Hacettepe Journal of Mathematics and Statistics
Volume 47 (1) (2018), 223 – 242

Robust $\bar{X}$ control chart for monitoring the skewed and contaminated process

Derya Karagöz*†

Abstract

In this paper, we propose the modified Shewhart, the modified weighted variance and the modified skewness correction methods by using trimmed mean and interquartile range estimators to construct the control limits of robust $\bar{X}$ control chart for monitoring the skewed and contaminated process. A comparison between the performances of the $\bar{X}$ chart for monitoring the process mean based on these three modified models is made in terms of the Type I risk probabilities and the average run length values for the various levels of skewness as well as different contamination models.

Keywords: Skewed process, modified weighted variance method, modified skewness correction method.

Received : 15.07.2015 *Accepted* : 17.02.2016 *Doi* : 10.15672/HJMS.201611815892

1. Introduction

Control charts are among the most commonly used and powerful tools in statistical process control (1) to learn about a process, (2) to monitor a process for control and (3) to improve it sequentially. They are now widely accepted and applied in industry. The conventional Shewhart $\bar{X}$ and R control charts are based on the assumption that the distribution of the quality characteristic (also called process distribution) is normal or approximately normal. However, in many situations the normality assumption of process population is not valid. One case is that the distribution is skewed [3], [6] and [13]. For instance, the distributions of measurements in chemical processes, semiconductor processes, cutting tool wear processes and observations on lifetimes in accelerated life test samples are often skewed[10].

The $\bar{X}$ and R control charts are widely applied technique for monitoring the process. Control charts can be applied in a two-stages when the parameters of a quality characteristic of the process are unknown. In Phase I, control charts are used to study

*Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey, Email: deryacal@hacettepe.edu.tr
†Corresponding Author.

a historical data set and determine the samples that are out of control. Based on the resulting reference sample, the process parameters are estimated and control limits are calculated for Phase II. Control charts are used for real-time process monitoring in Phase II [15].

To deal with non normal underlying distributions, three methods using asymmetric control limits were proposed as alternatives to the Shewhart method. The Weighted Variance (WV) method proposed by [6], the Weighted Standard Deviations (WSD) proposed by [7] and the Skewness Correction (SC) method proposed by [5] take into consideration the skewness of the process distribution for constructing $\bar{X}$ and R charts. Moreover, [4] proposed a synthetic Scaled WV (SWV) control chart for monitoring the mean of skewed populations. The Scaled Weighted Variance method has been proven to be more efficient than the WV one [4]. Some of the other works on control charts for contaminated populations are made by: [20] considered robust estimators to obtain the control limits for $\bar{X}$ charts. Via simulation, they studied the seven different estimators of σ, one of which was based on absolute deviations from the mean, and three others were based on deviations from the median. [14] studied design schemes for the $\bar{X}$ control chart under non-normality. Different estimators of standard deviation were considered and the effect of the estimator on the performance of the control chart under non-normality was investigated. [1] presented a simple approach to robust estimation of the process standard deviation σ based on a very robust scale estimator, namely, the median absolute deviation from the sample median (MAD). The proposed method provides an alternative to the Shewhart S control chart.

[17] considered the interquartile range and the 25% trimmed mean of the interquartile ranges. [17] gave the practical details for the construction of the charts based on these estimators. [15] and [16] studied several estimators used to construct the standard deviation Phase II control chart. They found that Tatum's estimator is robust against diffuse disturbances but less robust against shifts in the process standard deviation in Phase I.

[16] studied alternative standard deviation estimators that serve as a basis to determine the $\bar{X}$ control chart limits used for real-time process monitoring (Phase II). Several robust estimation methods were considered. In addition, they proposed a new estimation method based on a Phase I analysis, that is, the use of a control chart to identify disturbances in a data set retrospectively. The method constructs a Phase I control chart derived from the trimmed mean of the sample interquartile ranges, which is used to identify out-of-control data.

In this paper, we propose the modified Shewhart (MS), the modified weighted variance (MWV) and the modified skewness correction (MSC) methods to construct the limits of $\bar{X}$ control chart for monitoring skewed and contaminated process . One contribution of this paper is to replace the overall mean by a trimmed mean and the estimator of the standard deviation based on the ranges by the interquartile ranges. For this new situations coefficients for establishing the control limits are given. Control chart constants are simulated for three skewed distributions. Another contribution is to correct the control limits for skewness. Again two alternatives are considered: one variant based on the traditional choices; the other based on the robust choices. We study the effect of the estimators on control chart performance under non-normality for moderate sample sizes. To evaluate the performance of control chart we obtain the Type I risk probabilities (p) and the average run lengths (ARL) of these control charts. The performance characteristics in the in-control situation can be derived as follows: The desired type I error probability p is $p = 0.0027$ and $ARL = 370.4$. By using Monte Carlo simulation, the p and the ARL of $\bar{X}$ control charts are compared with the classic estimators for the Shewhart, WV and SC methods and the robust estimators for the MS, MWV and MSC methods.

This paper is organized as follows. The estimators and modified methods are presented in Section 2. The effect of outliers on the accuracy of the conventional and robust estimators are evaluated by root mean square errors via simulation in Section 3.1. The control chart constants for each method are obtained in Section 3.2. The next Section 3.3 presents the simulation study that is given to compare the p and the ARL of $\bar{X}$ control chart with respect to different subgroup sizes for Weibull, gamma and lognormal skewed distributions. The results are presented in Section 4 . The study ends up with a conclusion in Section 5.

2. Skewed distributions, estimators and modified methods

In this section, the modified methods under skewed distributions, given in 2.1, using classic and robust estimators, given in Section 2.2. The proposed methods to construct the $\bar{X}$ control chart are explained in details in Section 2.3.

2.1. Skewed distributions. The Weibull, gamma and lognormal distributions are chosen since they can represent a wide variety of shapes from nearly symmetric to highly skewed.

- The probability density function of the Weibull distribution is defined as

$$f(x|\beta,\lambda) = \beta\lambda^{\beta}x^{\beta-1}\exp(-x\lambda)^{\beta}$$

for $x > 0$, where β is shape parameter and λ is a scale parameter.

- The probability density function of the gamma distribution is defined as

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}\exp(-\frac{x}{\beta})$$

for $x > 0$, where α is a shape parameter and β is a scale parameter.

- 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.

2.2. Classic and robust estimators. The main advantage of the classic estimator, is that, it can be regarded as truly representative of the data, since all data values are taken into account in its calculation, while the main disadvantage, is that, it is non-robust to slight deviations from normality and can be easily influenced by outliers. The breakdown point of the sample mean for a sample of size n is merely 1/n, that is, it can be destroyed by even a single outlier. According to Tukey, using the trimmean instead of the mean or the median gives a more useful assessment of location or centering ([15]). Robust statistical methods, of which the trimmed mean is a simple example, seek to outperform classical statistical methods in the presence of outliers, or, more generally, when underlying parametric assumptions are not quite correct.

In this paper, we will restrict attention to estimator that have an explicit formula, being easily computable, needs little computation time and have robustness properties that are high breakdown point and a bounded influence function.

In practice, the process parameters μ and σ are usually unknown. They must therefore be estimated from samples taken when the process is assumed to be in control (i.e., in Phase I). The resulting estimates are used to monitor the location of the process in Phase II. We define $\hat{\mu}$ and $\hat{\sigma}$ as unbiased estimates of μ and σ, respectively, based on k. Phase I samples of size n, which are denoted by X_{ij}, $i = 1, 2, ..., k$. The first location estimator that we consider is the mean of the sample means, $\bar{\bar{X}} = \frac{1}{k}\sum_{i=1}^{k}\bar{X}_i =$

$\frac{1}{k}\sum_{i=1}^{k}(\frac{1}{n}\sum_{j=1}^{n}X_{ij})$ where $i = 1, 2, ..., k$ and $j = 1, 2, ..., n$. We assume that X_{ij} are independent and that their distribution is skewed. This is the most efficient estimator for normally distributed data, but it is well known that it is not robust against outliers. Therefore, we also consider the mean of the sample trimmed means. Let $X_{i1}...Xin$ represent observations on a variable from ith random sample. We start by ordering the values of X_{ij} from lowest to highest for each sample, and determining the desired amount of trimming,$0 = \alpha < 0,5$ the mean is then calculated for all observations of each samples except the g smallest and largest observations $g = \frac{n\alpha}{2}$, where $\frac{n\alpha}{2}$ is rounded to the nearest integer. The formula for the trimmed mean can be written as

$$T\bar{M}_{\alpha} = \frac{1}{k}\sum_{i=1}^{k} T\bar{M}_{vi} \tag{2.1}$$

where $TM_{(vi)}$ denotes the vth ordered value of the sample trimmed means defined by

$$T\bar{M}_{vi} = \frac{1}{n - 2\lceil n\alpha\rceil}\left[\sum_{j=\lceil n\alpha\rceil+1}^{n-\lceil n\alpha\rceil} X_{(ij)}\right] \tag{2.2}$$

where α denotes the percentage of samples to be trimmed, $\lceil n\alpha\rceil$ denotes the ceiling function, i.e., the smallest integer not less than $n\alpha$. We consider the 20% trimmed mean, which trims the three smallest and the three largest sample trimmed means when k=30.

The higher the breakdown point (*bdp*) of an estimator, the more robust it is. The *bdp* cannot exceed 50% because if more than half of the observations are contaminated, it is not possible to distinguish between the underlying distribution and the contaminating distribution. Therefore, the maximum *bdp* is 0.5 and there are estimators which achieve such a *bdp*. A relatively robust measure of center is the trimmed mean, which reduces the impact of outliers or heavy tails by removing the observations at the tails of the distribution. The *bdp* of the trimmed mean is determined by the amount of trimming, and thus is *bdp* $= \alpha$. For more details, see [9] and [11].

The amount of trimming also determines the influence function. While the influence function of the mean is unbounded, the influence function for the trimmed mean is bounded. Its influence function can be written as

$$IF_{T_\alpha}(X) = \begin{cases} \frac{X_\alpha - \hat{\mu}_t}{1-2\alpha} & for X < X_\alpha \\ \frac{X_\alpha - \hat{\mu}_t}{1-2\alpha} & for X_\alpha < X < X_{1-\alpha} \\ \frac{X_{(1-\alpha)} - \hat{\mu}_t}{1-2\alpha} & for X > X_{1-\alpha} \end{cases} \tag{2.3}$$

where $\hat{\mu}_t$ is the trimmed mean (see [19]). The relative efficiency of the trimmed mean depends on the distribution. If the distribution is normal and too much trimming is done, precision will be reduced because it results in greater spread relative to the smaller n, thus increasing the estimate of the 12 spread of its sampling distribution. On the other hand, if the distribution has heavy tails and extreme outliers, trimming can result in improved efficiency because the variance of X and hence the estimated variance of the sampling distribution of its mean is decreased.

The first scale estimator is the mean of the sample range
$\bar{R} = \frac{1}{k}\sum_{i=1}^{k} R_i$ where R_i is the range of the ith sample. An unbiased estimator of σ is $\bar{R}/d_2(n)$. We also consider the mean of the sample interquartile ranges since the mean of the sample range not robust against outliers. The mean of the sample interquartile

ranges ($IQRs$) is defined by

$$(2.4) \qquad I\bar{Q}R = \frac{1}{k}\sum_{i=1}^{k} IQR_i$$

where IQR_i is the interquartile range of sample i: $IQR_i = Q_{75,i} - Q_{25,i}$; $Q_{r,i}$ is the rth percentile of the values in sample i. Q_{75} and Q_{25} are found by solving the following integrals

$$(2.5) \qquad Q_{75} = \int_{-\infty}^{Q_3} f(x)dx \quad and \quad Q_{25} = \int_{-\infty}^{Q_1} f(x)dx$$

The function $f(x)$ is continuous over the support of X that satisfies the two properties, $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$. The IQR for Weibull, gamma and lognormal distributions are obtained by taking difference between the quantiles in 2.5 after some integration calculations by [18] and are given respectively

$$\begin{aligned}
IQR_{weib} &= \Bigg[-1/\beta \ln(0.25)\Bigg]^{1/\lambda} - \Bigg[-1/\beta \ln(0.75)\Bigg]^{1/\lambda} = 1/\beta^{1/\lambda}\ln(4)^{1/\lambda} - \ln(4/3)^{1/\lambda} \\
IQR_{gamma} &= \sum_{x=0}^{\alpha-1} \frac{Q_1/\beta^x \exp(-Q_1/\beta)}{x!} - \sum_{x=0}^{\alpha-1} \frac{Q_3/\beta^x \exp(-Q_3/\beta)}{x!} \\
IQR_{logn} &= \exp(\mu)\Bigg[\exp(0.6745\sigma) - \exp(-0.6745\sigma)\Bigg]
\end{aligned}$$

where $\sigma > 0$ by [18].

The IQR is a set of bounded influence measures of scale that can have a very high breakdown point. The difference between the .25 and .75 quantiles produces the IQR, which, with a $bdp = 0.25$, is the most robust and thus most commonly used of the quantile ranges [19]. The influence function for the IQR is given by the influence function at the third quartile minus the influence function at the first quartile

$$(2.6) \qquad IF_{IQR}(X) = \begin{cases} \frac{1}{f(x_{0.25})} - C & if X < X_{0.25} or X > X_{0.75} \\ -C & if X_{0.25} \leq X \leq X_{0.75} \end{cases}$$

where $C = q(\frac{1}{f(x_{0.25})} + \frac{1}{f(x_{0.75})})$, here q is the quantile of the distribution. IQR has the high bdp and bounded influence function which are are desirable properties.

Theorem 1. The probability distribution function for interquartile range is

$$\begin{aligned}
f_Y(y) &= \int_a^{b-y} f_{(Y,Z)}(y,z)dz \\
&= \int_a^{b-y} \frac{n!}{(\frac{n}{4}-1)!(\frac{3n}{4}-\frac{n}{4}-1)!(n-\frac{3n}{4})!} \\
(2.7) \qquad &* \quad (F(z))^{\frac{n}{4}-1}(F(y+z)-F(z))^{\frac{3n}{4}-\frac{n}{4}-1}(1-(F(y+z))^{n-\frac{3n}{4}} f(z)f(y+z)dz.
\end{aligned}$$

Proof 1. Given a random sample, $X_1, ...X_n$, the sample order statistics $X_{(1)} < .. < X_{(m)} < ...X_{(k)} < ... < X_{(n)}$ are the sample values placed in ascending order,

$$\begin{aligned} X_{(1)} &= \min_{1\le i\le n} (X_i), \\ X_{(m)} &= \text{the first quantile} \quad X_{\frac{n}{4}}, \\ X_{(k)} &= \text{the third quantile} \quad X_{\frac{3n}{4}}, \\ X_{(n)} &= \max_{1\le i\le n} (X_i). \end{aligned}$$

The event $A = \left\{ X_m \le x_1, X_k \le X_2 \right\}$ is a union of some disjoint events

$$a_{m,k,n-m-k} = \left\{ \begin{array}{l} \text{m elements of the sample fall into } (-\infty, x_1], \\ k\text{elements fall into interval } (x_1, x_2], \text{and} \\ (n-m-k)\text{elements lie to the right of} x_2 \end{array} \right\}$$

To construct A one has to take all $a_{m,k,n-m-k}$ such that $r \le m \le n, j \ge 0$ and $s \le m + n \le n$ [2].

The joint distribution of two order statistics X_m and X_k is given by [2] as following:

$$\begin{aligned} f_{X_m,Y_k}(x_1, x_2) &= \frac{n!}{(m-1)!(m-k-1)!(n-k)!} \\ &* \quad (F(x_1))^{m-1}(F(x_2) - F(x_1))^{m-k-1}(1 - (F(x_2))^{n-k} f(x_1)f(x_2). \end{aligned} \tag{2.8}$$

Hence the distribution function of two order statistics X_m and X_k is given by [2] as following:

$$F_{X_m,X_k}(x_1, x_2) = \sum_{m=r}^{n} \sum_{k=\max\{0,s-m\}}^{n-m} P\{A_{m,k,n-m-k}\} \tag{2.9}$$

where $P\{A_{m,k,n-m-k}\} = \frac{n!}{m!k!(n-m-k)!}(F(x_1))^m(F(x_2) - F(x_1))^k(1 - F(x_2))^{n-m-k}$. To find the distribution of the IQR: Let $Y = X_{\frac{3n}{4}} - X_{\frac{n}{4}}$ and $Z = X_{\frac{n}{4}}$. $X_{\frac{n}{4}} = Z$ and $X_{\frac{3n}{4}} = Y + Z$. The Jacobian matrix $\mathbf{J}$, $\mathbf{J} = \frac{d\mathbf{f}}{d\mathbf{x}} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix}$ and the jacobian determinant is $|\mathbf{J}| = 1$ and so $f_{(Y,Z)}(y, z) = f_{(Y_{\frac{n}{4}}, Y_{\frac{3n}{4}})}(z, y + z) \, |J|$. By using Eq: 2.8

$$\begin{aligned} f_{(Y,Z)}(y, z) &= \frac{n!}{(\frac{n}{4} - 1)!(\frac{3n}{4} - \frac{n}{4} - 1)!(n - \frac{3n}{4})!} \\ &* \quad (F(z))^{\frac{n}{4}-1}(F(y+z) - F(z))^{\frac{3n}{4}-\frac{n}{4}-1}(1 - (F(y+z))^{n-\frac{3n}{4}} f(z)f(y+z) \end{aligned} \tag{2.10}$$

We have $f_{(Y,Z)}(y, z)$ distribution function. So we can find the probability distribution function for the IQR $Y = X_{\frac{3n}{4}} - X_{\frac{n}{4}}$ by using $f_Y(y) = \int_{\min(z)}^{\max(z)} f_{(Y,Z)}(y, z)dz$. Since $a < X_{\frac{n}{4}} < X_{\frac{3n}{4}} < b$,and $a < z < y + z < b \Longrightarrow a < z < b - y$.

The probability distribution function for IQR is obtained as following:

$$\begin{aligned} f_Y(y) &= \int_a^{b-y} f_{(Y,Z)}(y,z)dz \\ &= \int_a^{b-y} \frac{n!}{(\frac{n}{4}-1)!(\frac{3n}{4}-\frac{n}{4}-1)!(n-\frac{3n}{4})!} \\ &* \quad (F(z))^{\frac{n}{4}-1}(F(y+z)-F(z))^{\frac{3n}{4}-\frac{n}{4}-1}(1-(F(y+z))^{n-\frac{3n}{4}}f(z)f(y+z)dz \end{aligned}$$

In this study we consider Weibull, gamma and lognormal distributions. We can obtain the distribution of IQR for this three distributions by using their pdf distributions in Eq: 2.7.

2.3. Modified methods for $\bar{X}$ control chart. The robust methods are one of the most commonly used statistical methods when the underlying normality assumption is violated. These methods offer useful and viable alternative to the traditional statistical methods and can provide more accurate results, often yielding greater statistical power and increased sensitivity and yet still be efficient if the normal assumption is correct [1].

We propose modifications to the Shewhart, weighted variance and skewness correction methods using simple robust estimators to construct $\bar{X}$ control chart for skewed and contaminated process. In this section, we construct the control limits of $\bar{X}$ control chart for skewed populations under the MS, MWD and MSC methods. We estimate μ_x, μ_R and P_X by using robust estimators. The μ_x is estimated using the trimmed mean of the subgroup trimmed means TM_α and μ_R is estimated using the mean of the subgroup interquartile ranges $I\bar{Q}R$. The control limits are derived by assuming that the parameters of the process are unknown.

We first consider the Shewhart method proposed by [12]. The control limits of $\bar{X}$ chart for Shewhart method are given as follows:

$$UCL_{\bar{X}_{Shewhart}} = \bar{\bar{X}} + \frac{3}{d_2\sqrt{n}}\bar{R} \tag{2.11}$$

$$LCL_{\bar{X}_{Shewhart}} = \bar{\bar{X}} - \frac{3}{d_2\sqrt{n}}\bar{R}. \tag{2.12}$$

where d_2 is constant that depends on the subgroup size n, and is calculated when the distribution is normal [12].

The control limits of the $\bar{X}$ chart for MS method are defined as follows:

$$UCL_{\bar{X}_{MS}} = T\bar{M}_\alpha + \frac{3}{d_2^Q\sqrt{n}}I\bar{Q}R, \tag{2.13}$$

$$LCL_{\bar{X}_{MS}} = T\bar{M}_\alpha - \frac{3}{d_2^Q\sqrt{n}}I\bar{Q}R \tag{2.14}$$

where d_2^Q is a constant that depends on the subgroup size n, and is calculated when the distribution is skewed.

The second method investigated is the WV method proposed by [6]. The WV method decompose 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 WV method are defined by [3] as follows:

$$(2.15)\qquad \begin{aligned} UCL_{\bar{X}_{WV}} &= \bar{\bar{X}} + 3\tfrac{\bar{R}}{d_2^*\sqrt{n}}\sqrt{2\hat{P}_x} \\ LCL_{\bar{X}_{WV}} &= \bar{\bar{X}} - 3\tfrac{\bar{R}}{d_2^*\sqrt{n}}\sqrt{2(1-\hat{P}_x)} \end{aligned}$$

where d_2^* is the control chart constant for $\bar{X}$ chart based on WV and $P_X = P(X \leq \bar{X})$ is the probability that the quality variable X will be less than or equal to its mean $\bar{X}$. The constant d_2^* which is defined as the mean of relative range $E\left(\frac{R}{\sigma}\right)$ has been obtained under the non-normality assumption. This value can be computed via numerical integration once the distribution is specified [3].

The control limits of $\bar{X}$ chart for MWV method are defined as follows:

$$(2.16)\qquad UCL_{\bar{x}_{MWV}} = T\bar{M}_\alpha + 3\frac{I\bar{Q}R}{d_2^Q\sqrt{n}}\sqrt{2\hat{P}_x^R}$$

$$(2.17)\qquad LCL_{\bar{x}_{MWV}} = T\bar{M}_\alpha - 3\frac{I\bar{Q}R}{d_2^Q\sqrt{n}}\sqrt{2(1-\hat{P}_x^R)}.$$

where d_2^Q is the control chart constant for $\bar{X}$ chart based on MWV method. This constant, defined as the mean of interquartile range, $d_2^Q = E\left(\frac{IQR}{\sigma}\right)$ is obtained under the non-normality assumption as following:

$$(2.18)\qquad d_2^Q = E\left(\frac{IQR}{\sigma}\right) = \int_{R_{IQR}} \frac{IQR}{\sigma} f_Y(y)dy$$

where R_{IQR} is interval range for IQR and $f_Y(y)$ is the probability density function of interquartile range in Eq. 2.7. As seen it is not easy to obtain this constant for each skewed distribution. Because of the difficulty of numerical integration in Eq. 2.18, this constant based on classic and robust estimators are obtained via simulation for each skewed distribution. Eq. 2.17 allows the probability to be estimated from

$$\hat{P}_X^R = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}\delta\left(T\bar{M}_\alpha X - X_{ij}\right)}{nk}$$

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.

The last method being considered is the SC method proposed by [5] for constructing the $\bar{X}$ and R control charts under skewed distributions. It's asymmetric control limits are obtained by taking into consideration the degree of skewness estimated from subgroups, and making no assumptions about distributions. When the distribution is symmetric, $\bar{X}$ chart is closer to the Shewhart chart.

The control limits of the $\bar{X}$ chart for SC method are defined by [5] as follows:

$$(2.19)\qquad \begin{aligned} UCL_{\bar{X}_{SC}} &= \bar{\bar{X}} + (3 + c_4^*)\tfrac{\bar{R}}{d_2^*\sqrt{n}} \\ LCL_{\bar{X}_{SC}} &= \bar{\bar{X}} + (-3 + c_4^*)\tfrac{\bar{R}}{d_2^*\sqrt{n}} \end{aligned}$$

where c_4^* and d_2^* are the control chart constants for the SC method. The constant c_4^* is obtained as follows:

$$(2.20)\qquad c_4^* = \frac{\frac{4}{3}k_3(\bar{X})}{1 + 0.2k_3^2(\bar{X})}$$

where $k_3(\bar{X})$ is the skewness of the subgroup mean $\bar{X}$ [5].

The control limits of the $\bar{X}$ chart for MSC method are defined as follows:

$$UCL_{\bar{X}_{MSC}} = T\bar{M}_{\alpha} + (3 + c_4^Q)\frac{I\bar{Q}R}{d_2^Q\sqrt{n}} \tag{2.21}$$

$$LCL_{\bar{X}_{MSC}} = T\bar{M}_{\alpha} + (-3 + c_4^Q)\frac{I\bar{Q}R}{d_2^Q\sqrt{n}} \tag{2.22}$$

where c_4^Q is the control chart constant for the MSC method. The constant c_4^Q is obtained as follows:

$$c_4^Q = \frac{\frac{4}{3}k_3(TM_{\alpha})}{1 + 0.2k_3^2(TM_{\alpha})} \tag{2.23}$$

where $k_3(TM_{\alpha})$ is the skewness of the subgroup trimmed means TM_{α}.

A comparison between the performances of the $\bar{X}$ control chart for monitoring the process based on these three modified methods is made in terms of the Type I risk probabilities and the average run length values.

Let E_i denote the event that the ith sample mean is beyond the limits. Further, denote by $P(Ei|\bar{\bar{X}}, \hat{\sigma})$ the conditional probability that for given $\bar{\bar{X}}$ and $\hat{\sigma}$, the sample mean $\bar{X}_i$ is beyond the control limits

$$P(Ei|\bar{\bar{X}}, \hat{\sigma}) = P(\bar{X}_i < LCL \quad or \quad \bar{X}_i > UCL) \tag{2.24}$$

Given $\bar{\bar{X}}$ and $\hat{\sigma}$, the events E_s and E_t $(s \neq t)$ are independent. Therefore, the run length has a geometric distribution with parameter $P(Ei|\bar{\bar{X}}, \hat{\sigma})$. When we take the expectation over the estimation data X_{ij} we get the unconditional probability of one sample showing a Type I false alarm

$$P(Ei) = E(P(Ei|\bar{\bar{X}}, \hat{\sigma})) \tag{2.25}$$

and, similarly, the unconditional average run length (ARL)

$$ARL = E(1/P(Ei|\bar{\bar{X}}, \hat{\sigma})). \tag{2.26}$$

These expectations are simulated by generating 10 000 times k data samples of size n, computing for each data set the conditional value and averaging the conditional values over the data sets. Note that for the calculation of the control limits in Phase I the process is considered to be in-control, thus outliers are omitted in this phase [14].

3. Simulation study

We suggest to use robust estimators for the μ and σ coupled with the MS, MWV and MSC methods for skewed distributions. The Monte Carlo simulation study is considered in this section: The effects of outliers on the classic and robust estimations are evaluated in terms of their root mean-square errors in Section 3.1. The control chart constants are obtained for skewed distributions in Section 3.2. The performance of the control chart is compared using the Type I risk probabilities and average run lengths of these control charts in Section 3.3 , when the contamination is considered in Phase I and Phase II procedures.

3.1. Effect of outliers on estimations. In this section, we evaluate the effect of outliers on the accuracy of the conventional and robust estimators by means of simulation. $(M = 50.000)$ simulation runs of 30 $(k = 30)$ subgroups each of size n=5,10 are performed to generate data on skewed distributions. The distributions of the generated data are from Weibull, lognormal and gamma distributions with different parameters. The process dispersion is estimated by both classic and robust methods. We consider four model in the case of no outliers and outliers like [8],

- **Model 1:** The reference distribution parameters are selected with respect to skewness of distribution given in Table 1.
- **Model 2:** The case of 10% replacement outliers coming from another Weibull distribution with a different scale parameter ($\lambda_1 = 0.2$) and a shape parameter of ($\beta_1 = 02*\beta$) , another lognormal distribution with a different location parameter ($\mu_1 = 0.2$) and a scale parameter of ($\sigma_1 = 2*\sigma$) and another gamma distribution with a different shape parameter ($\alpha_1 = 2\alpha$) and a scale parameter of ($\beta_1 = 0.2$).
- **Model 3:** A case with 10% replacement outliers from a uniform distribution on $[0, 20]$.
- **Model 4:** A more extreme case with 10% of outliers placed at 50.

We thus allow that some observations come from a different skewed population and, in the last two models, we permit the occurrence of gross errors.

Table 1. The values of the P_X, the skewness and the parameters of distributions

	Lognormal		*Weibull*		*Gamma*	
k_3	σ	P_X	β	P_X	α	P_X
0.50	0.16	0.53	2.15	0.54	16.00	0.53
1.00	0.32	0.56	1.57	0.57	4.00	0.57
1.50	0.44	0.59	1.20	0.61	1.80	0.60
2.00	0.54	0.61	1.00	0.63	1.00	0.63
2.50	0.66	0.63	0.86	0.66	0.64	0.66
3.00	0.72	0.64	0.77	0.68	0.44	0.69

We run the simulation $M = 50.000$ times and generate $k = 30$ samples of size $n = 5$ and $n = 10$ according to different simulation schemes. For each sample, we compute the location estimate $\hat{\mu}_j$ and the scale estimate $\hat{\sigma}_j$, for $j = 1, \ldots, M$. For each simulation setting and each type of estimator, we compute the root mean squared error

$$RMSE_\mu = \sqrt{\frac{1}{M}\sum_{j=1}^{M}(\hat{\mu}_j - \mu_0)^2} \ , \ RMSE_\sigma \sqrt{\frac{1}{M}\sum_{j=1}^{M}(\hat{\sigma}_j - \sigma_0)^2}.$$

The results for the Weibull, lognormal and gamma distributions are reported in Table 2, Table 3 and Table 4, respectively. The conclusions drawn from the study are as follows.

(i) When there is no contamination, the classic estimators of mean and scale perform best, as expected.
(ii) Contamination by extreme outliers causes a large increase in the $RMSE$ of the classic estimators especially for large samples $n = 10$, and a much smaller increase in the $RMSE$ of the robust alternatives.
(iii) For the estimation of mean, the trimmed mean estimator performs better for large sample size than the small sample size, especially when there is contamination by extreme outliers. This is true for all considered distribution.
(iii) For scale estimation, the interquartile range estimator performs better for large sample size than the small sample size across all distributions, especially when there is contamination by extreme outliers.
(iv) In the presence of outliers, the classic scale estimator has the highest $RMSE$ of all skewed distributions except the scale estimator for gamma distribution less than 2 for Model 1, when $n = 5$ (for small sample size).

(v) For the estimation of both mean and scale, the robust estimators have a lower $RMSE$ than the classical estimator in Model 3 and Model 4.

Table 2. $RMSE$ of the $\hat{\mu}$ and $\hat{\sigma}$ estimators for Weibull Distribution, n = 5 ,10

		$\hat{\mu}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.0355	0.0475	0.0643	0.0820	0.1035	0.1249	0.0253	0.0336	0.0457	0.0576	0.0728	0.0887
	Robust	0.0494	0.0875	0.1497	0.2184	0.3012	0.3842	0.0345	0.0620	0.1075	0.1585	0.2211	0.2852
Model 2	**Classic**	0.0493	0.0668	0.0904	0.1131	0.1405	0.1682	0.0371	0.0506	0.0677	0.0846	0.1049	0.1250
	Robust	0.0418	0.0574	0.0980	0.1509	0.2196	0.2908	0.0286	0.0380	0.0653	0.1026	0.1522	0.2054
Model 3	**Classic**	0.7766	0.7767	0.7743	0.7706	0.7663	0.7628	0.6139	0.6131	0.6106	0.6091	0.6042	0.6024
	Robust	0.0528	0.0589	0.0800	0.1174	0.1712	0.2318	0.0434	0.0464	0.0526	0.0664	0.0929	0.1265
Model 4	**Classic**	3.9295	3.9285	3.9249	3.9205	3.9149	3.9082	3.1029	3.1029	3.0992	3.0973	3.0911	3.0853
	Robust	0.0545	0.0597	0.0782	0.1117	0.1612	0.2171	0.0452	0.0492	0.0546	0.0660	0.0858	0.1125
		$\hat{\sigma}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.0293	0.0446	0.0712	0.1048	0.1520	0.2036	0.1363	0.1899	0.2742	0.3741	0.5095	0.6592
	Robust	0.0337	0.0491	0.0738	0.1042	0.1460	0.1924	0.0290	0.0409	0.0623	0.0934	0.1453	0.2117
Model 2	**Classic**	0.0315	0.0492	0.0799	0.1176	0.1707	0.2291	0.1467	0.2093	0.3067	0.4230	0.5791	0.7509
	Robust	0.0382	0.0578	0.0901	0.1298	0.1836	0.2432	0.0347	0.0501	0.0686	0.0873	0.1154	0.1511
Model 3	**Classic**	1.5846	1.5864	1.5966	1.6086	1.6191	1.6268	2.5950	2.6168	2.6653	2.7317	2.8020	2.8895
	Robust	0.7251	0.7575	0.8178	0.8898	0.9804	1.0721	0.0591	0.0893	0.1281	0.1664	0.2067	0.2429
Model 4	**Classic**	8.4085	8.5658	8.8463	9.1533	9.5031	9.8179	13.3728	13.6361	14.1158	14.6384	15.2350	15.7725
	Robust	3.6826	3.8093	4.0516	4.3354	4.6874	5.0316	0.0621	0.0942	0.1386	0.1826	0.2323	0.2821

Table 3. $RMSE$ of the $\hat{\mu}$ and $\hat{\sigma}$ estimators for lognormal distribution, n = 5 ,10

		$\hat{\mu}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.0362	0.0768	0.1131	0.1491	0.2051	0.2377	0.0255	0.0541	0.0795	0.1052	0.1448	0.1688
	Robust	0.0373	0.0798	0.1187	0.1581	0.2197	0.2567	0.0315	0.0889	0.1576	0.2368	0.3648	0.4438
Model 2	**Classic**	0.1225	0.1333	0.1509	0.1779	0.2447	0.3044	0.0963	0.1018	0.1121	0.1277	0.1703	0.2078
	Robust	0.1236	0.1466	0.1753	0.2083	0.2644	0.2975	0.0680	0.1448	0.2188	0.2978	0.4226	0.4995
Model 3	**Classic**	0.6328	0.6267	0.6208	0.6176	0.6131	0.6146	0.4975	0.4924	0.4875	0.4836	0.4785	0.4764
	Robust	0.4677	0.4895	0.5038	0.5136	0.5113	0.5043	0.0398	0.0664	0.0882	0.1170	0.1827	0.2328
Model 4	**Classic**	3.7608	3.7554	3.7401	3.7308	3.7148	3.7037	2.9849	2.9779	2.9701	2.9619	2.9513	2.9396
	Robust	3.5358	3.5356	3.5202	3.5105	3.4914	3.4779	0.0485	0.0825	0.1032	0.1231	0.1549	0.1797
		$\hat{\sigma}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.0314	0.0752	0.1247	0.1839	0.2902	0.3610	0.1481	0.3266	0.5037	0.6972	1.0141	1.2212
	Robust	0.0352	0.0798	0.1263	0.1777	0.2675	0.3269	0.0298	0.0685	0.1118	0.1679	0.2716	0.3461
Model 2	**Classic**	0.1640	0.1279	0.1631	0.2442	0.4417	0.6190	0.3744	0.4560	0.6217	0.8501	1.3030	1.6448
	Robust	0.1018	0.1089	0.1473	0.2064	0.3297	0.4311	0.0394	0.0669	0.1034	0.1528	0.2474	0.3152
Model 3	**Classic**	1.3094	1.2203	1.1497	1.0904	1.0209	0.9931	2.1632	2.1307	2.1254	2.1454	2.2356	2.3170
	Robust	0.6167	0.6405	0.6767	0.7213	0.7938	0.8367	0.0519	0.1131	0.1655	0.2109	0.2725	0.3057
Model 4	**Classic**	8.0412	8.0930	8.1701	8.2778	8.4516	8.5472	12.8626	13.0311	13.2564	13.5063	13.9106	14.1228
	Robust	3.5883	3.7202	3.8852	4.0817	4.4061	4.6041	0.0557	0.1308	0.2028	0.2763	0.3863	0.4468

Table 4. $RMSE$ of the $\hat{\mu}$ and $\hat{\sigma}$ estimators for gamma distribution, n = 5 ,10

		$\hat{\mu}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.3273	0.1641	0.1095	0.0820	0.0654	0.0541	0.2296	0.1153	0.0776	0.0575	0.0461	0.0381
	Robust	0.3377	0.1708	0.1163	0.0889	0.0727	0.0619	0.2883	0.1969	0.1716	0.1550	0.1421	0.1297
Model 2	**Classic**	0.8229	0.2467	0.1363	0.0918	0.0698	0.0564	0.6464	0.1892	0.1012	0.0677	0.0509	0.0409
	Robust	0.8609	0.2776	0.1612	0.1131	0.0888	0.0739	0.6243	0.3352	0.2381	0.1919	0.1646	0.1438
Model 3	**Classic**	0.5539	0.5568	0.7118	0.7733	0.7987	0.8133	0.4226	0.4338	0.5605	0.6097	0.6291	0.6414
	Robust	0.5684	0.4788	0.5823	0.6184	0.6294	0.6337	0.4501	0.1300	0.0868	0.0652	0.0564	0.0523
Model 4	**Classic**	2.7260	3.6648	3.8401	3.9028	3.9324	3.9433	2.1568	2.9089	3.0475	3.0955	3.1175	3.1314
	Classic	2.6244	3.4622	3.6140	3.6660	3.6904	3.6972	0.4304	0.1734	0.0950	0.0644	0.0545	0.0506
		$\hat{\sigma}$											
		n=5						n=10					
	Model/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**Classic**	0.2813	0.1570	0.1203	0.1048	0.0966	0.0927	1.3256	0.6772	0.4724	0.3748	0.3227	0.2938
	Robust	0.3183	0.1698	0.1241	0.1045	0.0946	0.0896	0.2693	0.1425	0.1062	0.0940	0.0941	0.1023
Model 2	**Classic**	0.8501	0.1626	0.1228	0.1089	0.1004	0.0953	2.3528	0.7093	0.4487	0.3469	0.2969	0.2681
	Robust	0.6610	0.1803	0.1261	0.1098	0.0998	0.0939	0.3639	0.1428	0.1130	0.1143	0.1192	0.1267
Model 3	**Classic**	0.5513	0.8855	1.3340	1.6147	1.8183	2.0093	1.9356	1.8544	2.3800	2.7347	3.0083	3.2867
	Robust	0.4404	0.5739	0.7623	0.8931	1.0037	1.1248	0.2813	0.2239	0.1900	0.1663	0.1454	0.1270
Model 4	**Classic**	4.9489	7.6274	8.4630	9.1127	9.7395	10.4095	8.7237	12.5101	13.6815	14.6295	15.5769	16.6415
	Robust	2.5862	3.6174	3.9768	4.3167	4.7097	5.1889	0.5141	0.2969	0.2233	0.1815	0.1535	0.1318

Table 5. The values of the constants for the skewed distributions for n=3,5

	Weibull				Lognormal				Gamma			
	n =3											
k_3	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q
0.50	1.6880	0.3702	1.2660	0.3702	1.6776	0.3337	1.2582	0.3337	1.6791	0.3414	1.2593	0.3414
1.00	1.6447	0.6537	1.2335	0.6537	1.6352	0.6547	1.2264	0.6547	1.6406	0.6515	1.2305	0.6515
1.50	1.5726	0.9223	1.1795	0.9223	1.5860	0.8784	1.1895	0.8784	1.5804	0.9012	1.1853	0.9012
2.00	1.4995	1.1017	1.1246	1.1017	1.5335	1.0381	1.1502	1.0381	1.5001	1.1033	1.1251	1.1033
2.50	1.4221	1.2355	1.0665	1.2355	1.4587	1.1940	1.0940	1.1940	1.4102	1.2429	1.0577	1.2429
3.00	1.3552	1.3162	1.0164	1.3162	1.4174	1.2529	1.0630	1.2529	1.3157	1.3386	0.9868	1.3386
	n=5											
k_3	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q
0.50	2.3088	0.2879	1.3332	0.2879	2.3092	0.2602	1.3094	0.2602	2.3089	0.2669	1.3122	0.2669
1.00	2.2559	0.5173	1.2921	0.5173	2.2575	0.5223	1.2665	0.5223	2.2595	0.5163	1.2773	0.5163
1.50	2.1702	0.7529	1.2201	0.7529	2.1974	0.7207	1.2166	0.7207	2.1827	0.7362	1.2218	0.7362
2.00	2.0831	0.9283	1.1459	0.9283	2.1346	0.8787	1.1656	0.8787	2.0827	0.9281	1.1457	0.9281
2.50	1.9903	1.0764	1.0672	1.0764	2.0423	1.0527	1.0932	1.0527	1.9758	1.0811	1.0587	1.0811
3.00	1.9102	1.1819	1.0003	1.1819	1.9911	1.1274	1.0539	1.1274	1.8621	1.2011	0.9635	1.2011

3.2. Determination of the control charts constants. An assumption of non-normality is incorporated into the constants d_2 and c_4 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.

The WV method constant d_2^* is calculated by taking the mean of range $\left(\frac{R}{\sigma}\right)$. In this study, we consider the modified WV method constant d_2^Q which is calculated by taking the mean of interquartile range $\left(\frac{IQR}{\sigma}\right)$. The SC method constant c_4^* is calculated by using Eq: 2.20. We consider the MSC method constant c_4^Q, which is calculated using Eq: 2.23. All constants are obtained for three skewed distributions via simulation. We obtain $E(I\bar{Q}R)$ by simulation: we generate 100.000 times k samples of size n, compute

Table 6. The values of the constants for the skewed distributions for n=7,10

	Weibull				Lognormal				Gamma			
					n=7							
k_3	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q
0.50	2.6721	0.2415	1.3345	0.2592	2.6877	0.2229	1.2947	0.2291	2.6858	0.2253	1.2992	0.2351
1.00	2.6172	0.4418	1.2861	0.4921	2.6381	0.4480	1.2423	0.4468	2.6328	0.4413	1.2602	0.4655
1.50	2.5340	0.6522	1.1988	0.7211	2.5790	0.6268	1.1826	0.6063	2.5531	0.6368	1.1974	0.6820
2.00	2.4499	0.8158	1.1084	0.8854	2.5159	0.7773	1.1211	0.7319	2.4502	0.8146	1.1083	0.8848
2.50	2.3601	0.9660	1.0130	1.0248	2.4231	0.9512	1.0359	0.8707	2.3417	0.9647	1.0044	1.0569
3.00	2.2808	1.0758	0.9317	1.1208	2.3696	1.0323	0.9895	0.9341	2.2306	1.0931	0.8893	1.2006
					n=10							
k_3	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q	d_2^*	c_4^*	d_2^Q	c_4^Q
0.50	3.0213	0.2044	1.3415	0.2169	3.0640	0.1826	1.2928	0.1850	3.0587	0.1881	1.2982	0.1938
1.00	2.9709	0.3708	1.2892	0.4052	3.0225	0.3787	1.2350	0.3739	3.0050	0.3726	1.2572	0.3871
1.50	2.8990	0.5531	1.1929	0.6012	2.9701	0.5368	1.1698	0.5117	2.9258	0.5423	1.1893	0.5711
2.00	2.8301	0.7046	1.0930	0.7529	2.9145	0.6748	1.1030	0.6242	2.8287	0.7032	1.0926	0.7529
2.50	2.7530	0.8464	0.9875	0.8859	2.8300	0.8432	1.0113	0.7547	2.7323	0.8450	0.9786	0.9142
3.00	2.6842	0.9586	0.8980	0.9850	2.7806	0.9246	0.9617	0.8153	2.6348	0.9715	0.8504	1.0598

IQR for each instance and take the average of the values. The results for all constants for $k = 30$ are presented in Table 5 for $n = 3, 5$ and Table 6 for $n = 7, 10$.

3.3. Performance of modified methods. 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%. However, due to the control limits, about 0.0027 of all control points will be false alarms and have no assignable cause of variation. The ARL is the number of points plotted within the control limits before one exceeds the limits. Under the normality assumption and for the Shewhart control charts, it is expected that 370.4 points would be plotted on the chart within the 3σ control limits, before one gets out. If the process is in-control, we want the in-control average run length, ARL_0, to be large. If the process is out-of-control, we want the out-of-control average run length, ARL_1, to be small.

In this section, we consider design schemes for the $\bar{X}$ control chart for non-contaminated and contaminated skewed distributed data. We use the mean and the trimmed mean estimators of mean and the range and the interquartile range estimators of the standard deviation for the Shewhart, WV and SC methods. To evaluate the control chart performance we obtain p and the in-control ARL for moderate sample size (30 subgroups of 3-10) for each skewed distribution. The simulation consists of two Phases. The steps of each Phase are described as following.

Phase I:

1.a. Generate n *i.i.d.* Weibull $(\beta, 1)$, gamma$(\alpha, 1)$ and lognormal$(1, \sigma)$ varieties for $n = 3, 5, 7, 10$.

1.b. Repeat step 1.a 30 times $(k = 30)$.

1.c. By using classic estimators compute the control limits for Shewhart, the WV and the SC methods. By using robust estimators compute the control limits for the MS, the MWV and the MSC methods.

Phase II:

2.a. Generate n *i.i.d.* Weibull$(\beta, 1)$, gamma$(\alpha, 1)$ and lognormal $(1, \sigma)$ 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 $\bar{X}$ chart for the Shewhart, WV and SC methods. Compute the robust estimator interquartile range IQR for the MS, MWV and MSC 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.

In the simulation study, we consider non-contaminated and contaminated data set in Phase I and Phase II. We consider the 20% trimmed mean, which trims the six smallest and the six largest sample trimmed means when $k = 30$.

- **Non-contaminated case:** The reference distribution parameters are selected with respect to skewness of distribution given in Table 1.
- **Contaminated case:** The more extreme case of 10% of outliers placed at 50.

The simulation results of p for the $\bar{X}$ control chart for non-contaminated data under skewed distributions are given in Table 7 for small sample sizes and Table 8 for large sample size. The results of ARL for the $\bar{X}$ control chart for non-contaminated data under skewed distributions are given in Table 9 for small sample sizes and Table 10 for large sample size. The results of p and ARL for the $\bar{X}$ control chart for contaminated Weibull, lognormall and gamma distrubuted data are given in Table 11, Table 12 and Table 13, respectively.

Table 7. Results of the p for the $\bar{X}$ control chart based on classic and robust estimators for small sample sizes

		n=3											
		Classic Estimators						Robust Estimators					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	0.0050	0.0080	0.0119	0.0152	0.0186	0.0203	0.0053	0.0089	0.0134	0.0172	0.0206	0.0233
	WV	0.0042	0.0058	0.0081	0.0100	0.0120	0.0131	0.0047	0.0068	0.0096	0.0120	0.0143	0.0161
	SC	0.0034	0.0035	0.0037	0.0043	0.0051	0.0060	0.0035	0.0035	0.0038	0.0045	0.0056	0.0068
Lognormal	**Shewhart**	0.0055	0.0086	0.0116	0.0142	0.0170	0.0181	0.0058	0.0094	0.0129	0.0157	0.0188	0.0201
	WV	0.0051	0.0068	0.0086	0.0103	0.0121	0.0129	0.0054	0.0077	0.0101	0.0120	0.0142	0.0152
	SC	0.0046	0.0052	0.0058	0.0062	0.0065	0.0066	0.0046	0.0050	0.0055	0.0058	0.0063	0.0067
Gamma	**Shewhart**	0.0049	0.0079	0.0121	0.0153	0.0178	0.0189	0.0046	0.0059	0.0075	0.0095	0.0113	0.0134
	WV	0.0041	0.0057	0.0081	0.0100	0.0115	0.0120	0.0044	0.0051	0.0061	0.0072	0.0082	0.0094
	SC	0.0033	0.0033	0.0038	0.0042	0.0048	0.0052	0.0042	0.0043	0.0038	0.0026	0.0020	0.0025
		n=5											
		Classic Estimators						Robust Estimators					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	0.0042	0.0061	0.0090	0.0115	0.0144	0.0167	0.0046	0.0067	0.0101	0.0131	0.0161	0.0184
	WV	0.0036	0.0043	0.0056	0.0070	0.0085	0.0099	0.0040	0.0049	0.0066	0.0082	0.0100	0.0114
	SC	0.0032	0.0032	0.0033	0.0035	0.0039	0.0045	0.0035	0.0033	0.0034	0.0036	0.0042	0.0049
Lognormal	**Shewhart**	0.0045	0.0067	0.0091	0.0113	0.0140	0.0153	0.0048	0.0073	0.0098	0.0123	0.0152	0.0166
	WV	0.0042	0.0051	0.0064	0.0077	0.0094	0.0102	0.0045	0.0057	0.0072	0.0087	0.0106	0.0115
	SC	0.0039	0.0042	0.0047	0.0052	0.0057	0.0059	0.0041	0.0043	0.0046	0.0049	0.0053	0.0055
Gamma	**Shewhart**	0.0042	0.0062	0.0093	0.0118	0.0141	0.0152	0.0048	0.0069	0.0098	0.0131	0.0163	0.0195
	WV	0.0036	0.0044	0.0058	0.0071	0.0083	0.0089	0.0044	0.0053	0.0067	0.0083	0.0098	0.0114
	SC	0.0032	0.0032	0.0034	0.0035	0.0038	0.0039	0.0040	0.0039	0.0037	0.0037	0.0041	0.0050

Table 8. Results of the p for the $\bar{X}$ control chart based on classic and robust estimators for large sample sizes

		n=7											
		Classic Estimators						Robust Estimators					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	0.0038	0.0053	0.0077	0.0100	0.0122	0.0144	0.0054	0.0068	0.0086	0.0088	0.0081	0.0065
	WV	0.0033	0.0038	0.0047	0.0057	0.0068	0.0081	0.0051	0.0058	0.0068	0.0067	0.0059	0.0046
	SC	0.0032	0.0031	0.0033	0.0035	0.0036	0.0041	0.0045	0.0037	0.0030	0.0022	0.0016	0.0011
Lognormal	**Shewhart**	0.0041	0.0058	0.0078	0.0097	0.0123	0.0137	0.0045	0.0065	0.0086	0.0104	0.0124	0.0133
	WV	0.0039	0.0045	0.0054	0.0064	0.0079	0.0088	0.0042	0.0050	0.0062	0.0072	0.0085	0.0091
	SC	0.0037	0.0039	0.0044	0.0047	0.0054	0.0058	0.0037	0.0036	0.0037	0.0039	0.0043	0.0046
Gamma	**Shewhart**	0.0040	0.0054	0.0079	0.0100	0.0121	0.0133	0.0059	0.0068	0.0080	0.0088	0.0091	0.0089
	WV	0.0036	0.0039	0.0048	0.0057	0.0067	0.0073	0.0057	0.0060	0.0065	0.0066	0.0065	0.0060
	SC	0.0033	0.0033	0.0035	0.0034	0.0035	0.0035	0.0052	0.0043	0.0031	0.0021	0.0019	0.0017
		n=10											
		Classic Estimators						Robust Estimators					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	0.0037	0.0047	0.0065	0.0086	0.0105	0.0120	0.0056	0.0061	0.0067	0.0069	0.0066	0.0063
	WV	0.0033	0.0035	0.0040	0.0047	0.0055	0.0063	0.0053	0.0051	0.0049	0.0048	0.0043	0.0040
	SC	0.0032	0.0033	0.0034	0.0035	0.0035	0.0039	0.0050	0.0040	0.0027	0.0019	0.0015	0.0013
Lognormal	**Shewhart**	0.0038	0.0052	0.0067	0.0084	0.0107	0.0119	0.0049	0.0051	0.0053	0.0053	0.0051	0.0049
	WV	0.0036	0.0041	0.0046	0.0054	0.0066	0.0073	0.0047	0.0044	0.0042	0.0040	0.0037	0.0034
	SC	0.0035	0.0038	0.0041	0.0045	0.0051	0.0054	0.0044	0.0036	0.0027	0.0020	0.0014	0.0012
Gamma	**Shewhart**	0.0040	0.0051	0.0068	0.0083	0.0101	0.0114	0.0045	0.0063	0.0087	0.0105	0.0119	0.0124
	WV	0.0036	0.0038	0.0042	0.0045	0.0053	0.0059	0.0041	0.0048	0.0058	0.0063	0.0067	0.0067
	SC	0.0035	0.0035	0.0035	0.0034	0.0034	0.0034	0.0037	0.0035	0.0034	0.0033	0.0033	0.0031

Table 9. Results of the ARL for the $\bar{X}$ control chart based on classic and robust estimators for small sample sizes

		n=3											
		Classic Estimators						Robust Estimators					
	k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	201.4504	124.3626	84.1255	65.8111	53.6251	49.2029	187.7335	112.9663	74.8671	58.0953	48.5753	42.9697
	WV	236.1833	173.2502	124.1003	100.2406	83.3681	76.3475	213.2378	147.2624	104.3515	83.2404	70.1671	62.2944
	SC	289.9391	288.6003	267.2368	232.6664	196.3479	166.7779	283.6075	285.9676	263.4213	220.2110	177.5726	147.6799
Lognormal	**Shewhart**	181.0217	115.7501	86.1876	70.4072	58.9299	55.2129	172.2030	106.2236	77.7267	63.8659	53.3294	49.7867
	WV	195.7522	146.6233	116.7229	97.5096	82.8947	77.8053	183.5132	129.5874	99.2339	83.2646	70.3284	65.7921
	SC	218.3978	192.6634	172.1467	160.0384	152.8865	150.6206	215.4197	198.1728	181.5640	172.5923	158.0303	150.1998
Gamma	**Shewhart**	206.1091	126.7990	82.5430	65.4095	56.0582	53.0453	172.3811	110.3327	76.3161	58.1061	47.6917	40.4073
	WV	241.8614	175.4879	123.4583	100.2486	87.2007	83.3021	186.4211	138.9024	103.5508	83.2903	70.6230	61.2329
	SC	302.7184	298.7482	263.5532	237.2085	206.2791	190.8615	224.8303	232.6068	241.8906	221.2536	178.7662	142.6737
		n=5											
		Classic Estimators						Robust Estimators					
Weibull	**Shewhart**	239.2917	158.8058	110.6562	83.4376	69.5894	58.8582	219.0245	148.4296	99.1405	76.2207	61.9602	54.2041
	WV	277.1619	228.2063	176.7721	138.3892	117.1097	100.1201	250.0188	204.4321	152.1422	121.6871	99.7914	87.5695
	SC	308.6420	303.8590	318.0662	280.8200	253.1005	219.2982	289.0925	299.7422	297.2828	274.4086	236.6136	202.9221
Lognormal	**Shewhart**	221.2389	149.9790	110.0606	88.2488	71.2560	65.3202	206.2323	137.1629	101.5713	81.1293	65.9191	60.1063
	WV	239.9981	194.4239	156.0184	129.4230	106.2699	97.9489	221.4790	173.9221	139.1866	114.9293	94.4136	86.5883
	SC	257.3075	237.2423	212.7886	192.4039	174.1311	168.7849	243.9560	232.0724	218.4455	204.6748	189.8722	182.4185
Gamma	**Shewhart**	240.8362	161.2279	107.6739	84.8731	71.1081	65.7086	209.4592	144.0320	101.6322	76.1278	61.4881	51.3168
	WV	277.9322	229.6159	171.5796	141.6712	120.5342	112.1152	226.9014	188.6081	150.2494	120.9278	102.2275	87.8418
	SC	311.1775	310.1929	293.1520	286.4837	266.3896	255.8526	252.7806	258.6185	272.9332	271.7022	242.1366	198.6808

Table 10. Results of the ARL for the $\bar{X}$ control chart based on classic and robust estimators for large sample sizes

		n=7											
		Classic Estimators						Robust estimators					
	k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull	**Shewhart**	263.4352	187.5117	129.9883	99.9300	81.7127	69.5749	185.6838	147.4861	116.3982	113.7229	122.8818	153.4213
	WV	299.0431	264.2706	214.5923	175.4078	146.9076	124.1773	195.3774	173.1722	146.9076	149.6334	168.1520	219.4234
	SC	312.9890	319.6931	305.5301	289.0173	278.7845	242.1894	223.3639	268.5285	334.8289	459.8547	619.8859	939.9380
Lognormal	**Shewhart**	241.6276	172.0312	128.2150	102.9824	81.4233	73.1507	223.3240	154.6623	116.5488	96.3419	80.4466	74.9895
	WV	258.6987	222.0101	184.2401	156.2940	126.2499	114.1826	240.8942	198.7360	162.2955	138.0167	117.3778	110.0219
	SC	272.7248	255.3952	229.2999	211.1397	185.5976	173.6986	269.1355	276.9853	271.8573	257.6257	230.8989	217.0421
Gamma	**Shewhart**	250.0438	183.6446	125.9287	100.1422	82.8995	75.4245	170.0912	146.0110	124.6090	114.2583	110.2196	112.7332
	WV	281.3969	256.4366	206.6799	175.5433	148.6171	136.4685	174.9567	166.5584	154.3925	150.9434	153.6641	166.1323
	SC	299.6344	299.8591	289.3686	295.0549	285.4777	284.0022	190.7851	233.7049	325.3831	468.5596	530.7011	589.4836
		n=10											
		Classic Estimators						Robust Estimators					
Weibull	**Shewhart**	273.1494	211.5507	154.3210	116.6181	95.3107	83.0496	179.1088	162.9673	148.8716	145.3848	150.3850	157.9854
	WV	302.4803	284.5760	253.0364	211.7747	182.4818	157.7785	188.4979	195.3850	202.9015	210.2077	229.9432	249.3393
	SC	307.9766	303.4901	297.4420	288.6003	282.4061	256.3445	201.5154	248.6634	371.2090	523.6425	678.2420	787.0917
Lognormal	**Shewhart**	260.3828	193.3899	150.2268	119.4172	93.8069	84.1128	206.0624	195.4117	190.2081	188.8004	194.5374	203.7739
	WV	277.4926	244.0274	215.1880	184.4916	152.3879	136.8738	213.5292	225.9019	239.1944	250.5136	272.0644	290.9937
	SC	285.4126	264.4313	246.2387	223.7737	196.1631	184.2197	224.8454	280.5994	369.6721	488.7347	694.2034	812.9420
Gamma	**Shewhart**	249.4574	197.9257	146.0750	120.3920	98.7147	87.4027	223.8138	157.8034	115.1145	95.0480	84.3526	80.9454
	WV	275.2092	263.9637	238.7490	219.9784	187.5574	169.4083	243.5460	209.3364	173.1902	158.1028	149.7679	149.2983
	SC	282.7415	281.9045	283.9860	296.1822	291.5282	293.9793	271.0762	286.6972	291.2056	306.9368	303.7667	321.1304

Table 11. Results of the p and ARL for the $\bar{X}$ control chart for contaminated Weibull distribution

		n=5											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**MS**	0.0033	0.0049	0.0072	0.0094	0.0116	0.0132	300.9601	205.6302	137.9805	106.3140	86.4013	75.7708
	MWV	0.0030	0.0035	0.0046	0.0057	0.0069	0.0079	337.8378	284.8029	215.8895	174.4379	144.4127	127.2734
	MSC	0.0027	0.0024	0.0023	0.0023	0.0026	0.0030	376.1520	411.4718	433.9713	433.2380	382.4092	335.5254
Model 2	**MS**	0.1792	0.1541	0.1181	0.0820	0.0469	0.0263	5.5800	6.4873	8.4662	12.1966	21.3049	38.0451
	MWV	0.2030	0.1237	0.0752	0.0425	0.0195	0.0097	4.9252	8.0827	13.2959	23.5185	51.2768	102.5694
	MSC	0.1647	0.1240	0.0684	0.0281	0.0086	0.0034	6.0725	8.0670	14.6236	35.5745	115.6283	290.3853
Model 3*	**MS**	0.2811	0.2492	0.1605	0.0882	0.0386	0.0163	0.0036	0.0040	0.0062	0.0113	0.0259	0.0613
	MWV	0.6631	0.6425	0.6101	0.5401	0.3588	0.1521	0.0015	0.0016	0.0016	0.0019	0.0028	0.0066
	MSC	0.2279	0.1183	0.0309	0.0065	0.0010	0.0002	0.0044	0.0085	0.0324	0.1541	1.0508	4.2337
		n=10											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1*	**MS**	0.0043	0.0047	0.0050	0.0051	0.0049	0.0045	0.2317	0.2141	0.1981	0.1945	0.2062	0.2224
	MWV	0.0041	0.0039	0.0037	0.0035	0.0031	0.0027	0.2430	0.2573	0.2717	0.2886	0.3239	0.3650
	MSC	0.0039	0.0031	0.0020	0.0013	0.0009	0.0008	0.2540	0.3267	0.5101	0.7807	1.0861	1.3307
Model 2	**MS**	0.0050	0.0060	0.0073	0.0080	0.0082	0.0079	201.9508	165.8760	136.0859	124.6976	122.2240	126.7363
	MWV	0.0045	0.0047	0.0051	0.0053	0.0051	0.0048	221.1460	212.5037	195.3812	189.8109	194.2238	209.1569
	MSC	0.0042	0.0036	0.0028	0.0023	0.0020	0.0017	239.0229	279.4623	357.1173	437.0247	509.1909	593.7537
Model 3	**MS**	0.0048	0.0059	0.0075	0.0085	0.0093	0.0099	208.8206	168.0983	133.6380	117.4936	107.1260	101.4188
	MWV	0.0043	0.0046	0.0051	0.0055	0.0059	0.0061	230.2715	217.6468	195.8365	180.4175	170.1201	163.4147
	MSC	0.0040	0.0035	0.0028	0.0025	0.0024	0.0024	247.7701	285.7878	351.4197	399.9840	417.9204	417.7109

4. Results

In this section, the performance of different design schemes is evaluated. When the process in control, it is expected that p is to be as low as possible and ARL is to be as high as possible. The desired ARL value of 370 indicates that the control limits are chosen to provide a p of 0.0027. First we consider the design scheme where the process has a skewed distribution and the Phase I data are non-contaminated. Tables 7,8, 9 and 10 present the p and ARL values for the $\bar{X}$ control chart under the skewed distributions. The tables indicates the following points:

- When the distribution is approximately symmetric ($k_3 = 0.5$), the p of the SC, WV and Shewhart charts are comparable, while the SC $\bar{X}$ chart has a noticeable

Table 12. Results of the p and ARL for the $\bar{X}$ control chart for contaminated Lognormal distribution

		n=5											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1*	**MS**	0.2715	0.1555	0.0696	0.0293	0.0101	0.0067	0.0037	0.0064	0.0144	0.0342	0.0989	0.1500
	MWV	0.6595	0.5970	0.4204	0.1921	0.0311	0.0077	0.0015	0.0017	0.0024	0.0052	0.0321	0.1299
	MSC	0.2050	0.0522	0.0086	0.0016	0.0005	0.0004	0.0049	0.0192	0.1169	0.6117	1.9577	2.4820
Model 2*	**MS**	0.2729	0.1547	0.0705	0.0295	0.0102	0.0068	0.0037	0.0065	0.0142	0.0339	0.0984	0.1480
	MWV	0.6602	0.5972	0.4198	0.1914	0.0306	0.0078	0.0015	0.0017	0.0024	0.0052	0.0327	0.1276
	MSC	0.2060	0.0513	0.0086	0.0016	0.0005	0.0004	0.0049	0.0195	0.1166	0.6143	1.9701	2.4172
Model 3*	**MS**	0.2726	0.1549	0.0696	0.0291	0.0101	0.0066	0.0037	0.0065	0.0144	0.0343	0.0993	0.1508
	MWV	0.6598	0.5972	0.4204	0.1921	0.0306	0.0078	0.0015	0.0017	0.0024	0.0052	0.0327	0.1275
	MSC	0.2053	0.0521	0.0085	0.0016	0.0005	0.0004	0.0049	0.0192	0.1178	0.6339	1.8911	2.4420
		n=10											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**MS**	0.0037	0.0044	0.0051	0.0055	0.0056	0.0055	273.6502	224.9769	194.5904	182.0930	178.7374	180.7566
	MWV	0.0036	0.0039	0.0041	0.0041	0.0041	0.0040	276.3500	257.0628	243.0665	241.0742	246.0509	253.1069
	MSC	0.0033	0.0030	0.0025	0.0020	0.0016	0.0015	299.3743	334.2246	396.6995	488.0429	610.1281	673.6275
Model 2	**MS**	0.0051	0.0063	0.0071	0.0071	0.0062	0.0054	194.3370	159.6526	141.5268	140.6035	160.8131	185.4565
	MWV	0.0048	0.0051	0.0053	0.0051	0.0042	0.0035	207.0951	194.5260	187.7018	195.3697	236.5352	282.8614
	MSC	0.0045	0.0040	0.0034	0.0027	0.0017	0.0013	219.9446	249.0226	293.5823	373.9856	575.9705	789.0791
Model 3	**MS**	0.0045	0.0059	0.0072	0.0084	0.0097	0.0103	222.0101	169.1332	138.5445	119.2066	103.5990	97.5572
	MWV	0.0041	0.0047	0.0053	0.0060	0.0068	0.0072	241.8497	213.0288	188.2920	167.1794	147.4274	139.4953
	MSC	0.0038	0.0037	0.0034	0.0033	0.0034	0.0035	260.3624	273.6427	291.5962	299.0073	291.7919	284.0828

Table 13. Results of the p and ARL for the $\bar{X}$ control chart for contaminated gamma distribution

		n=5											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**MS**	0.0028	0.0061	0.0102	0.0142	0.0181	0.0214	355.1515	163.1162	97.5705	70.2336	55.1155	46.8305
	MWV	0.0030	0.0045	0.0067	0.0088	0.0109	0.0125	333.9456	220.1916	148.8117	113.6454	91.6137	79.7118
	MSC	0.0020	0.0028	0.0035	0.0040	0.0049	0.0058	499.8251	356.5952	285.2741	247.6106	204.3235	172.0282
Model 2	**MS**	0.0047	0.0176	0.0512	0.0816	0.0975	0.1029	212.1341	56.7038	19.5456	12.2533	10.2563	9.7216
	MWV	0.0048	0.0108	0.0269	0.0422	0.0497	0.0503	206.7910	92.3344	37.2108	23.7184	20.1147	19.8702
	MSC	0.0052	0.0074	0.0179	0.0278	0.0320	0.0308	192.2929	134.5605	55.8572	35.9442	31.2426	32.4377
Model 3	**MS**	0.0038	0.0393	0.0797	0.0878	0.0719	0.0507	261.1989	25.4221	12.5483	11.3872	13.9121	19.7187
	MWV	0.0053	0.1967	0.4482	0.5400	0.5612	0.5541	187.0977	5.0828	2.2313	1.8518	1.7817	1.8048
	MSC	0.0018	0.0077	0.0100	0.0066	0.0029	0.0011	557.8178	130.2456	100.1161	152.1005	340.6459	902.1200
		n=10											
		p values						ARL values					
	Method/k_3	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Model 1	**MS**	0.0037	0.0043	0.0054	0.0068	0.0076	0.0078	272.8662	229.9115	185.0995	147.9596	131.1785	127.6569
	MWV	0.0036	0.0037	0.0041	0.0046	0.0048	0.0047	277.4002	267.9313	246.2751	218.0549	207.3785	212.6800
	MSC	0.0033	0.0027	0.0020	0.0018	0.0017	0.0016	304.1178	375.0094	491.5696	549.9340	581.3953	634.8803
Model 2	**MS**	0.0044	0.0062	0.0072	0.0081	0.0085	0.0086	226.4288	160.8001	138.0129	123.4294	117.9830	116.7583
	MWV	0.0043	0.0051	0.0052	0.0053	0.0052	0.0049	231.2994	196.8233	193.8172	187.0592	192.1451	203.6784
	MSC	0.0042	0.0039	0.0030	0.0023	0.0020	0.0017	240.4135	253.6204	334.4705	435.1989	512.6891	578.8712
Model 3	**MS**	0.0045	0.0059	0.0074	0.0085	0.0091	0.0094	221.9953	170.9285	136.0304	117.8217	109.8901	106.0524
	MWV	0.0041	0.0046	0.0051	0.0055	0.0056	0.0054	243.1847	218.2263	194.9622	181.8579	179.3915	183.5266
	MSC	0.0038	0.0036	0.0030	0.0025	0.0021	0.0020	260.9535	281.1042	330.2946	407.3652	465.9181	505.3312

smaller p values. As the skewness increases, the p values of the Shewhart method increases too much and are quite higher than that of the other methods.

- Under non-normality, when skewness increases, ARL decreases and therefore p increases.
- When skewness increases, the Shewhart ARL values decreases significantly, while the ARL of the SC and MSC methods are not effected by the skewness. In particular, for $n \geq 5$ the ARL of the SC and MSC provide desirable values as skewness increases.

- The WV method produces better results than the Shewhart method, and the SC method produces better results than both the Shewhart and WV, while the skewness increases.
- The results for the Weibull, gamma and lognormal distributions are more or less similar with respect to the p and ARL values.
- When skewness increases, the p values of classic estimators increases for all methods. However, the p values of the MSC method decreases and reach the desirable value for gamma distribution for n=3,7 and 10.
- For the large sample size n=10; the p values of MWV and MSC methods decrease for Weibull and lognormal distributions when skewness increases, reaching the desirable value 0.0027. These modified methods work very well when skewness increases.
- In general, the ARL values of the SC and MSC methods are higher than those of the Shewhart and WV methods for all design schemes. Therefore, the SC and MSC methods have the best overall performance.

We investigate the effect of non-normality on estimated limits under contamination. We present the results of the simulation for $n = 5, 10$. Table 11, 12 and 13 give the results of the p and ARL for the $\bar{X}$ control chart for contaminated Weibull, lognormal and gamma distributions, respectively. The main points from these data are as follows:

- As skewness increases, the p decreases and so the ARL increases.
- The results for contaminated Weibull distribution are as follows: The MSC method has the best performance for Model 1 and Model 2 for n=5. However, its performance is deteriorated for Model 3, while skewness increases. For Model 1, MWV performs better than the others, while the performance of the MSC is deteriorated when $k_3 > 2$. For Model 2 and Model 3, the MSC has the lowest p values than the MWV and MS methods. All three methods can be used. However, where there are large outliers, the MSC method gives the desirable results when n=10. MWV can be used as an alternative.
- The results for contaminated lognormal distribution : For small sample sizes, the MS method performs better than the other modified methods when $k_3 > 1.5$. However, for large sample sizes, the MSC method has the best performance especially when $k_3 > 1$. Moreover the MWV method can be used as alternative to the MSC method for Model 2.
- The results for contaminated gamma distribution: When skewness increases, the p values of the MSC method decrease and so the ARL values increase. The MSC method performs better than the other method for n=5 and n=10. The MWV method produces the desirable results for n=10.
- For large sample size, the MSC method has the lowest p values and the highest ARL values for all skewed distributions. This modified method has the best performance.
- The MS and MWV methods can be used as alternatives to MSC method.

5. Conclusion

In this paper, three modified methods to construct the robust $\bar{X}$ control chart limits are suggested to monitor the skewed and contaminated process . We propose the MS, MWV and MSC methods using the simple and robust estimators, which are the trimmed mean and interquartile range. We have studied the effect of the estimators on control chart performance under non-normal distributed data for small and large sample sizes. The effect of outliers on the accuracy of conventional and robust estimators have been evaluated by root mean square errors via simulation. Contamination by extreme outliers result in a large increase in the $RMSE$ of the classic estimators, especially for the large samples (n = 10) and a much smaller increase in the $RMSE$s of the robust alternatives. The control chart constants for each method are obtained. To evaluate control chart performance, we obtain the p and ARL values of this control chart and the results are used to compare the methods. We analyse design schemes in which the Phase I and the Phase II data are non-contaminated and contaminated, respectively. The results can be summed up as follows: for non-contaminated data, as skewness increases, the p values of the classical estimators also increase and so the ARL values decrease in all methods. In contrast, the p values of the MSC method decrease and reach the desirable value (0.0027) for gamma distribution for n = 3,7 and 10. The WV method provides better results than the Shewhart method, and SC provides better results than both the Shewhart and WV methods, as skewness increases. The SC and MSC methods have the best performance out of all the design schemes analysed. For large sample sizes (n = 10), the MWV and MSC methods work very well for both Weibull and lognormal distributions, as skewness increases. Under these conditions, the use of these MWV and MSC methods is strongly recommended. The results can be summed up as follows: for contaminated data with large outliers and small sample sizes, the MS method performs better than the other modified methods when $k_3 > 1.5$. For large sample sizes, the MSC method has the best performance, especially when $k_3 > 1$. We strongly recommend use of the MSC method for large sample sizes, while the MWV can be used as an alternative. When the process distribution is in some neighbourhood of Weibull, lognormal or gamma, SC and MSC control charts have a p (i.e. probability of a false alarm) closer to 0.0027. Consequently, the proposed method for the robust $\bar{X}$ control chart can be a favourable substitute in process monitoring when the mean of a skewed population is contaminated in Phase I and Phase II.

References

[1] Abu-Shawiesh, M.O.A. *A Simple Robust Control Chart Based on MAD*, Journal of Mathematics and Statistics, 4(2), 102-107, 2008.

[2] Ahsanullah, M., Nevzorov, V.B., Shakil, M., *An Introduction to Order Statistics*, 246 p., Hardcover ISBN: 978-94-91216-82-4, A product of Atlantis Press, 2013.

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

[4] Castagliola, P. and Khoo, M.B.C. *A Synthetic Scaled Weighted Variance Control Chart for Monitoring the Process Mean of Skewed Populations*, Communications in Statistics: Simulation and Computation, 38, 1659 – 1674, 2009

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

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

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

[8] He, X., and Fung, W.K. *Method of medians for lifetime data with Weibull models*, Stat Med., 18, 1993-2009, 1999.

[9] Huber, P. J. *Robust Statistics*, John Wiley & Sons, 1981.

[10] Karagöz, D. and Hamurkaroğlu C. *Control Charts for Skewed Distributions: Weibull, Gamma, and Lognormal*, Metodoloski zvezki - Advances in Methodology and Statistics, 9(2), 95-106, 2012.

[11] Maronna, R. A., Martin, D. R. and Yohai, V. J. *Robust Statistics: Theory and Methods*, Wiley, 2006.

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

[13] Nelson, P.R. *Control Charts for Weibull Processes with Standards Given*, IEEE Transactions on Reliability, 28, 283-387, 1979.

[14] Schoonhoven, M. and Does, R.J.M.M. *The $\bar{X}$ control chart under non-normality*, Quality and Reliability Engineering International , 26: 167-176, 2010.

[15] Schoonhoven, M., Nazir, H.Z., Riaz, M. and Does, R.J.M.M. *Robust location estimations for the $\bar{X}$ control chart*, Journal of Quality Technology, 43: 363-379, 2010.

[16] Schoonhoven, M. and Does R.J.M.M. *A Robust $\bar{X}$ Control Chart*, Quality and Reliability Engineering International, 29: 951-970, 2013.

[17] Rocke, D.M. *$\hat{X}_Q$ and R_Q Charts: Robust Control Charts*, The Statistician, 41: 97-104, 1992.

[18] Whaley, D. L. *The Interquartile Range: Theory and Estimation*, Master Thesis Presented to the Faculty of the Department of Mathematics East Tennessee State University, 2005.

[19] Wilcox, R. *Introduction to Robust Estimation and Hypothesis Testing*, Academic press, University of Southern California, 2005.

[20] Wu, C., Zhao, Y., Wang, Z. *The median absolute deviations and their application to Shewhart $\bar{X}$ control charts*, Communication in Statistics -Simulation and Computation, 31: 425-442, 2002.