An alternative to the bivariate control chart for process dispersion

Riaz, M.; Does, R.J.M.M.

Published in:
Quality Engineering

DOI:
10.1080/08982110802445579

Citation for published version (APA):
An Alternative to the Bivariate Control Chart for Process Dispersion

Muhammad Riaz¹, Ronald J. M. M. Does²
¹Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
²Institute for Business and Industrial Statistics, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT In this article, we propose a Shewhart-type control chart for monitoring changes in the process variability of a bivariate process. The sample Gini mean differences based matrix \( \bar{G} \) is used as an estimate of the population variance–covariance matrix \( \Sigma \). The newly proposed control chart, denoted by \( \bar{G} \)-chart, is based on the generalized Gini mean differences \( \bar{G} \). For the case of two correlated quality characteristics \( Y \) and \( X \), the design structure of the proposed \( \bar{G} \)-chart is developed assuming bivariate normality of \((Y, X)\). The performance of the proposed \( \bar{G} \)-chart is compared with that of the \( S \)-chart (a sample generalized variance based control chart).

KEYWORDS average run length, control charts, generalized variance, non-normality, normality, process variability

INTRODUCTION

Multivariate statistical process control (MSPC) is used to simultaneously monitor multiple measurements from a process. Multiple process variables might be measured such as temperatures, pressures, concentrations, flow rates, voltages, et cetera. With multiple measurements, each can be monitored in its own control chart. However, this has two disadvantages (see Runger, 2007). One is that it is difficult to control the number of false alarms. The other is that there are often important relations between the variables that should be considered for MSPC. In a multivariate setup, the variance–covariance matrix \( \Sigma \) and the mean vector \( \mu \) are generally used to refer to the spread and location parameters respectively of the distribution of a random vector \( X \). Note that we shall use the following manner of writing: an italic character represents a univariate quantity; e.g., \( Y \); an underlined character represents a vector; e.g., \( \mu \); a tilde under a character represents a matrix; e.g., \( \Sigma \); \( |\Sigma| \) represents the determinant of the matrix \( \Sigma \); and a hat above a character represents an estimate; e.g., \( \hat{\Sigma}^{1/2} \).

Several papers are available in the quality control literature that provide an extensive review of multivariate control charts (cf. Bersimis et al., 2007; Lowry and Montgomery, 1995; Wierda, 1994; Yeh et al., 2006).

Multivariate control charts for controlling the process mean were introduced by Hotelling (1947). Process variability is summarized in the
variance–covariance matrix $\Sigma$. There are two single-number quantities for measuring the overall variability of a set of multivariate data. These are (1) the determinant of the variance–covariance matrix, $|\Sigma|$, which is called the generalized variance; (2) the trace of the variance–covariance matrix, tr$\Sigma$, which is the sum of the variances of the variables.

The issue of multivariate monitoring of process variability is addressed by different researchers (cf. Alt, 1985; Alt and Smith, 1988; Aparaisi et al., 1999, 2001; Khoo and Quah, 2004; Menzlyfriecke, 2007; Montgomery and Wadsworth, 1972). Alt (1985) and Alt and Smith (1988) gave different approaches for monitoring variability of normally distributed quality characteristics in a process, of which the sample generalized variance based control chart (i.e., the $|S|$-chart) is the one that is commonly used. The $|S|$-chart is the standard multivariate control chart for dispersion in statistical software programs, like Minitab 15. In this article we shall also use the generalized variance as a measure of the process dispersion. Although the generalized variance is a widely used measure of variability, it can be misleading in some cases (cf. Lowry and Montgomery, 1995). The reason is that the values of the generalized variance do not represent unique correlations for the underlying variables.

The focus of this article is on monitoring the variability parameter with respect to two correlated quality characteristics having a bivariate normal distribution, following Alt (1985), Alt and Smith (1988), Aparaisi et al. (1999), and Khoo and Quah (2004). They have given a relationship between $|S|^{1/2}$ and $|\Sigma|^{1/2}$ to develop the design structure of the $|S|$-control chart, where $|S|^{1/2}$ and $|\Sigma|^{1/2}$ are the square roots of the determinants. Under bivariate normality $A = 2(n - 1)|S|^{1/2}/|\Sigma|^{1/2}$ has a chi-square distribution with $2n - 4$ degrees of freedom (cf. Khoo and Quah, 2004). Based on this property the probability and the 3-sigma limits based design structure for the $|S|$-chart is easily obtained.

The $|S|$-chart performs well when the vector $X$, of the correlated quality characteristics of interest, follows a multivariate normal distribution. In cases of contaminated multivariate normal distributions and departures from multivariate normality, the $|S|$-chart loses its efficiency. To overcome these problems with the $|S|$-chart, this article proposes a Shewhart-type control chart based on the sample Gini mean differences based matrix, say $G$, as an estimate of the population variance-covariance matrix $\Sigma$. The generalized Gini mean differences, say $|G|$, based control chart, denoted by $|G|$-chart, is proposed to monitor the population-generalized variance $|\Sigma|$. The motivation for this is to obtain control limits that are more robust so that these are less affected by departures from normality. Note that Riaz and Saghir (2007) proposed a Gini mean differences based univariate control chart for monitoring the scale parameter of a normally distributed quality characteristic.

One may object that multivariate methods have not gained much popularity on the shop floor. This is probably due to an important drawback: the interpretation of out-of-control situations signaled by a multivariate chart is usually difficult and involves further statistical evaluation of the data. However, in chapters 7 and 8 in Mason and Young (2002), one finds guidelines on interpretation multivariate control charts and interpretation after a signal. Does et al. (1999) Woodall and Ncube (1985) and demonstrate in their papers that simultaneous univariate charts often perform as well as multivariate charts. Note that we can always use univariate control charts for variability as a supplement to any control chart based on the generalized variance.

This article contains the following topics:

1. The design structure of the proposed $|G|$-chart is developed for the bivariate case assuming bivariate normality. A comparison of the $|G|$-chart is made with the $|S|$-chart in terms of average run length (ARL).

2. The robustness against departures from bivariate normality is examined on the design structures of the $|G|$- and $|S|$-charts. To examine the robustness of the design structures of the $|G|$- and $|S|$-charts, the affected ARLs (i.e., when the parent distribution is either bivariate t, bivariate chi square or bivariate exponential) have been compared with the respective original ARLs (i.e., when the parent distribution is bivariate normal).

**THE PROPOSED CONTROL CHART**

$(|G|$-CHART)

Let $Y$ and $X$ be two correlated quality characteristics that follow a bivariate normal distribution; i.e., $(Y, X) \sim N_2(\mu, \Sigma)$ where $\mu = \left(\begin{array}{c} \mu_y \\ \mu_x \end{array}\right)$ and...
\[ \Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \] (symbolically, the same may also be written in another way as: \((Y, X) \sim N_2(\mu_y, \mu_x, \sigma_y^2, \sigma_x^2, \rho_{xy})\)). The most commonly used measure of multivariate process variability is the variance-covariance matrix \(\Sigma\). A single number representation for the variation expressed by the matrix \(\Sigma\) is its determinant, known as the generalized variance. The population generalized variance is denoted by \(|\Sigma|\) and our interest in this study lies in monitoring \(|\Sigma|^{1/2}\).

Let \((y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)\) be a random sample of size \(n\) from \(N_2(\mu, \Sigma)\), then the sample Gini mean differences based matrix \(\mathcal{G}\) is defined as (cf. David, 1968; Olkin and Yitzhaki, 1992; Riaz and Saghir, 2007; Schechtman and Yitzhaki, 1987; Yitzhaki, 2003):

\[
\mathcal{G} = \begin{bmatrix} G_y^2 & G_{yx} \\ G_{xy} & G_x^2 \end{bmatrix},
\]

with

\[
\begin{align*}
G_y &= (\sqrt{\pi}/2)\text{Cov}(Y, F(Y)) \\
G_x &= (\sqrt{\pi}/2)\text{Cov}(X, F(X)) \\
G_{yx} &= (\sqrt{\pi}/2)\text{Cov}(Y, F(Y))G_x \\
G_{xy} &= (\sqrt{\pi}/2)\text{Cov}(X, F(Y))G_y
\end{align*}
\]

where the symbols \(F\) and \(\text{Cov}\) represent the cumulative distribution function and the covariance, respectively. The sample versions of all the quantities in the formulas of \(G_y, G_x, G_{yx}, \) and \(G_{xy}\) are used here. The generalized Gini mean differences, denoted by \(|G|\), are defined as \(|G| = G_y^2G_x^2 - G_{yx}G_{xy}\). The elements of the matrix \(\mathcal{G}\) estimate the respective elements of the variance-covariance matrix \(\Sigma\) and hence the quantity \(|\mathcal{G}|\) is used as an estimate of the \(|\Sigma|\). In this study, the quantity \(|\mathcal{G}|^{1/2}\) is used for monitoring the quantity \(|\Sigma|^{1/2}\) and for developing the design structure of the proposed \(|\mathcal{G}|\)-chart. The quantity \(|\mathcal{G}|^{1/2}\) represents the square root of the determinant of the Gini’s mean differences based matrix \(\mathcal{G}\).

To develop the design structure of the proposed \(|\mathcal{G}|\)-chart, let \(B\) be a random variable that defines a relationship between \(|\mathcal{G}|^{1/2}\) and \(|\Sigma|^{1/2}\) as:

\[
B = 2(n - 1)|\mathcal{G}|^{1/2}/|\Sigma|^{1/2},
\]

The distributional behavior of \(B\) (in terms of its mean, standard deviation, and quantile points) is required for the development of the design structure of the proposed \(|\mathcal{G}|\)-chart.

### Some Distributional Results for the Proposed \(|\mathcal{G}|\)-Chart

Assuming \((Y, X) \sim N_2(\mu, \Sigma)\), we consider here, without loss of generality, a standard bivariate normal distribution (i.e., \((Y, X) \sim N_2(0, \rho)\) where \(0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\) and \(\rho = \begin{pmatrix} 1 & \rho_{xy} \\ \rho_{yx} & 1 \end{pmatrix}\)). When \((Y, X)\) follow a bivariate normal distribution, the distributional behavior of \(B\) (cf. [2]) entirely depends on \(n\). The distributional properties of the Gini mean differences-based estimators have been discussed by different researchers (cf. David, 1968; Lomnicki, 1952; Nair, 1956; Riaz and Saghir, 2007; Yitzhaki, 2003). From Riaz and Saghir (2007) it follows that \(G_y\) is an unbiased estimate of \(\sigma_y\) in the univariate case. They recommended \(\overline{G}_y\), mean of \(G_y\), computed from an initial set of stable points for an unbiased estimation of \(\sigma_y\) in the dispersion control charts.

However, the distributional behavior of \(G_y^2G_x^2 - G_{yx}G_{xy}\), \((G_y^2G_x^2 - G_{yx}G_{xy})^{1/2}\) and hence \(B\) is not easy to obtain analytically. Therefore, the Monte Carlo simulation technique has been used to explore the distributional behavior of \(B\). “In practice, simulation methods are often used to evaluate the expectation of a statistic,” according to Ross (1990). A detailed discussion regarding the number of simulations required in control chart Monte Carlo simulation studies may be found in Schaffer and Kim (2007). They examined recently published studies to develop recommendations for the minimum number of replications necessary to reproduce the reported results within a specified degree of accuracy. In many cases, only 5,000 replications or fewer were required. In general, the number of replications required to reproduce the target ARL decreased as the shift size increased.

Let \(b_0\), \(b_1\), and \(b_a\), respectively, represent the mean, standard deviation, and \(a\)th quantile point (i.e., the point that has \(a\)% area below it is completed) of the distribution of \(B\), which entirely depends on \(n\) in the case of bivariate normality. The values of \(b_0\), \(b_1\), and \(b_a\) have been obtained as function of \(n\) using the simulation approach.

To conduct a Monte Carlo experiment, we have generated 10,000 random samples, of a given size \(n\), from the standard bivariate normal distribution without loss of generality. For each sample, we have computed the values of the random variable \(B\).
TABLE 1 Table of Coefficients for the $|G_{r}|$-chart

<table>
<thead>
<tr>
<th>$n$</th>
<th>$b_0$</th>
<th>$b_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.540</td>
<td>2.849</td>
</tr>
<tr>
<td>4</td>
<td>4.679</td>
<td>3.405</td>
</tr>
<tr>
<td>5</td>
<td>6.776</td>
<td>4.072</td>
</tr>
<tr>
<td>6</td>
<td>8.851</td>
<td>4.512</td>
</tr>
<tr>
<td>7</td>
<td>10.858</td>
<td>4.974</td>
</tr>
<tr>
<td>8</td>
<td>12.885</td>
<td>5.358</td>
</tr>
<tr>
<td>9</td>
<td>14.815</td>
<td>5.725</td>
</tr>
<tr>
<td>10</td>
<td>16.821</td>
<td>5.985</td>
</tr>
<tr>
<td>11</td>
<td>18.950</td>
<td>6.357</td>
</tr>
<tr>
<td>12</td>
<td>20.878</td>
<td>6.686</td>
</tr>
<tr>
<td>13</td>
<td>22.983</td>
<td>7.007</td>
</tr>
<tr>
<td>14</td>
<td>24.917</td>
<td>7.308</td>
</tr>
<tr>
<td>15</td>
<td>26.865</td>
<td>7.618</td>
</tr>
<tr>
<td>16</td>
<td>28.844</td>
<td>7.743</td>
</tr>
<tr>
<td>17</td>
<td>31.042</td>
<td>8.206</td>
</tr>
<tr>
<td>18</td>
<td>32.935</td>
<td>8.398</td>
</tr>
<tr>
<td>19</td>
<td>34.898</td>
<td>8.570</td>
</tr>
<tr>
<td>20</td>
<td>37.047</td>
<td>8.843</td>
</tr>
<tr>
<td>25</td>
<td>46.964</td>
<td>9.968</td>
</tr>
<tr>
<td>30</td>
<td>56.660</td>
<td>10.808</td>
</tr>
<tr>
<td>50</td>
<td>97.057</td>
<td>14.197</td>
</tr>
<tr>
<td>100</td>
<td>197.341</td>
<td>20.235</td>
</tr>
</tbody>
</table>

followed by its descriptive statistics to obtain $b_0$, $b_1$, and $B_a$. To take care of random variability due to simulation, we have repeated the above procedure 1,000 times. Based on these repeated computations for $b_0$, $b_1$, and $B_a$, we have obtained the mean values of these quantities and their standard errors to report the precision of the results obtained for these quantities (see Tables 1 and 2).

The standard errors for the results of each cell in the aforementioned tables varied between 0.004 and 0.010. Similar results for $b_0$, $b_1$, and $B_a$ can easily be obtained for any value of $n$.

The quantities $b_0$, $b_1$, and $B_a$ are needed to determine the control limits and the power of the proposed $|G_{r}|$-chart to detect shifts in the process $|\sum_{i=1}^{n}|^{1/2}$.

**Design Structure of the Proposed $|G_{r}|$-Chart**

Let $\mu_{|G_{r}|^{1/2}}$ and $\sigma_{|G_{r}|^{1/2}}$ denote the mean and the standard deviation of the distribution of the sample

TABLE 2 Table of Quantile Points for the $|G_{r}|$-chart

<table>
<thead>
<tr>
<th>$n$</th>
<th>$B_{0.05}$</th>
<th>$B_{0.025}$</th>
<th>$B_{0.01}$</th>
<th>$B_{0.005}$</th>
<th>$B_{0.001}$</th>
<th>$B_{0.0005}$</th>
<th>$B_{0.00025}$</th>
<th>$B_{0.00005}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.410</td>
<td>6.375</td>
<td>8.188</td>
</tr>
<tr>
<td>4</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.027</td>
<td>1.849</td>
<td>7.107</td>
<td>9.094</td>
<td>11.134</td>
</tr>
<tr>
<td>5</td>
<td>0.000</td>
<td>0.000</td>
<td>1.692</td>
<td>2.361</td>
<td>3.389</td>
<td>9.758</td>
<td>12.174</td>
<td>14.400</td>
</tr>
<tr>
<td>6</td>
<td>0.000</td>
<td>1.231</td>
<td>2.168</td>
<td>2.938</td>
<td>3.762</td>
<td>5.010</td>
<td>12.310</td>
<td>14.892</td>
</tr>
<tr>
<td>7</td>
<td>1.418</td>
<td>2.130</td>
<td>4.152</td>
<td>5.166</td>
<td>6.629</td>
<td>14.695</td>
<td>17.615</td>
<td>20.258</td>
</tr>
<tr>
<td>25</td>
<td>22.580</td>
<td>24.882</td>
<td>26.628</td>
<td>31.837</td>
<td>34.830</td>
<td>38.376</td>
<td>55.018</td>
<td>64.477</td>
</tr>
<tr>
<td>30</td>
<td>30.033</td>
<td>32.451</td>
<td>34.017</td>
<td>39.669</td>
<td>43.328</td>
<td>47.588</td>
<td>65.568</td>
<td>70.709</td>
</tr>
<tr>
<td>50</td>
<td>58.195</td>
<td>63.378</td>
<td>66.201</td>
<td>74.642</td>
<td>79.084</td>
<td>84.759</td>
<td>108.858</td>
<td>115.514</td>
</tr>
<tr>
<td>100</td>
<td>140.972</td>
<td>149.265</td>
<td>154.062</td>
<td>165.363</td>
<td>172.214</td>
<td>180.193</td>
<td>213.957</td>
<td>223.752</td>
</tr>
</tbody>
</table>

M. Riaz and R. J. M. M. Does

66
statistic $|G|^{1/2}$, respectively. Applying expectations to [2] gives:

$$E(B) = E(2(n - 1)|G|^{1/2}/|\Sigma|^{1/2}) = (2(n - 1)/|\Sigma|^{1/2})E(|G|^{1/2}).$$

In [3], $E(|G|^{1/2})$ can be replaced by its estimate $\overline{|G|^{1/2}}$ (the mean of sample $|G|^{1/2}$ s), using an appropriate number of random samples, from the process under study when the process is in the state of statistical control (cf. Hillier, 1969; Yang and Hillier, 1970). Thus, from [3] an estimate of $|\Sigma|^{1/2}$, after rearranging the terms, is given as:

$$|\Sigma|^{1/2} = 2(n - 1) \frac{\overline{|G|^{1/2}}}{b_0}. \quad [4]$$

The expression for $|\Sigma|^{1/2}$ given in [4] is useful for an unbiased estimation of $|\Sigma|^{1/2}$ using $\overline{|G|^{1/2}}$ and the coefficient $b_0$ provided in Table 1 as a function of $n$.

Also from [3] we have:

$$E(|G|^{1/2}) = \frac{b_0|\Sigma|^{1/2}}{2(n - 1)}. \quad [5]$$

Replacing the estimate of $|\Sigma|^{1/2}$ (given in [4]) in [5] and simplification gives:

$$\hat{\mu}_{|G|^{1/2}} = \overline{|G|^{1/2}}. \quad [6]$$

Also, taking the variance of $B$ and then simplification gives the expression for $\sigma_B$ as:

$$\sigma_B = 2(n - 1)\sigma_{|G|^{1/2}}/|\Sigma|^{1/2}. \quad [7]$$

Rearrangement of (7) yields the following result for $\sigma_{|G|^{1/2}}$:

$$\sigma_{|G|^{1/2}} = \frac{b_1|\Sigma|^{1/2}}{2(n - 1)}. \quad [8]$$

Substituting the estimate for $|\Sigma|^{1/2}$ given in [4] into [8], the estimate for $\sigma_{|G|^{1/2}}$ is given as:

$$\hat{\sigma}_{|G|^{1/2}} = \frac{b_1\overline{|G|^{1/2}}}{b_0}. \quad [9]$$

The expression in [9] is similar to the expression for $\hat{\sigma}_R$ of the $R$-chart as provided in Alwan (2000).

### Parameters of the Proposed $|G|^{-}$-Chart

The central line ($CL$), lower control limit ($LCL$), and upper control limit ($UCL$) are the three parameters of any Shewhart-type control chart. There are two approaches to express these parameters; namely, the probability limit approach and the 3-sigma limit approach. In case of an asymmetric distributional behavior of a relevant estimator, the probability limits approach is preferred. If the distributional behavior of a relevant estimator is nearly symmetric, then the 3-sigma limits approach is a good alternative. In this article we use the probability limits approach because of the asymmetric distributional behavior of the $|G|^{-}$-chart.

#### Probability Limits Approach

The value $\overline{|G|^{1/2}}$ corresponds to the $CL$ of the proposed $|G|^{-}$-chart. Assuming that the probability of making a Type-I error is less than a specified value, say $z$, the control limits (which are actually the true probability limits) for the proposed $|G|^{-}$-chart are defined as:

$$LCL = |\overline{G}^{1/2}| with F_n\left( |\overline{G}^{1/2} = |G_l^{1/2}| \right) \leq z_l \quad \{10\}$$

$$UCL = |\overline{G}^{1/2}| with F_n\left( |\overline{G}^{1/2} = |G_u^{1/2}| \right) \geq 1 - z_u$$

where $|G_l^{1/2}$ and $|G_u^{1/2}$ are the quantile points of the distribution of $|\overline{G}^{1/2}$ below which the areas are $z_l$ and $1 - z_u$ respectively, and $z = z_l + z_u$ and $F_n(X = x)$ represent the cumulative distribution function of $X$ at point $x$, for a given value of $n$.

Now using [2] and [4] in [10] and simplification gives the following:

$$LCL = |G_l^{1/2}| = B_l \frac{\overline{|G|^{1/2}}}{b_0} with F_n(B = B_l) \leq z_l \quad \{11\}$$

$$UCL = |G_u^{1/2}| = B_u \frac{\overline{|G|^{1/2}}}{b_0} with F_n(B = B_u) \geq 1 - z_u$$

Thus, the quantile points of the distribution of $B$, the average of the sample $|G|^{1/2}$ s (i.e., $|\overline{G}|^{1/2}$), and the values of $b_0$ allow setting the true probability limits.
of the proposed \(|G|\)-chart. Similarly, the 3-sigma limits for the proposed \(|\Sigma|\)-chart may easily be defined using \([6]\) and \([9]\).

Once we have computed the control limits of the proposed \(|G|\)-chart for a given significance level by either probability limits approach or the 3-sigma limits approach, the sample statistic \(|G|^{1/2}\) is plotted against the time order of the samples. If all of the sample \(|G|^{1/2}\)'s lie within the control limits, there is reasonable evidence to conclude that there is no shift in the process \(|\Sigma|^{1/2}\) and that the process is stable. Otherwise, some assignable cause(s) are at work causing a shift in the process \(|\Sigma|^{1/2}\).

To address specifically small and moderate shifts: (i) the runs rules (as discussed by Nelson, 1984) may be supplemented to the basic structure of the proposed \(|G|\)-chart of this article, resulting into an increased false alarm rate; (ii) EWMA and CUSUM schemes may be developed based on the statistic \(|G|\) or \(|G|^{1/2}\) (cf. Woodall and Neube, 1985).

For more than two correlated quality characteristics, the design structure of the proposed \(|G|\)-chart may be extended on similar lines. For guidelines regarding more than two correlated quality characteristics of interest, see Gnanadesikan and Gupta (1970) and Khoo and Quah (2004).

**PERFORMANCE EVALUATIONS AND COMPARISONS**

In this section, the performance of the proposed \(|G|\)-chart given earlier is compared with that of the \(|\Sigma|\)-chart given in Khoo and Quah (2004). The comparisons are made for two different situations; namely, the case of a bivariate normal distribution and the case of non-normal bivariate distributions. For comparison purposes, some selective cases of shifts in the parameter values have been considered and the average run length (ARL) has been computed as the performance measure, as is usually done in quality control literature for comparisons among different methods (cf. Hawkins and Maboudou, 2007; Khoo and Quah, 2004).

**The Case of a Bivariate Normal Distribution**

The efficiency of the proposed \(|G|\)-chart as compared to that of the \(|\Sigma|\)-chart has been examined here for the case of a bivariate normal distribution, using the ARL as a performance measure. Probability limits of the \(|G|\)-chart (cf. \([11]\)) and the \(|\Sigma|\)-chart have been obtained for different combinations of \(\alpha\) and \(n\). In addition, the ARLs for the two charts have been computed. The ARLs for some values of \(n\) using \(\alpha = 0.005\) are provided here in Tables 3 and 4 for comparisons between the \(|G|\)- and the \(|\Sigma|\)-charts. For the ARL computations, the shifts in \(|\Sigma|^{1/2}\) are considered in terms of \(\delta|\Sigma|^{1/2}\).

One may observe from these tables that (i) for small values of \(n\), the ARLs of the proposed \(|G|\)-chart are slightly less than those of the \(|\Sigma|\)-chart for small values of \(\delta\), and the two ARLs almost coincide when \(\delta\) increases; (ii) for large values of \(n\), the ARLs of the proposed \(|G|\)-chart are almost the same as those of the \(|\Sigma|\)-chart for all the choices of \(\delta\).

Thus, the proposed \(|G|\)-chart is a competitor to the \(|\Sigma|\)-chart in case of a bivariate normal parent population for detecting shifts in the process \(|\Sigma|^{1/2}\).

**The Case of Non-Normal Bivariate Distributions**

Until now we have assumed that the samples are drawn from a normal distribution. In case this is not true, then an option is to employ a control chart appropriately designed for some particular parent distribution. But in practice, we prefer to have control chart structures that are not much affected by the departures from normality. We examine here departures from normality for our proposed \(|G|\)-chart and the traditional \(|\Sigma|\)-chart. To study the effect of non-normality, two situations are considered: one by disturbing the symmetry and the other by disturbing the kurtosis (peak) of the parent distribution.

For the case of disturbances in symmetry, we use the bivariate exponential and bivariate chi-square distribution. This can be generated by disturbing the symmetry and the other by disturbing the kurtosis (peak) of the parent distribution. Two independent standard normal variables, \(Z_1\) and \(Z_2\), are considered. The results are shown in Tables 3 and 4 for various values of \(\alpha\) and \(n\). The results indicate that the \(|G|\)-chart is more efficient than the \(|\Sigma|\)-chart in detecting shifts in the process \(|\Sigma|^{1/2}\).

**TABLE 3 ARL Values for the \(|G|\)-Chart at \(\alpha = 0.005\) for the Bivariate Normal Parent Distribution**

<table>
<thead>
<tr>
<th>(\delta)</th>
<th>(n = 5)</th>
<th>(n = 10)</th>
<th>(n = 20)</th>
<th>(n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>200.011</td>
<td>199.996</td>
<td>200.023</td>
<td>199.992</td>
</tr>
<tr>
<td>1.50</td>
<td>29.636</td>
<td>12.919</td>
<td>5.050</td>
<td>1.761</td>
</tr>
<tr>
<td>2.00</td>
<td>9.971</td>
<td>3.379</td>
<td>1.555</td>
<td>1.022</td>
</tr>
<tr>
<td>2.50</td>
<td>5.106</td>
<td>1.859</td>
<td>1.117</td>
<td>1.001</td>
</tr>
<tr>
<td>3.00</td>
<td>3.167</td>
<td>1.383</td>
<td>1.027</td>
<td>1.000</td>
</tr>
<tr>
<td>4.00</td>
<td>2.060</td>
<td>1.105</td>
<td>1.001</td>
<td>1.000</td>
</tr>
</tbody>
</table>
distributions; for the case of disturbance in the kurtosis we use the bivariate $t$ distribution.

To examine the effect of departures from bivariate normality, a bivariate random vector, say $U$, is simulated 10,000 times from a bivariate exponential distribution; a bivariate random vector, say $V$, is simulated 10,000 times from a bivariate $t$-distribution; and a bivariate random vector, say $W$, is simulated 10,000 times from a bivariate chi-square distribution such that the mean vectors and the variance-covariance matrices of $U$, $V$, and $W$ are the same as that of a comparable random vector from a bivariate normal distribution. Then the calculations are carried out for the charting characteristic $B$ (cf. [2]) based on the simulated random vectors $U$, $V$, and $W$, and the distribution of $B$ is obtained for the three cases under consideration. This is repeated 1,000 times and mean values of the quantile points of the distribution $B$ are obtained. The rejection region for given $\alpha$ is decided, using mean values of the quantile points of the distributions of $B$ derived from $U$, $V$, and $W$, by the probability limits approach and ARLs of the proposed $|G|$-chart are computed for different shifts in the process $|\Sigma|^{1/2}$ using the quantile points of Table 2, which are obtained assuming bivariate normality. The same is done, along similar lines, for

### Table 4
ARL Values for the $|S|$-Chart at $\alpha = 0.005$ for the Bivariate Normal Parent Distribution

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$n = 5$</th>
<th>$n = 10$</th>
<th>$n = 20$</th>
<th>$n = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>200.016</td>
<td>200.007</td>
<td>200.001</td>
<td>199.999</td>
</tr>
<tr>
<td>1.50</td>
<td>25.126</td>
<td>77.232</td>
<td>18.692</td>
<td>7.351</td>
</tr>
<tr>
<td>2.00</td>
<td>4.895</td>
<td>9.559</td>
<td>4.163</td>
<td>1.640</td>
</tr>
<tr>
<td>2.50</td>
<td>2.354</td>
<td>3.503</td>
<td>2.132</td>
<td>1.139</td>
</tr>
<tr>
<td>3.00</td>
<td>1.600</td>
<td>2.069</td>
<td>1.505</td>
<td>1.055</td>
</tr>
<tr>
<td>4.00</td>
<td>1.169</td>
<td>1.306</td>
<td>1.142</td>
<td>1.142</td>
</tr>
</tbody>
</table>

### Table 5
Original and Affected ARLs of the $|G|$- and $|S|$-Charts at $\alpha = 0.005$ for $n = 5$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>G-chis5</th>
<th>S-chis5</th>
<th>G-chis20</th>
<th>S-chis20</th>
<th>G-Expo</th>
<th>S-Expo</th>
<th>G-t5</th>
<th>S-t5</th>
<th>G-t20</th>
<th>S-t20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>200.016</td>
<td>200.007</td>
<td>200.001</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
</tr>
<tr>
<td>2.00</td>
<td>4.895</td>
<td>9.559</td>
<td>4.163</td>
<td>1.640</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
</tr>
<tr>
<td>2.50</td>
<td>2.354</td>
<td>3.503</td>
<td>2.132</td>
<td>1.139</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
</tr>
<tr>
<td>3.00</td>
<td>1.600</td>
<td>2.069</td>
<td>1.505</td>
<td>1.055</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
</tr>
<tr>
<td>4.00</td>
<td>1.169</td>
<td>1.306</td>
<td>1.142</td>
<td>1.142</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
</tr>
</tbody>
</table>

### Table 6
Original and Affected ARLs of the $|G|$- and $|S|$-Charts at $\alpha = 0.005$ for $n = 10$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>G-chis5</th>
<th>S-chis5</th>
<th>G-chis20</th>
<th>S-chis20</th>
<th>G-Expo</th>
<th>S-Expo</th>
<th>G-t5</th>
<th>S-t5</th>
<th>G-t20</th>
<th>S-t20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>200.016</td>
<td>200.007</td>
<td>200.001</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
</tr>
<tr>
<td>2.00</td>
<td>4.895</td>
<td>9.559</td>
<td>4.163</td>
<td>1.640</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
</tr>
<tr>
<td>2.50</td>
<td>2.354</td>
<td>3.503</td>
<td>2.132</td>
<td>1.139</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
</tr>
<tr>
<td>3.00</td>
<td>1.600</td>
<td>2.069</td>
<td>1.505</td>
<td>1.055</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
</tr>
<tr>
<td>4.00</td>
<td>1.169</td>
<td>1.306</td>
<td>1.142</td>
<td>1.142</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
</tr>
</tbody>
</table>

### Table 7
Original and Affected ARLs of the $|G|$- and $|S|$-Charts at $\alpha = 0.005$ for $n = 20$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>G-chis5</th>
<th>S-chis5</th>
<th>G-chis20</th>
<th>S-chis20</th>
<th>G-Expo</th>
<th>S-Expo</th>
<th>G-t5</th>
<th>S-t5</th>
<th>G-t20</th>
<th>S-t20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>200.016</td>
<td>200.007</td>
<td>200.001</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
<td>200.013</td>
<td>199.999</td>
</tr>
<tr>
<td>2.00</td>
<td>4.895</td>
<td>9.559</td>
<td>4.163</td>
<td>1.640</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
<td>13.010</td>
<td>2.142</td>
</tr>
<tr>
<td>2.50</td>
<td>2.354</td>
<td>3.503</td>
<td>2.132</td>
<td>1.139</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
<td>3.387</td>
<td>2.043</td>
</tr>
<tr>
<td>3.00</td>
<td>1.600</td>
<td>2.069</td>
<td>1.505</td>
<td>1.055</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
<td>5.077</td>
<td>1.730</td>
</tr>
<tr>
<td>4.00</td>
<td>1.169</td>
<td>1.306</td>
<td>1.142</td>
<td>1.142</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
<td>1.297</td>
<td>1.132</td>
</tr>
</tbody>
</table>

Control Chart for Process Dispersion
the $|S|$-chart using the quantile points of a chi-square distribution with $2n - 4$ degrees of freedom for comparison purposes.

Tables 5 through 8 provide the ARLs of the $|G|$- and $|S|$-charts when the samples are generated from the comparable bivariate exponential, $t$ and chi-square distributions, for some selective values of $n$ using $a = 0.005$.

In the above tables, G(S)-Chi5, G(S)-Chi20, G(S)-t5, G(S)-t20, and G(S)-Expo refer to the ARLs of the $|G|$ ($|S|$)-chart when the parent distributions are bivariate chi-square with 5 (respectively 20) degrees of freedom, bivariate $t$ with 5 (respectively 20) degrees of freedom, and bivariate exponential, respectively.

From Tables 5 through 8 we may conclude that the ARLs of the proposed $|G|$-chart are less affected by non-normality compared to those of the $|S|$-chart. A similar behavior is observed for the other values of $n$. Thus, the proposed $|G|$-chart provides a reasonably robust design structure that can be used even if the behavior of the correlated quality characteristics $(Y, X)$ depart from normality.

### An Interesting Phenomenon for the Exponential Distribution

It has been observed that in the case of exponential distribution, the values of ARL increase instead of decrease for small shifts (e.g., $\delta = 1.5$) for both $|G|$- and $|S|$-charts. This increase in the ARL values is larger for the $|S|$-chart compared to that of the $|G|$-chart, especially for small values of $n$, as is obvious from Tables 5 through 7. A similar type of observation for the exponential distribution was also made by Vermaat and Does (2006) in their study regarding a semi-Bayesian method for Shewhart control charts.

Moreover, another aspect of the proposed $|G|$-chart is covered in Riaz (2008), which shows that this proposed chart is not unduly affected by the presence of contaminations (outliers and special causes).

### CONCLUSIONS

This article shows that the proposed $|G|$-chart is a satisfactory competitor to the $|S|$-chart, in case of a normal parent population. Moreover, the design structure of the proposed $|G|$-chart shows more robust behavior compared to that of the $|S|$-chart against departures from normality.

### ACKNOWLEDGEMENTS

The authors thank the referees and the editor for their constructive criticism, which has improved the article substantially.

### ABOUT THE AUTHORS

Ronald J.M.M. Does obtained his MS degree (cum laude) in Mathematics at the University of Leiden in 1976. In 1982 he defended his PhD entitled “Higher Order Asymptotics for Simple Linear Rank Statistics” at the same university. From 1981–1989, he worked at the University of Maastricht, where he became Head of the Department of Medical Informatics and Statistics. In that period his main research interests were medical statistics and psychometrics. In 1989 he joined Philips Electronics as a senior consultant in Industrial Statistics.

Since 1991 he became Professor of Industrial Statistics at the University of Amsterdam. In 1994 he founded the Institute for Business and Industrial Statistics, which operates as an independent consultancy firm within the University of Amsterdam. The projects at this institute involve the implementation, training and support of Lean Six Sigma, among others.
His current research activities lie in the design of control charts for nonstandard situations and the improvement of statistical methods in Lean Six Sigma.

Muhammad Riaz was born in Rawalpindi (Pakistan) on April 19, 1976. He earned his B.Sc., with Statistics and Mathematics A&B as major subjects from the Government Gordon College Rawalpindi, University of the Punjab: Lahore, Pakistan in 1998, M.Sc., in Statistics from the Department of Mathematics and Statistics, Quaid-e-Azam University; Islamabad, Pakistan in 2001 and M.Phil., in Statistics from the Department of Mathematics and statistics, Allama Iqbal Open University, Islamabad, Pakistan in 2006. The title of his M. Phil Dissertation was “Use of Probability weighted moments in Statistical Process Control”. He obtained his Doctoral degree in 2008 from the Institute for Business and Industrial Statistics, University of Amsterdam, Amsterdam, The Netherlands. The title of his PhD thesis was “Improved and Robust Monitoring in Statistical Process Control”.

He served as a Statistical Officer in the Ministry of Food, Agriculture and Livestock, Islamabad, Pakistan during 2002–2003, as a Staff Demographer in the Pakistan Institute of Development Economics, Islamabad, Pakistan during 2003–2004, as a Lecturer in the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan during 2004-2007. He is serving as an Assistant Professor in the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan from 2007-Present.

REFERENCES


