applied sciences

MDPI

Article

On the Subrange and Its Application to the R-Chart

En Xie [1], Yizhong Ma [1], Linhan Ouyang [2] and Chanseok Park [3,*]

[1] School of Economics & Management, Nanjing University of Science and Technology, Nanjing 210094, China; jinyu2016@njust.edu.cn (E.X.); yzma@njust.edu.cn (Y.M.)
[2] College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; ouyang@nuaa.edu.cn
[3] Department of Industrial Engineering, Pusan National University, Busan 46241, Korea
* Correspondence: cp@pusan.ac.kr

Abstract: The conventional sample range is widely used for the construction of an R-chart. In an R-chart, the sample range estimates the standard deviation, especially in the case of a small sample size. It is well known that the performance of the sample range degrades in the case of a large sample size. In this paper, we investigate the sample subrange as an alternative to the range. This subrange includes the range as a special case. We recognize that we can improve the performance of estimating the standard deviation by using the subrange, especially in the case of a large sample size. Note that the original sample range is biased. Thus, the correction factor is used to make it unbiased. Likewise, the original subrange is also biased. In this paper, we provide the correction factor for the subrange. To compare the sample subranges with different trims to the conventional sample range or the sample standard deviation, we provide the theoretical relative efficiency and its values, which can be used to select the best trim of the subrange with the sense of maximizing the relative efficiency. For a practical guideline, we also provide a simple formula for the best trim amount, which is obtained by the least-squares method. It is worth noting that the breakdown point of the conventional sample range is always zero, while that of the sample subrange increases proportionally to a trim amount. As an application of the proposed method, we illustrate how to incorporate it into the construction of the R-chart.

Keywords: subrange; distribution; unbiasing factors; relative efficiency; breakdown point

check for updates

Citation: Xie, E.; Ma, Y.; Ouyang, L.; Park, C. On the Subrange and Its Application to the R-Chart. *Appl. Sci.* **2021**, *11*, 11632. https://doi.org/10.3390/app112411632

Academic Editor: Pentti Nieminen

Received: 13 October 2021
Accepted: 2 December 2021
Published: 8 December 2021

1. Introduction

The control chart is a widely used and powerful graphical tool in quality control that is used to measure, monitor, and control a process over time. Usually, the control charts are in pairs. For example, an $\overline{X}$-chart monitors the average of the manufacturing process while an R-chart monitors the variation of the process [1]. Generally, there are two phases for constructing control charts [2]. In Phase-I the goal is to obtain reliable control limits from the process data. Then, in Phase-II monitor the process by comparing the statistical properties of the future observation to the control limits, which are achieved in Phase-I [3,4]. The performance of the control charts constructed in Phase-I will determine the performance of the results in Phase-II. Thus, the data quality in Phase-I plays an important role in statistical process control (SPC). However, for $\overline{X} - R$ charts, the sample mean and range are susceptible to the outliers, which is also called data contamination. Thus, the conventional control charts may be invalidated in the case of data contamination. To solve this problem, we use a robust estimator to construct control charts in Phase-I.

Robust statistics can provide good performance when there is the presence of outliers and departures from the model assumption. In robust design, when the collected data are contaminated, the robust estimators are employed to reduce or even avoid the influence of outliers on the results [5–9]. In statistical process control, Park et al. [10] proposed the use of robust scale estimators (e.g., median absolute deviation (MAD)) [11] and Shamos [12]) used an alternative to the standard deviation to construct the control charts in Phase–I when the data is contaminated. Shaweish and Saeed modified the S-control chart based on the decile mean standard deviation, which possessed a better performance when the sample size was small in most realistic applications [13]. When the process data in Phase-I diverge from normality and outliers are present, the robust S-chart based on robust scale estimators is more preferable than a traditional S-chart [10,13]. The reweighted robust standard deviation estimator is adopted to modify the S-chart by Mutlu and Alakent [14]. To improve robustness and be more efficient, in the case of the presence of outliers in the data set, the trimmed mean of mean, standard deviation, and range are adopted to construct the control charts [15–17]. Shaweish et al. [18] and Riaz and Saghir [19] pointed out that the control chart based on the robust scale estimator provides good performance not only without outliers, but also within outliers in the data set. Cumulative sum (CUSUM) control charts for special causes are also under the assumption of normal distribution. Given this, many attentions are paid to the robust CUSUM chart when the ideal assumption is violated [20–22]. Considering the multivariate non-normality problems, the distribution-free method [23] and Bayesian seemingly unrelated regression model [24] can give us some inspirations.

The trimmed mean (also called truncated mean) is the mean of the part after removing the x% largest and smallest observations [25]. For example, to reduce and mitigate the influence of bias in subjective judgements, the highest and lowest scores are discarded and the middle three are summed for a five-judge panel [26,27] for many Olympics games (e.g., diving and gymnastics etc.).

However, there is seldom research about the R-chart under data contamination in the literature. This motivates us to develop a robust method for control charts when the data is contaminated. In this paper, we propose a robust scale estimator subrange as an alternative to the sample range—as an estimator of the standard deviation to improve the control chart in the case of data contamination. The calculation of subrange is very similar to the trimmed mean. However, the subrange is not unbiased.

As we know, order statistics are widely used in nonparametric inferences and robust procedures. Order statistics theories can also be used to compute simple statistics such as median, the empirical cumulative distribution function (cdf), and the sample range [28–30]. The range of a sample dispersion tells us how far apart the largest and smallest values are. The traditional sample range is the distance between the largest and the smallest observation values from a sample. However, estimating the standard deviation by means of the sample range is not unbiased. Thus, it is necessary to use the unbiasing factor d_2 for the original sample range to estimate the standard deviation properly, which is a function of the sample size [31]. The method is widely seen in quality control literature [32,33].

Based on the joint pdf, the distribution of the sample range has been extensively studied [34–36]. However, in the earlier literature, many papers focus on the sample range. If there is an outlier in a sample, obviously the sample range will be easily influenced by even a single outlier. Given this, we can say that the sample range clearly has a zero-breakdown point, which is a popular measure for the robustness property of a statistical estimator. The breakdown point of a statistical estimator is the fraction of data that can be given arbitrary values without making the estimator arbitrarily bad [10,36]. The sample range is almost as efficient as the sample standard deviation for a small sample under the normal distribution, which is generally assumed in many engineering applications. Howefer, Patnaik [37] pointed out that when the sample size is larger than 20, even a small departure from normality in the tail of the parental distribution has a considerable effect on the distribution of the sample range. In order to make a good performance of the sample range for a large sample, Cadwell [38] proposed a method of dividing the sample into a few subgroups and used the average of the range of these subgroups.

Since the sample subrange depends on the trim amount, there may be several sub-ranges to consider under different trims. Therefore, natural questions arise: which of the subranges is preferred? Then, by what criteria, shall we make the choice? Serfling [39] and

Lehmann [40] proposed that the smaller the variance in its sampling distribution, the more "efficient" is that estimator. The relative efficiency (RE) value is considered in order to decide how to choose the optimal trim amount of the subrange by comparing the variance of the sample. Firstly, we give the distribution of the sample subrange and provide the unbiasing factors, which depend on the sample size and the number of trims to make the estimator unbiased. Next, we provide a criterion that is based on the RE value to choose the optimal trims in the sample. Lastly, we consider the breakdown point of the subrange. As we know, the inter-quartile range (IQR) is widely used to estimate the standard deviation when the sample has an outlier or extreme values. The proposed subrange includes the IQR as a special case. As well, the subrange is robust to outliers. The contribution of this article is as follows.

We give the distribution of the sample subrange and provide the unbiasing factor for the subrange of a sample through a Monte Carlo simulation. We assume the data are a normal distribution in order to calculate the unbiasing factor and the distribution for the subrange because the commonly used control charts are under the independent and normality assumption [1,12,41,42].

The RE values are calculated by using two different kinds of baseline estimators, from which we discover that the results coincide with choosing the best performance of subranges. Some interesting points are found through comparing the breakdown points of a range and a subrange, by which we conclude that the subrange has a positive breakdown point of k/n, and the asymptotic breakdown point of the subrange can be increased to around $1/2$ in theory.

In previous studies, researchers paid little attention to the sample range due to its limitations, such as its narrow application and susceptibility to outliers. In this paper, we improve the range by extending its application and improving its breakdown point. Then we use the subrange to construct a control chart for monitoring the changes in the progress when the data are contaminated. Through the Monte Carlo simulation, we investigate the properties of the subrange. We offer unbiasing factors and the relative efficiency values for the sample size of 50 or less, firstly.

In this paper, the sample subrange is proposed as an alternative to the sample range as an estimate of the standard deviation. A correction factor for the subrange is provided in order to get unbiasedness. An application of the proposed method is illustrated by incorporating it into the construction of the R-chart. In Section 2, calculations are carried out for a random sample from a standard normal distribution, and then from a non-standard normal distribution. In Section 3, the central limit theorem is used to construct control charts by adopting the subrange. Some discussions about the relative efficiency value and the breakdown point of the subrange are provided in Section 4. Additionally, concluding remarks are given in Section 4.

2. The Distribution of the Subrange

Let $X_1, X_2, \cdots X_n$ be a random sample with continuous cdf $F(x)$ and pdf $f(x)$. Let $X_{(1)}, X_{(2)}, \cdots X_{(n)}$ be the order statistics of a random sample. Then the joint pdf of $X_{(i)}$ and $X_{(j)}$ for $1 \leq i < j \leq n$ is given by

$$f_{(i,j)}(u,v) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f(u)f(v)[F(u)]^{i-1}[F(v)-F(u)]^{j-i-1}[1-F(v)]^{n-j}, \quad (1)$$

for $-\infty < u < v < \infty$. For more details on the above result, refer to Casella and Berger [28], and Hogg et al. [30].

We consider the symmetrically trimmed subrange. That is, we exclude the k smallest and k largest values. Thus, the symmetrically trimmed subrange is defined as $R_{[k]} = X_{(n-k)} - X_{(k+1)}$ where $k = 0, 1, 2, \cdots, \lfloor n/2 \rfloor - 1$ and $n \geq 2$. Here $\lfloor x \rfloor$ is the largest integer less than or equal to x. Note that $R_{[k]}$ becomes the regular range with $k = 0$.

We assume that a random sample is from a normal distribution with mean μ and standard deviation σ. It is well known that the range is not unbiased to estimate σ. By dividing R by the unbiasing factor d_2, we can easily make this R unbiased. The values of d_2 are provided in the quality control literature—for more details refer to Shewhart [32] and Oakland [33].

In this paper, we will provide the unbiasing factor which makes $R_{[k]}$ unbiased for σ. We denote this factor by $d_2(n,k)$. It should be noted that the factor d_2 depends on the sample size n and the number of trims k in the sample.

For notational convenience, we denote the joint pdf of $X_{(k+1)}$ and $X_{(n-k)}$ by $f(u,v)$ Then we have

$$f_k(u,v) = \frac{n!}{(k!)^2(n-2k-2)!} f(u)f(v)[F(u)]^k[1-F(v)]^k[F(v)-F(u)]^{n-2k-2}.$$

Let $Z_1, Z_2, \cdots Z_n$ be a random sample from a standard normal distribution with pdf $\phi(z)$ and cdf $\Phi(z)$. For notational convenience, we denote $W_1 = Z_{(k+1)}$ and $W_2 = Z_{(n-k)}$. Using (1), we have the joint pdf of W_1 and W_2.

$$f_k(w_1,w_2) = \frac{n!}{(k!)^2(n-2k-2)!} \phi(w_1)\phi(w_2)[\Phi(w_1)]^k[1-\Phi(w_2)]^k[\Phi(w_2)-\Phi(w_1)]^{n-2k-2}.$$

The goal is to derive the distribution of the subrange, $W_2 - W_1 = Z_{(n-k)} - Z_{(k+1)}$ where the Z_i are from the standard normal distribution, $N(0,1)$. Next, we consider the new random variables given by $Y_1 = W_1$ and $Y_2 = W_2 - W_1$. Notice that the random variable Y_2 is the subrange. The inverse transformations are easily obtained by $w_1 = y_1$ and $w_1 = y_2 + y_1$. Then, using the bivariate transformations, the joint pdf of Y_1 and Y_2, denoted by $g_k(y_1,y_2)$ is given by

$$g_k(y_1,y_2) = \frac{n!}{(k!)^2(n-2k-2)!} \phi(y_1)\phi(y_1+y_2)\Phi(y_1)^k[1-\Phi(y_1+y_2)]^k[\Phi(y_1+y_2)-\Phi(y_1)]^{n-2k-2}|J|,$$

where $-\infty < y_1 < \infty, y_2 > 0$ and J is the determinant of the Jacobian matrix given by

$$J = \det\begin{bmatrix} \frac{\partial w_1}{\partial y_1} & \frac{\partial w_1}{\partial y_2} \\ \frac{\partial w_2}{\partial y_1} & \frac{\partial w_2}{\partial y_2} \end{bmatrix} = \det\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = 1$$

For more details, see Casella and Berger [28]. Then the pdf $Y_2 = R_{[k]} = Z_{(n-k)} - Z_{(k+1)}$ is just the marginal pdf of Y_2, which is given by

$$g_{[k]}(y_2) = \int_{-\infty}^{\infty} \frac{n!}{(k!)^2(n-2k-2)!} \phi(y_1)\phi(y_1+y_2)\Phi(y_1)^k[1-\Phi(y_1+y_2)]^k[\Phi(y_1+y_2)-\Phi(y_1)]^{n-2k-2} dy_1 \qquad (2)$$

where $k = 0, 1, 2, \cdots, \lfloor n/2 \rfloor - 1$ and $n \geq 2$ again.

2.1. The Unbiasing Factors for the Subrange

Using the pdf of the range in (2), we can obtain the l-th moment of the range by calculating the expectation as follows:

$$E\left(Y_2^l\right) = \int_0^{\infty} y_2^l g_{[k]}(y_2) dy_2$$

We denote

$$d_2(n,k) = E(Y_2). \qquad (3)$$

Let X_i be a sample from the normal distribution $N(\mu, \sigma^2)$. Since σ is a scale parameter, we have $X_i = \sigma Z_i + \mu$, where Z_i is a sample from the standard normal distribution $N(0,1)$.

Then we have $E\left(R_{[k]}\right) = E\left(X_{(n-k)} - X_{(k+1)}\right) = d_2(n,k)\sigma$. Thus, we can obtain the unbiased scale estimator of σ using the subrange which is given by

$$\widehat{\sigma}_{[k]} = \frac{R_{[k]}}{d_2(n,k)}. \tag{4}$$

Note that the unbiasing factor, d_2 in the quality control literature is obtained by $d_2 = d_2(n,0)$, so that $\widehat{\sigma}_0 = R/d_2$. Next, we denote

$$d_3(n,k) = \mathrm{SD}(Y_2) = \sqrt{E(Y_2^2) - E(Y_2)^2}. \tag{5}$$

Then, we have $\mathrm{SD}\left(R_{[k]}\right) = d_3(n,k)\sigma$. Also, the unbiasing factor d_3 in the quality control literature is also given by $d_3 = d_3(n,0)$.

In Tables A1 and A2, we provide the values of $d_2(n,k)$ and $d_3(n,k)$, respectively. For brevity, we provide the values for $n = 2,3,\cdots 50$ and $k = 0,1,\cdots,9$ in the tables. The R language program is provided in the online supplement. Subrange. Available online: https://github.com/jin-yuyu/subrange.git (accessed on 1 December 2021).

2.2. The Relative Efficiency of the Subrange

We proposed the R-chart based on the subrange. Given this, a natural question is which subrange should be selected? This is essentially the same as how to choose k. We suggest the choice of k based on the RE value. In the statistics literature, the RE is defined as

$$\mathrm{RE}\left(\widehat{\theta}_2 \middle| \widehat{\theta}_1\right) = \frac{\mathrm{Var}\left(\widehat{\theta}_1\right)}{\mathrm{Var}\left(\widehat{\theta}_2\right)} \times 100\%, \tag{6}$$

where $\widehat{\theta}$ is often a reference or baseline estimator. For more details, see Serfling [39] and Lehmann [40]. We consider the RE of the unbiased scale estimator $\widehat{\sigma}_{[k]}$ using $\widehat{\sigma}_{[0]}$ as a baseline estimator, which is given by

$$\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \widehat{\sigma}_{[0]}\right) = \frac{\mathrm{Var}\left(\widehat{\sigma}_{[0]}\right)}{\mathrm{Var}\left(\widehat{\sigma}_{[k]}\right)} \times 100\% = \left(d_3(n,0)^2/d_2(n,0)^2\right)/\left(d_3(n,k)^2/d_2(n,k)^2\right) \times 100\%. \tag{7}$$

We calculated the values of $\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \widehat{\sigma}_{[0]}\right)$ in Table A3 for $n = 2,3,\cdots 50$ and $k = 0,1,\cdots,9$. It is easily seen that the choice of $k = 0$ gives the best performance for $n = 2,3,\cdots,17$, and the choice of $k = 1$ does for $n = 18,19,\cdots,31$ etc. In Table 1, we also summarize the best choice of k which provides the maximum RE value. We also obtain the simple regression line using the least-squares method. The fitted line equation is given by

$$k = 0.65806 + 0.06926 \cdot n.$$

Table 1. The number of trimmings, k^*, with the best relative efficiency.

$2 \le n \le 17$	$18 \le n \le 31$	$32 \le n \le 45$	$46 \le n \le 60$	$61 \le n \le 74$
$k^* = 0$	$k^* = 1$	$k^* = 2$	$k^* = 3$	$k^* = 4$
$75 \le n \le 88$	$89 \le n \le 103$	$104 \le n \le 118$	$119 \le n \le 132$	$133 \le n \le 146$
$k^* = 5$	$k^* = 6$	$k^* = 7$	$k^* = 8$	$k^* = 9$
$147 \le n \le 162$	$163 \le n \le 175$	$176 \le n \le 188$	$189 \le n \le 203$	$204 \le n \le 217$
$k^* = 10$	$k^* = 11$	$k^* = 12$	$k^* = 13$	$k^* = 14$

Thus, when the sample size n is very large, the value of k can be approximately selected using the above-fitted value. For example, when $n = 500$, we have $k \approx 34$.

Using the unbiased estimator $S_n/c_4(n)$ of σ as a baseline, we also calculated the RE of the unbiased scale estimator $\widehat{\sigma}_{[k]}$ again. Here, $c_4(n)$ is given by

$$c_4(n) = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2} - \frac{1}{2}\right)}.$$

It should be noted that $E\left[S_n^2\right] = \sigma^2$ and $E[S_n] = c_4(n)\sigma$. Thus, we have $\mathrm{Var}(S_n/c_4(n)) = \left\{1/c_4(n)^2 - 1\right\}\sigma^2$ and

$$\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \frac{S_n}{c_4(n)}\right) = \frac{\mathrm{Var}\left(\frac{S_n}{c_4(n)}\right)}{\mathrm{Var}\left(\widehat{\sigma}_{[k]}\right)} \times 100\% = \frac{1/c_4(n)^2 - 1}{d_3(n,k)^2/d_2(n,k)^2} \times 100\%.$$

We also calculated the values of $\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \frac{S_n}{c_4(n)}\right)$ in Table A3 for $n = 2, 3, \cdots 50$ and $k = 0, 1, \cdots, 9$. It should be noted that $\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \frac{S_n}{c_4(n)}\right) = 100\%$ when $n = 2$. We investigated this specific case in more detail as follows. When $n = 2$, k should be zero. Thus, in this case, it is immediately from (2) that the pdf of $Y_2 = R_{[0]} = Z_2 - Z_1$ is given by

$$g_{[0]} = \int_{-\infty}^{\infty} 2\phi(y_1)\phi(y_1+y_2)dy_1 = 2\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_1+y_2)^2}{2}\right) dy_1 = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{y_2^2}{4}\right) \int_{-\infty}^{\infty} g^*(y_1)dy_1.$$

where $g^*(y_1) = \frac{1}{\sqrt{2\pi}\cdot 1/\sqrt{2}} \exp\left(-\frac{1}{2} \cdot \frac{(y_1+y_2/2)^2}{1/2}\right)$.

Science $g^*(y_1)$ is the pdf of the normal distribution with $N\left(-y_2/2, 1/\sqrt{2}\right)$, we have

$$g_{[0]} = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{y_2^2}{4}\right).$$

which results in

$$d_2(2,0) = \int_0^{\infty} y_2 g_{[0]} dy_2 = \int_0^{\infty} \frac{1}{\sqrt{\pi}} y_2 \exp\left(-\frac{y_2^2}{4}\right) dy_2 = \frac{1}{\sqrt{\pi}} \left[-2\exp\left(-\frac{y_2^2}{4}\right)\right]_0^{\infty} = \frac{2}{\sqrt{\pi}}.$$

Then we have

$$\widehat{\sigma}_{[0]} = \frac{R_{[0]}}{d_2(2,0)} = \frac{\sqrt{\pi}}{2}\left(X_{(2)} - X_{(1)}\right). \tag{8}$$

On the other hand, we have

$$c_4(2) = \sqrt{2}\frac{\Gamma(2)}{\Gamma(1/2)} = \frac{\sqrt{2}}{\sqrt{\pi}}$$

Since $S_2^2 = \frac{1}{2-1}\sum_{i=1}^{2}\left(X_i - \overline{X}_2\right)^2 = (X_1 - X_2)^2/2$, we have

$$\frac{S_2}{c_4(2)} = \frac{|X_1 - X_2|/\sqrt{2}}{\sqrt{2}/\sqrt{\pi}} = \frac{\sqrt{\pi}}{2}|X_1 - X_2| = \frac{\sqrt{\pi}}{2}\left(X_{(2)} - X_{(1)}\right) \tag{9}$$

It is immediate from (8) and (9) that

$$\widehat{\sigma}_{[0]} = \frac{S_2}{c_4(2)}$$

Thus, as aforementioned, we have $\mathrm{RE}\left(\widehat{\sigma}_{[k]}\middle|\frac{S_n}{c_4(n)}\right) = 100\%$ when $n = 2$ (also $k = 0$ in this case, as mentioned earlier).

2.3. The Breakdown Points

Another way of choosing k is considering the finite-sample breakdown point, denoted by ε_n. This is the maximum proportion of arbitrarily extreme observations where an estimator results in a reasonable value. For more detail see [6,10]. This finite-sample breakdown point is generally a function of sample size n. The proposed estimator, based on the subrange, has additional merit because of a positive breakdown point. It is clear that $\widehat{\sigma}_{[k]}$ has a positive breakdown point of k/n for $k > 0$,$\hat{\sigma}_{[0]}$ has a zero one. Thus, we can choose k based on the value k/n.

It should be noted that the maximum attainable value of the asymptotic breakdown point is given as follows. Since the maximum value of k is $\lfloor n/2 \rfloor - 1$, the maximum finite-sample breakdown point available is given by

$$\varepsilon_n = \frac{\lfloor n/2 \rfloor - 1}{n}.$$

Since $\lfloor n/2 \rfloor$ can be expressed by $\lfloor n/2 \rfloor = n/2 - \delta$ with $0 \le \delta < 1$, we have

$$\varepsilon_n = \frac{1}{2} - \frac{\delta + 1}{n}.$$

Thus, taking the limit of ε_n as $n \to \infty$, we have the asymptotic breakdown point

$$\varepsilon = \lim_{n\to\infty} \varepsilon_n = \frac{1}{2},$$

which is the maximum attainable asymptotic breakdown point.

It is likely that the RE value tends to decrease as the finite-sample breakdown point increases. Thus, when we choose k, we need to consider the RE value and the finite-sample breakdown point as well.

3. The Construction of Control Charts

In this section, we provide the method for incorporating the proposed methods into constructing the control charts. We use the proposed subrange as an alternative of the range and standard deviation when there are outliers in the data of Phase I. We provided the $\overline{X}$-chart first, and then the R-chart.

For the case of the $\overline{X}$-chart, it is immediately from the central limit theorem that we have

$$\frac{\overline{X} - E(\overline{X})}{\mathrm{SE}(\overline{X})} \sim N(0,1),$$

where $\overline{X}$-is the sample mean from a sample of size n, and $\mathrm{SE}(\cdot)$ is the standard deviation from a sample of size n. Solving the below for $\overline{X}$

$$\left(\overline{X} - E(\overline{X})\right)/\mathrm{SE}(\overline{X}) = \pm 3,$$

we can construct the $\mathrm{CL} \pm 3 \cdot \mathrm{SE}$ control limits.

$$E(\overline{X}) \pm 3 \cdot \mathrm{SE}(\overline{X}) = \mu \pm \frac{3\sigma}{\sqrt{n}}.$$

Then we have UCL $= \mu + \frac{3\sigma}{\sqrt{n}}$, CL $= \mu$, and LCL $= \mu - \frac{3\sigma}{\sqrt{n}}$. Since μ and σ are unknown in practice, we need to estimate them. Suppose that there are m samples of size n. The most widely used estimator of μ is $\overline{\overline{X}}$ where

$$\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \cdots + \overline{X}_m}{m}.$$

The scale σ can be estimated by using the subrange. Let $R_{[k],i}$ be the subrange with k trims from the ith sample. Then, as shown in (4), $R_{[k],i}/d_2(n,k)$ is an unbiased scale estimator of σ. Thus, it is quite reasonable to use $\overline{R}_{[k]}/d_2(n,k)$ for the scale estimator, where

$$\overline{R}_{[k]} = \frac{R_{[k],1} + R_{[k],2} + \cdots + R_{[k],m}}{m}.$$

With the estimators of μ and σ, we have the following control limits:

$$\begin{aligned} \text{UCL} &= \overline{\overline{X}} + \tfrac{3\overline{R}_{[k]}}{d_2(n,k)\sqrt{n}} \\ \text{CL} &= \overline{\overline{X}} \\ \text{LCL} &= \overline{\overline{X}} - \tfrac{3\overline{R}_{[k]}}{d_2(n,k)\sqrt{n}}. \end{aligned}$$

It should be noted that when $k = 0$, we have $\overline{R}_{[k]} = \overline{R}$ and $d_2(n,k) = d_2$. The above control limits then become

$$\begin{aligned} \text{UCL} &= \overline{\overline{X}} + \tfrac{3\overline{R}}{d_2\sqrt{n}} \\ \text{CL} &= \overline{\overline{X}} \\ \text{LCL} &= \overline{\overline{X}} - \tfrac{3\overline{R}}{d_2\sqrt{n}}. \end{aligned}$$

which is the traditional $\overline{X}$-chart provided in the quality control literature.

For the case of the R-chart, we briefly reviewed the conventional R-chart and proposed a new $R_{[l]}$-chart. Solving the below

$$\frac{R - E(R)}{\text{SE}(R)} = \pm 3$$

for R, we can construct the CL $\pm 3 \cdot$ SE control limits given by

$$E(R) \pm 3 \cdot \text{SE}(R) = d_2\sigma \pm 3d_3\sigma.$$

Then we have UCL $= d_2\sigma + 3d_3\sigma$, CL $= d_2\sigma$ and LCL $= d_2\sigma - 3d_3\sigma$.

Analogous to the above construction, we can calculate the $R_{[l]}$-chart with the limits given by

$$E\left(R_{[l]}\right) \pm 3\text{SE} \cdot \left(R_{[l]}\right) = d_2(n,l)\sigma \pm 3d_3(n,l)\sigma$$

In practice, the scale σ is unknown. Thus, we need to estimate σ using $\overline{R}_{[k]}/d_2(n,k)$, as above. We then have

$$\begin{aligned} \text{UCL} &= d_2(n,l)\tfrac{\overline{R}_{[k]}}{d_2(n,k)} + 3d_3(n,l)\tfrac{\overline{R}_{[k]}}{d_2(n,k)} \\ \text{CL} &= d_2(n,l)\tfrac{\overline{R}_{[k]}}{d_2(n,k)} \\ \text{LCL} &= d_2(n,l)\tfrac{\overline{R}_{[k]}}{d_2(n,k)} - 3d_3(n,l)\tfrac{\overline{R}_{[k]}}{d_2(n,k)}. \end{aligned}$$

When $l = 0$ and $k = 0$, we have the following limits which are essentially the same as the conventional R-chart

$$\begin{aligned} \mathrm{UCL} &= \overline{R}_{[0]} + 3d_3(n,0)\frac{\overline{R}_{[0]}}{d_2(n,0)} \\ \mathrm{CL} &= \overline{R}_{[0]} \\ \mathrm{LCL} &= \overline{R}_{[0]} - 3d_3(n,0)\frac{\overline{R}_{[0]}}{d_2(n,0)} \end{aligned}$$

Also, if we select $k = l$, we can simplify the above $\overline{R}_{[l]}$-chart

$$\begin{aligned} \mathrm{UCL} &= \overline{R}_{[l]} + 3d_3(n,l)\frac{\overline{R}_{[l]}}{d_2(n,l)} \\ \mathrm{CL} &= \overline{R}_{[l]} \\ \mathrm{LCL} &= \overline{R}_{[l]} - 3d_3(n,l)\frac{\overline{R}_{[l]}}{d_2(n,l)} \end{aligned}$$

Then a natural question is how to choose k for the $R_{[l]}$-chart? The following proposition provides how to select k for the best performance

Proposition 1. *Let $X_1, X_2, \cdots X_n$ be a random sample from the normal distribution with scale σ, we then have*

$$\mathrm{Var}\Big(\overline{R}_{[k]}/d_2(n,k)\Big) \geq \mathrm{Var}\Big(\overline{R}_{[k^*]}/d_2(n,k^*)\Big),$$

where k^* is the choice of k, $\mathrm{RE}\Big(\hat{\sigma}_{[k]}\Big|\hat{\sigma}_{[0]}\Big)$ in (7), for $k = 0, 1, \cdots, \lfloor n/2 \rfloor - 1$.

Proof of Proposition 1. It is immediately from (5) that we can obtain $\mathrm{Var}\Big(\overline{R}_{[k]}/d_2(n,k)\Big) = d_3(n,k)^2\sigma^2$ and $\mathrm{Var}\Big(\overline{R}_{[k^*]}/d_2(n,k^*)\Big) = d_3(n,k^*)^2\sigma^2$.

Using the above, $\mathrm{Var}\Big(\overline{R}_{[k]}/d_2(n,k)\Big)/\mathrm{Var}\Big(\overline{R}_{[k^*]}/d_2(n,k^*)\Big)$ can be formulated as

$$\frac{d_3(n,0)^2/d_2(n,0)^2}{d_3(n,k^*)^2/d_2(n,k^*)^2} \Big/ \frac{d_3(n,0)^2/d_2(n,0)^2}{d_3(n,k)^2/d_2(n,k)^2}.$$

It is easily seen from (7) that k^* provides the maximum RE value so that we have

$$\mathrm{Var}\Big(\overline{R}_{[k]}/d_2(n,k)\Big)/\mathrm{Var}\Big(\overline{R}_{[k^*]}/d_2(n,k^*)\Big) = \mathrm{RE}\Big(\widehat{\sigma}_{[k^*]}\Big|\widehat{\sigma}_{[0]}\Big)/\mathrm{RE}\Big(\widehat{\sigma}_{[k]}\Big|\widehat{\sigma}_{[0]}\Big) \geq 1.$$

Thus, we get $\mathrm{Var}\Big(\overline{R}_{[k]}/d_2(n,k)\Big) \geq \mathrm{Var}\Big(\overline{R}_{[k^*]}/d_2(n,k^*)\Big)$, which completes the proof. □

4. Discussion

As mentioned above, the choice of the trim k^* is based on the relative efficiency under different sample sizes. From Table 1, we can see that when the sample size is $18 \leq n \leq 31$, the trim $k^* = 1$ is the best choice. This means that the subrange is robust in the case of an outlier in the data, under the trim k^* when its set to be 1. Given this, the subrange performs better than the traditional range. However, we also can see that when the sample size is $18 \leq n \leq 31$, the largest breakdown point is $1/18 \approx 5.6\%$. This breakdown point of the subrange is much smaller than the MAD is. The proposed subrange is very useful for constructing a control chart when the data are contaminated in Phase-I. Although the breakdown point of the subrange is not satisfactory, it is robust under data contamination.

5. Conclusions

In this paper, we proposed a method of estimating the scale estimator using the subrange. Using this, we can construct the $\overline{X}$ and R charts, which are widely used in the manufacturing process. By using the proposed control charts based on the subrange, we can gain statistical efficiency along with a robustness property.

We also provided the method of choosing the trim amount of the subrange, in the sense of gaining more efficiency compared to the conventional range. For the R program used in the paper, one can refer to the URL. Subrange. Available online: https://github.com/jin-yuyu/subrange.git (accessed on 1 December 2021). The simulation results are shown in the Appendix A. Tables A1 and A2 give the unbiasing factors for the sample size of 50 or less. Tables A3 and A4 give the information on relative efficiency values for the sample size of 50 or less—in terms of the relative efficiency values.

Note that the subrange has a positive breakdown point, while the breakdown point of the conventional range is always zero. The proposed subrange in this paper is robust. It can also be used when the sample size is larger (I.e., greater than 25).

However, for the $\overline{X}$ chart, the centerline is the sample mean, which is also the sensitivity to outliers. For future research, we suggest using a robust location estimator (e.g., median and Hodges-Lehmann [43]) to construct the $\overline{X}$-chart. As discussed in the last section, selecting suitable or proper criteria to improve the breakdown may be an interesting topic.

Author Contributions: Conceptualization, E.X. and C.P.; methodology and software, C.P.; formal analysis, L.O.; funding acquisition, Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: The work of was supported by the National Natural Science Foundation of China, grants number NSFC-71871119/NSFC-71811540414/NSFC-71702072, the National Research Foundation of Korea (NRF), grant funded by the Korea government (No. NRF-2017R1A2B4004169).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Unbiasing factor, $d_2(n,k)$.

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
2	1.1284									
3	1.6926									
4	2.0588	0.5940								
5	2.3259	0.9900								
6	2.5344	1.2835	0.4031							
7	2.7044	1.5147	0.7054							
8	2.8472	1.7044	0.9456	0.305						
9	2.9700	1.8646	1.1439	0.5491						
10	3.0775	2.0027	1.3121	0.7515	0.2453					
11	3.1729	2.1238	1.4577	0.924	0.4498					
12	3.2585	2.2315	1.5857	1.0737	0.6245	0.2052				
13	3.3360	2.3282	1.6997	1.2057	0.7767	0.3810				
14	3.4068	2.4158	1.8023	1.3235	0.9111	0.5346	0.1763			
15	3.4718	2.4959	1.8954	1.4298	1.0314	0.6706	0.3306			
16	3.532	2.5695	1.9805	1.5263	1.1400	0.7924	0.4675	0.1546		
17	3.5879	2.6376	2.0589	1.6148	1.2389	0.9027	0.5904	0.2920		
18	3.6401	2.7008	2.1315	1.6963	1.3296	1.0032	0.7017	0.4155	0.1376	
19	3.6890	2.7599	2.1989	1.7717	1.4132	1.0954	0.8033	0.5275	0.2614	
20	3.7349	2.8152	2.2619	1.8420	1.4908	1.1806	0.8967	0.6299	0.3739	0.1240
21	3.7783	2.8672	2.3209	1.9076	1.5630	1.2596	0.983	0.7241	0.4768	0.2367
22	3.8194	2.9163	2.3765	1.9692	1.6305	1.3333	1.0631	0.8112	0.5716	0.3399
23	3.8583	2.9627	2.4289	2.0271	1.6939	1.4023	1.1379	0.8922	0.6593	0.4351

Table A1. *Cont.*

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
24	3.8953	3.0068	2.4785	2.0818	1.7536	1.4671	1.2080	0.9678	0.7409	0.5233
25	3.9306	3.0486	2.5255	2.1336	1.8100	1.5281	1.2738	1.0387	0.8172	0.6054
26	3.9643	3.0885	2.5702	2.1827	1.8634	1.5858	1.3359	1.1053	0.8887	0.6821
27	3.9965	3.1265	2.6128	2.2294	1.9141	1.6404	1.3945	1.1682	0.9560	0.7541
28	4.0274	3.1629	2.6535	2.2739	1.9623	1.6923	1.4502	1.2277	1.0195	0.8219
29	4.0570	3.1978	2.6924	2.3164	2.0083	1.7417	1.5030	1.2841	1.0796	0.8860
30	4.0855	3.2312	2.7296	2.3571	2.0522	1.7888	1.5533	1.3377	1.1367	0.9466
31	4.1129	3.2633	2.7654	2.3961	2.0942	1.8338	1.6013	1.3888	1.1909	1.0041
32	4.1393	3.2942	2.7997	2.4334	2.1344	1.8768	1.6472	1.4375	1.2426	1.0589
33	4.1648	3.324	2.8327	2.4694	2.173	1.9181	1.6911	1.4841	1.2919	1.1110
34	4.1894	3.3527	2.8646	2.5039	2.2102	1.9577	1.7332	1.5287	1.3391	1.1609
35	4.2132	3.3805	2.8952	2.5372	2.2459	1.9958	1.7736	1.5715	1.3843	1.2085
36	4.2362	3.4072	2.9249	2.5693	2.2803	2.0325	1.8125	1.6126	1.4276	1.2542
37	4.2586	3.4332	2.9535	2.6003	2.3135	2.0678	1.8499	1.6521	1.4693	1.2980
38	4.2802	3.4583	2.9812	2.6303	2.3456	2.1019	1.8860	1.6902	1.5094	1.3402
39	4.3012	3.4826	3.008	2.6593	2.3766	2.1348	1.9208	1.7269	1.548	1.3807
40	4.3216	3.5062	3.0340	2.6874	2.4066	2.1666	1.9544	1.7623	1.5852	1.4198
41	4.3414	3.5292	3.0593	2.7146	2.4356	2.1974	1.9870	1.7965	1.6211	1.4574
42	4.3606	3.5514	3.0838	2.7410	2.4638	2.2273	2.0184	1.8296	1.6558	1.4938
43	4.3794	3.5731	3.1075	2.7666	2.4911	2.2562	2.0489	1.8616	1.6894	1.5290
44	4.3976	3.5941	3.1307	2.7915	2.5176	2.2843	2.0785	1.8927	1.7219	1.5630
45	4.4154	3.6147	3.1532	2.8157	2.5434	2.3115	2.1072	1.9228	1.7535	1.5959
46	4.4328	3.6346	3.1751	2.8392	2.5684	2.3380	2.1350	1.9520	1.7840	1.6278
47	4.4497	3.6541	3.1964	2.8622	2.5928	2.3637	2.1621	1.9804	1.8137	1.6588
48	4.4662	3.6731	3.2172	2.8845	2.6165	2.3888	2.1884	2.0079	1.8425	1.6888
49	4.4824	3.6916	3.2375	2.9062	2.6397	2.4132	2.2140	2.0347	1.8705	1.7180
50	4.4981	3.7097	3.2573	2.9275	2.6622	2.4369	2.2390	2.0608	1.8977	1.7464

Table A2. Unbiasing factor, $d_3(n,k)$.

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
2	0.8525									
3	0.8884									
4	0.8798	0.4990								
5	0.8641	0.5685								
6	0.848	0.5894	0.3548							
7	0.8332	0.5946	0.4245							
8	0.8198	0.5936	0.4538	0.2757						
9	0.8078	0.5899	0.4672	0.3403						
10	0.7971	0.5851	0.4732	0.3715	0.2256					
11	0.7873	0.5798	0.4753	0.3884	0.2844					
12	0.7785	0.5745	0.4753	0.3979	0.3153	0.191				
13	0.7704	0.5692	0.4740	0.4032	0.3335	0.2445				
14	0.7630	0.5642	0.4720	0.4060	0.3449	0.2743	0.1656			
15	0.7562	0.5593	0.4695	0.4071	0.3520	0.2929	0.2146			
16	0.7499	0.5546	0.4669	0.4073	0.3565	0.3050	0.2429	0.1462		
17	0.7441	0.5502	0.4641	0.4067	0.3593	0.3133	0.2613	0.1912		
18	0.7386	0.5460	0.4613	0.4057	0.3609	0.3189	0.2738	0.2181	0.1309	
19	0.7335	0.5420	0.4585	0.4044	0.3617	0.3228	0.2827	0.2360	0.1724	
20	0.7287	0.5383	0.4557	0.4029	0.3619	0.3254	0.2890	0.2487	0.1979	0.1185
21	0.7242	0.5346	0.4530	0.4013	0.3617	0.3271	0.2936	0.2578	0.2153	0.1571
22	0.7199	0.5312	0.4503	0.3995	0.3611	0.3282	0.2969	0.2646	0.2278	0.1812
23	0.7159	0.5280	0.4478	0.3978	0.3603	0.3287	0.2993	0.2696	0.2371	0.1980
24	0.7121	0.5248	0.4453	0.3960	0.3594	0.3289	0.3010	0.2735	0.2441	0.2103
25	0.7084	0.5219	0.4428	0.3942	0.3583	0.3288	0.3022	0.2764	0.2495	0.2196

Table A2. *Cont.*

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
26	0.7050	0.5190	0.4405	0.3924	0.3572	0.3285	0.3029	0.2785	0.2537	0.2267
27	0.7017	0.5163	0.4382	0.3906	0.3560	0.3280	0.3033	0.2801	0.2569	0.2323
28	0.6986	0.5136	0.4360	0.3889	0.3548	0.3274	0.3035	0.2813	0.2595	0.2368
29	0.6956	0.5111	0.4339	0.3872	0.3535	0.3267	0.3035	0.2822	0.2615	0.2403
30	0.6927	0.5087	0.4319	0.3855	0.3523	0.3259	0.3033	0.2827	0.2630	0.2431
31	0.6899	0.5064	0.4299	0.3838	0.3510	0.3250	0.3030	0.2830	0.2641	0.2453
32	0.6873	0.5042	0.4280	0.3822	0.3497	0.3242	0.3025	0.2832	0.2650	0.2471
33	0.6847	0.5020	0.4261	0.3806	0.3484	0.3232	0.3020	0.2832	0.2656	0.2485
34	0.6822	0.4999	0.4243	0.3791	0.3472	0.3223	0.3015	0.2831	0.2661	0.2497
35	0.6799	0.4979	0.4226	0.3776	0.3460	0.3214	0.3009	0.2829	0.2663	0.2505
36	0.6776	0.4960	0.4209	0.3761	0.3447	0.3204	0.3002	0.2826	0.2665	0.2512
37	0.6754	0.4941	0.4192	0.3747	0.3435	0.3194	0.2995	0.2822	0.2665	0.2517
38	0.6733	0.4923	0.4176	0.3733	0.3423	0.3185	0.2988	0.2818	0.2664	0.2521
39	0.6712	0.4905	0.4161	0.3720	0.3412	0.3175	0.2981	0.2813	0.2663	0.2523
40	0.6692	0.4888	0.4146	0.3706	0.3400	0.3166	0.2974	0.2808	0.2661	0.2524
41	0.6673	0.4872	0.4131	0.3694	0.3389	0.3156	0.2966	0.2803	0.2658	0.2524
42	0.6654	0.4856	0.4117	0.3681	0.3378	0.3147	0.2959	0.2798	0.2655	0.2524
43	0.6636	0.4840	0.4103	0.3669	0.3367	0.3137	0.2951	0.2792	0.2652	0.2523
44	0.6618	0.4825	0.4090	0.3657	0.3356	0.3128	0.2943	0.2786	0.2648	0.2522
45	0.6601	0.4810	0.4077	0.3645	0.3346	0.3119	0.2936	0.2780	0.2644	0.2520
46	0.6584	0.4796	0.4064	0.3633	0.3336	0.3110	0.2928	0.2774	0.2639	0.2517
47	0.6568	0.4782	0.4051	0.3622	0.3326	0.3101	0.2921	0.2768	0.2635	0.2514
48	0.6552	0.4768	0.4039	0.3611	0.3316	0.3093	0.2913	0.2761	0.2630	0.2511
49	0.6536	0.4755	0.4027	0.3600	0.3306	0.3084	0.2906	0.2756	0.2625	0.2508
50	0.6521	0.4742	0.4016	0.3590	0.3297	0.3076	0.2898	0.2750	0.2620	0.2505

Table A3. Relative efficiency, $\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| \widehat{\sigma}_{[0]}\right) \times 100\%$.

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
2	**100**									
3	**100**									
4	**100**	25.9								
5	**100**	41.9								
6	**100**	53.1	14.5							
7	**100**	61.6	26.2							
8	**100**	68.4	36.0	10.1						
9	**100**	73.9	44.3	19.3						
10	**100**	78.6	51.6	27.5	7.9					
11	**100**	82.6	57.9	34.9	15.4					
12	**100**	86.1	63.5	41.6	22.4	6.6				
13	**100**	89.2	68.6	47.7	28.9	13.0				
14	**100**	92.0	73.2	53.3	35.0	19.1	5.7			
15	**100**	94.5	77.3	58.5	40.7	24.9	11.3			
16	**100**	96.7	81.1	63.3	46.1	30.4	16.7	5.0		
17	**100**	98.8	84.6	67.8	51.1	35.7	22.0	10.0		
18	100	**100.7**	87.9	72.0	55.9	40.7	27.0	14.9	4.5	
19	100	**102.5**	90.9	75.9	60.3	45.5	31.9	19.7	9.1	
20	100	**104.1**	93.8	79.6	64.6	50.1	36.6	24.4	13.6	4.2
21	100	**105.7**	96.4	83.0	68.6	54.5	41.2	29.0	18.0	8.3
22	100	**107.1**	98.9	86.3	72.4	58.7	45.5	33.4	22.4	12.5
23	100	**108.4**	101.3	89.4	76.1	62.7	49.8	37.7	26.6	16.6
24	100	**109.7**	103.5	92.4	79.6	66.5	53.8	41.9	30.8	20.7
25	100	**110.9**	105.7	95.2	82.9	70.2	57.7	45.9	34.8	24.7
26	100	112.0	107.7	97.9	86.1	73.7	61.5	49.8	38.8	28.6
27	100	**113.1**	109.6	100.4	89.1	77.1	65.2	53.6	42.7	32.5
28	100	**114.1**	111.4	102.9	92.0	80.4	68.7	57.3	46.4	36.3
29	100	**115.0**	113.2	105.2	94.8	83.6	72.1	60.9	50.1	40.0

Table A3. *Cont.*

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
30	100	**116.0**	114.8	107.5	97.6	86.6	75.4	64.4	53.7	43.6
31	100	**116.8**	116.4	109.6	100.2	89.6	78.6	67.7	57.2	47.1
32	100	117.7	**118.0**	111.7	102.7	92.4	81.7	71.0	60.6	50.6
33	100	118.5	**119.4**	113.7	105.1	95.2	84.7	74.2	63.9	54.0
34	100	119.3	**120.9**	115.7	107.5	97.8	87.6	77.3	67.2	57.3
35	100	120.0	**122.2**	117.6	109.7	100.4	90.5	80.4	70.3	60.6
36	100	120.7	**123.6**	119.4	111.9	103.0	93.3	83.3	73.4	63.8
37	100	121.4	**124.8**	121.1	114.1	105.4	95.9	86.2	76.5	66.9
38	100	122.1	**126.1**	122.8	116.2	107.8	98.6	89.0	79.4	69.9
39	100	122.7	**127.3**	124.5	118.2	110.1	101.1	91.7	82.3	72.9
40	100	123.4	**128.4**	126.1	120.1	112.3	103.6	94.4	85.1	75.9
41	100	124.0	**129.5**	127.6	122.0	114.5	106.0	97.0	87.9	78.7
42	100	124.6	**130.6**	129.1	123.9	116.6	108.4	99.6	90.6	81.5
43	100	125.1	**131.7**	130.6	125.7	118.7	110.7	102.1	93.2	84.3
44	100	125.7	**132.7**	132.0	127.4	120.7	112.9	104.5	95.8	87.0
45	100	126.2	**133.7**	133.4	129.1	122.7	115.1	106.9	98.3	89.7
46	100	126.7	134.7	**134.7**	130.8	124.7	117.3	109.2	100.8	92.3
47	100	127.2	135.6	**136.0**	132.4	126.5	119.4	111.5	103.2	94.8
48	100	127.7	136.5	**137.3**	134.0	128.4	121.4	113.8	105.6	97.3
49	100	128.2	137.4	**138.6**	135.5	130.2	123.5	115.9	108.0	99.8
50	100	128.6	138.3	**139.8**	137.1	131.9	125.4	118.1	110.3	102.2

Table 4. Relative efficiency, $\mathrm{RE}\left(\widehat{\sigma}_{[k]} \middle| S_n/c_4(n)\right) \times 100\%$.

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
2	**100.00**									
3	**99.19**									
4	**97.52**	25.24								
5	**95.48**	39.97								
6	**93.30**	49.55	13.48							
7	**91.12**	56.14	23.88							
8	**89.00**	60.84	32.05	9.03						
9	**86.95**	64.27	38.56	16.75						
10	**84.99**	66.80	43.83	23.33	6.74					
11	**83.13**	68.67	48.14	28.97	12.80					
12	**81.36**	70.07	51.69	33.82	18.21	5.36				
13	**79.68**	71.09	54.65	38.00	23.04	10.32				
14	**78.09**	71.83	57.12	41.64	27.34	14.88	4.44			
15	**76.57**	72.35	59.20	44.80	31.19	19.05	8.62			
16	**75.13**	72.69	60.95	47.57	34.63	22.86	12.54	3.78		
17	**73.76**	72.89	62.44	50.00	37.71	26.34	16.19	7.40		
18	72.46	**72.98**	63.70	52.14	40.48	29.52	19.58	10.83	3.30	
19	71.21	**72.98**	64.76	54.03	42.97	32.42	22.73	14.06	6.47	
20	70.02	**72.91**	65.67	55.70	45.23	35.09	25.65	17.10	9.51	2.92
21	68.88	**72.78**	66.43	57.19	47.26	37.53	28.36	19.96	12.41	5.75
22	67.79	**72.59**	67.08	58.51	49.11	39.76	30.88	22.64	15.16	8.48
23	66.75	**72.37**	67.62	59.68	50.78	41.82	33.21	25.16	17.77	11.09
24	65.75	**72.11**	68.07	60.73	52.31	43.72	35.38	27.52	20.24	13.60
25	64.78	**71.82**	68.45	61.66	53.69	45.46	37.40	29.73	22.57	15.99
26	63.86	**71.52**	68.76	62.49	54.96	47.07	39.28	31.81	24.78	18.28
27	62.97	**71.19**	69.00	63.24	56.11	48.56	41.03	33.76	26.87	20.45
28	62.11	**70.86**	69.20	63.90	57.17	49.94	42.66	35.59	28.85	22.52
29	61.29	**70.51**	69.35	64.49	58.13	51.21	44.19	37.31	30.72	24.49
30	60.49	**70.15**	69.46	65.01	59.01	52.39	45.61	38.93	32.49	26.36
31	59.72	**69.78**	69.53	65.48	59.82	53.49	46.94	40.45	34.16	28.15
32	58.97	69.41	**69.57**	65.90	60.56	54.50	48.19	41.89	35.74	29.85
33	58.25	69.03	**69.58**	66.26	61.24	55.44	49.36	43.24	37.24	31.46

Table A4. *Cont.*

n	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
34	57.56	68.65	**69.57**	66.59	61.86	56.32	50.45	44.51	38.66	33.00
35	56.88	68.27	**69.53**	66.87	62.43	57.13	51.48	45.71	40.01	34.46
36	56.23	67.89	**69.48**	67.12	62.95	57.89	52.44	46.85	41.29	35.86
37	55.60	67.51	**69.41**	67.34	63.43	58.60	53.34	47.92	42.50	37.19
38	54.98	67.13	**69.32**	67.53	63.86	59.26	54.19	48.93	43.66	38.46
39	54.38	66.76	**69.21**	67.69	64.26	59.87	54.99	49.89	44.75	39.67
40	53.80	66.38	**69.10**	67.82	64.63	60.44	55.74	50.80	45.79	40.82
41	53.24	66.00	**68.97**	67.94	64.96	60.97	56.44	51.66	46.78	41.92
42	52.69	65.63	**68.83**	68.03	65.27	61.46	57.11	52.47	47.72	42.97
43	52.16	65.26	**68.69**	68.10	65.54	61.93	57.73	53.24	48.61	43.97
44	51.64	64.89	**68.53**	68.16	65.80	62.35	58.32	53.96	49.47	44.93
45	51.14	64.53	**68.37**	68.20	66.03	62.75	58.87	54.66	50.28	45.85
46	50.64	64.17	68.20	**68.22**	66.23	63.13	59.39	55.31	51.05	46.72
47	50.16	63.81	68.02	**68.24**	66.42	63.48	59.89	55.93	51.79	47.56
48	49.69	63.46	67.84	**68.23**	66.59	63.80	60.35	56.55	52.49	48.36
49	49.24	63.11	67.66	**68.22**	66.74	64.10	60.79	57.08	53.16	49.13
50	48.79	62.76	67.47	**68.20**	66.87	64.38	61.20	57.61	53.80	49.86

References

1. Tesfay, Y.Y. Process Control Charts. In *Developing Structured Procedural and Methodological Engineering Designs*; Springer: Cham, Switzerland, 2021. [CrossRef]
2. Montgomery, D.C. *Statistical Quality Control: An Modern Introduction*, 7th ed.; John Wiley & Sons: Singapore, 2013.
3. Benneyan, J.C. Statistical quality control methods in infection control and hospital epidemiology, part I introduction and basic theory. *Infect. Control. Hosp. Epidemiol.* **1998**, *19*, 194–214. [CrossRef]
4. Birnbaum, D.; Benneyan, J.C. Statistical quality control methods in infection control and hospital epidemiology, part II: Chart use, statistical properties, and research issues. *Infect. Control. Hosp. Epidemiol.* **1998**, *19*, 265–283. [CrossRef]
5. Gao, X.; Huang, G.; Zhao, Q.; Cao, C.; Jiang, H. Robust optimization model for medical staff rebalancing problem with data contamination during COVID-19 pandemic. *Int. J. Prod. Res.* **2021**, *1*, 1–30. [CrossRef]
6. Ouyang, L.; Park, C.; Byun, J.H.; Leeds, M. *Robust Design in the Case of Data Contamination and Model Departure. Statistical Quality Technologies*; Springer: Cham, Switzerland, 2019; pp. 347–373.
7. Gao, X.; Cui, C. A note on the warehouse location problem with data contamination. *RAIRO Rech. Opérationnelle* **2021**, *55*, 1113. [CrossRef]
8. Park, C.; Leeds, M. A highly efficient robust design under data contamination. *Comput. Ind. Eng.* **2016**, *93*, 131–142. [CrossRef]
9. Goetz, S.; Roth, M.; Schleich, B. Early Robust Design—Its Effect on Parameter and Tolerance Optimization. *Appl. Sci.* **2021**, *11*, 9407. [CrossRef]
10. Park, C.; Kim, H.; Wang, M. Investigation of finite-sample properties of robust location and scale estimators. *Commun. Stat. Simul. Comput.* **2021**, *1*, 1–27. [CrossRef]
11. Hampel, F.R. The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.* **1974**, *69*, 383–393. [CrossRef]
12. Shamos, M.I. Geometry and statistics: Problems at the interface. In *Algorithms and Complexity: New Directions and Recent Results*; Traub, J.F., Ed.; Academic Press: New York, NY, USA, 1976; pp. 251–280.
13. Abu-Shawiesh, M.O.A.; Saeed, N. Monitoring process variability using decile mean standard deviation. *Qual. Reliab. Eng. Int.* **2021**, *1*, 2996. [CrossRef]
14. Mutlu, E.C.; Alakent, B. Revisiting reweighted robust standard deviation estimators for univariate Shewhart S-charts. *Qual. Reliab. Eng. Int.* **2019**, *35*, 995–1009. [CrossRef]
15. Langenberg, P.; Iglewicz, B. Trimmed mean X and R charts. *J. Qual. Technol.* **1986**, *18*, 152–161. [CrossRef]
16. Dey, K.K.; Dhara, K.; Karmakar, B.; Sengupta, S. Univariate and data-depth based multivariate control charts using trimmed mean and winsorized standard deviation. *arXiv* **2012**, arXiv:1211.4262.
17. Sindhumol, M.R.; Srinivasan, M.M.; Gallo, M. Robust control charts based on modified trimmed standard deviation and Gini's mean difference. *J. Appl. Quant.* **2016**, *11*, 18–30.
18. Abu-Shawiesh, M.O.A. A simple robust control chart based on MAD. *J. Math. Stat.* **2008**, *4*, 102–107.
19. Riaz, M.; Saghir, A. A mean deviation-based approach to monitor process variability. *J. Stat. Comput. Simul.* **2009**, *79*, 1173–1193. [CrossRef]
20. Schoonhoven, M.; Nazir, H.Z.; Riaz, M.; Does, R.J.M.M. Robust location estimators for the control chart. *J. Qual. Technol.* **2021**, *43*, 363–379. [CrossRef]
21. Nazir, H.Z.; Riaz, M.; Does, R.J.; Abbas, N. Robust CUSUM control charting. *Qual. Eng.* **2013**, *25*, 211–224. [CrossRef]
22. Riaz, M. A dispersion control chart. *Commun. Stat. Simul. Comput.* **2008**, *37*, 1239–1261. [CrossRef]

23. Marchant, C.; Leiva, V.; Cysneiros, F.J.A.; Liu, S. Robust multivariate control charts based on Birnbaum–Saunders distributions. *J. Stat. Comput. Simul.* **2018**, *88*, 182–202. [CrossRef]
24. Ouyang, L.; Zhu, S.; Ye, K.; Park, C.; Wang, M. Robust Bayesian hierarchical modeling and inference using scale mixtures of normal distributions. *IISE Trans.* **2021**, *1*, 1–13. [CrossRef]
25. Tukey, J.W.; McLaughlin, D.H. Less Vulnerable Confidence and Significance Procedures for Location Based on a Single Sample: Trimming/Winsorization 1. *Sankhyā Indian J. Stat. Ser. A* **1963**, *25*, 331–352. Available online: http://www.jstor.org/stable/25049278 (accessed on 1 December 2021).
26. Osório, C.; António, M. Performance Evaluation: Subjectivity, Bias and Judgement Style in Sport Data. Bias and Judgement Style in Sport Data. 24 January 2019. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3321838 (accessed on 1 December 2021).
27. Bialik, C. Removing Judges' Bias Is Olympic-Size Challenge. The Wall Street Journal. Available online: https://www.wsj.com/articles/SB10000872396390443477104577551253521597214 (accessed on 7 September 2014).
28. George, C.; Berger, R.L. *Statistical Inference*; Pacific Grove: Duxbury, CA, USA, 2002; Volume 2.
29. David, H.A.; Nagaraja, H.N. *Order Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
30. Hogg, R.V.; Tanis, E.A.; Zimmerman, D.L. *Probability and Statistical Inference*; Pearson/Prentice Hall: Upper Saddle River, NJ, USA, 2010.
31. Tippett, L.H.C. On the extreme individuals and the range of samples taken from a normal population. *Biometrika* **1925**, *17*, 364–387. [CrossRef]
32. Shewhart, W.A. *Economic Control of Quality of Manufactured Product*; Macmillan: London, UK, 1931.
33. Oakland, R. *Quality, Processes and Control. Statistical Process Control*; Routledge: London, UK, 2018; pp. 3–23.
34. Siotani, M. Order statistics for discrete case with a numerical application to the binomial distribution. *Ann. Inst. Stat. Math.* **1956**, *8*, 95–104. [CrossRef]
35. Abdel-Aty, S.H. Ordered variables in discontinuous distributions. *Stat. Neerl.* **1954**, *8*, 61–82. [CrossRef]
36. Hartley, H.O.; David, H.A. Universal bounds for mean range and extreme observation. *Ann. Math. Stat.* **1954**, *25*, 85–99. [CrossRef]
37. Patnaik, P.B. The use of mean range as an estimator of variance in statistical tests. *Biometrika* **1950**, *37*, 78–87. [CrossRef]
38. Cadwell, J.H. The distribution of quasi-ranges in samples from a normal population. *Ann. Math. Stat.* **1953**, *24*, 603–613. [CrossRef]
39. Serfling, R. Asymptotic Relative Efficiency in Estimation. *Int. Encycl. Stat. Sci.* **2011**, *23*, 68–72.
40. Lehmann, E.L. *Elements of Large-Sample Theory*; Springer: Amsterdam, The Netherlands, 2004.
41. Yu, M.; Zhou, Y.; Tsung, F. Robust Online Detection in Serially Correlated Directed Network. *arXiv Prepr.* **2021**, arXiv:2111.02653.
42. Nidsunkid, S.; Lao, M.; Thaithanan, J. The impact of a random vector with variables from normal and non-normal distributions on multivariate control charts. *Qual. Reliab. Eng. Int.* **2021**, *1*, 1–25. [CrossRef]
43. Hodges, J.L., Jr.; Lehmann, E.L. Estimates of location based on rank tests. *Ann. Math. Stat.* **1963**, *34*, 598–611. [CrossRef]