Control.* John Wiley&Sons. Inc., USA.

Current address: Derya Karagöz : Department of Statistics, Hacettepe University, Ankara, 06800 Turkey.

E-mail address: deryacal@hacettepe.edu.tr

ORCID Address: http://orcid.org/0000-0002-8283-8573