THE SKEW-NORMAL DISTRIBUTION IN SPC

Authors: Fernanda Figueiredo
– CEAUL and Faculdade de Economia da Universidade do Porto, Portugal
otilia@fep.up.pt

M. Ivette Gomes
– Universidade de Lisboa, FCUL, DEIO and CEAUL, Portugal
ivette.gomes@fc.ul.pt

Abstract:

• Modeling real data sets, even when we have some potential (as)symmetric models for the underlying data distribution, is always a very difficult task due to some uncontrollable perturbation factors. The analysis of different data sets from diverse areas of application, and in particular from *statistical process control* (SPC), leads us to notice that they usually exhibit moderate to strong asymmetry as well as light to heavy tails, which leads us to conclude that in most of the cases, fitting a normal distribution to the data is not the best option, despite of the simplicity and popularity of the Gaussian distribution. In this paper we consider a class of skew-normal models that include the normal distribution as a particular member. Some properties of the distributions belonging to this class are enhanced in order to motivate their use in applications. To monitor industrial processes some control charts for skew-normal and bivariate normal processes are developed, and their performance analyzed. An application with a real data set from a cork stopper's process production is presented.

---

1. INTRODUCTION

The most commonly used standard procedures of *statistical quality control* (SQC), control charts and acceptance sampling plans, are often implemented under the assumption of normal data, which rarely holds in practice. The analysis of several data sets from diverse areas of application, such as, *statistical process control* (SPC), reliability, telecommunications, environment, climatology and finance, among others, leads us to notice that this type of data usually exhibit moderate to strong asymmetry as well as light to heavy tails. Thus, despite of the simplicity and popularity of the Gaussian distribution, we conclude that in most of the cases, fitting a normal distribution to the data is not the best option. On the other side, modeling real data sets, even when we have some potential (as)symmetric models for the underlying data distribution, is always a very difficult task due to some uncontrollable perturbation factors.

This paper focus on the parametric family of skew-normal distributions introduced by O'Hagan and Leonard (1976), and investigated with more detail by Azzalini (1985, 1986, 2005), among others.

Definition 1.1. A random variable (rv) Y is said to have a location-scale skew-normal distribution, with location at λ, scale at δ and shape parameter α, and we denote $Y \sim SN(\lambda, \delta^2, \alpha)$, if its probability density function (pdf) is given by

$$(1.1) \quad f(y; \lambda, \delta, \alpha) \,=\, \frac{2}{\delta}\,\phi\!\left(\frac{y-\lambda}{\delta}\right) \Phi\!\left(\alpha\,\frac{y-\lambda}{\delta}\right), \quad y \in \mathbb{R} \ \ (\alpha, \lambda \in \mathbb{R}, \ \delta \in \mathbb{R}^+)\,,$$

where ϕ and Φ denote, as usual, the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. If $\lambda = 0$ and $\delta = 1$, we obtain the standard skew-normal distribution, denoted by $SN(\alpha)$.

This class of distributions includes models with different levels of skewness and kurtosis, apart from the normal distribution itself ($\alpha = 0$). In this sense, it can be considered an extension of the normal family. Allowing departures from the normal model, by the introduction of the extra parameter α that controls the skewness, its use in applications will provide more robustness in inferential methods, and perhaps better models to fit the data, for instance, when the empirical distribution has a shape similar to the normal, but exhibits a slight asymmetry. Note that even in potential normal situations there is some possibility of having disturbances in the data, and the skew-normal family of distributions can describe the process data in a more reliable and robust way. In applications it is also important to have the possibility of regulating the thickness of the tails, apart of the skewness.

The cdf of the skew-normal rv Y defined in (1.1) is given by

$$(1.2) \quad F(y; \lambda, \delta, \alpha) = \Phi\left(\frac{y - \lambda}{\delta}\right) - 2\, T\left(\frac{y - \lambda}{\delta}, \alpha\right), \quad y \in \mathbb{R} \; (\alpha, \lambda \in \mathbb{R},\; \delta \in \mathbb{R}^+)\,,$$

where $T(h, b)$ is the Owen's T function (integral of the standard normal bivariate density, bounded by $x = h$, $y = 0$ and $y = b\,x$), tabulated in Owen (1956), and that can be defined by $T(h, b) = \dfrac{1}{2\pi} \displaystyle\int_0^b \left\{ \mathrm{e}^{-\frac{1}{2} h^2(1 + x^2)} / (1 + x^2) \right\} dx$, $(b, h) \in \mathbb{R} \times \mathbb{R}$.

Although the pdf in (1.1) has a very simple expression the same does not happen with the cdf in (1.2), but this is not a problem that leads us to avoid the use of the skew-normal distribution. We have access to the R package 'sn' (version 0.4-17) developed by Azzalini (2011), for instance, that provides functions related to the skew-normal distribution, including the density function, the distribution function, the quantile function, random number generators and maximum likelihood estimates. The moment generating function of the rv Y is given by $M_Y(t) = 2 \exp\left(\lambda t + \delta^2 t^2 / 2\right) \Phi(\theta \delta t)$, $\forall\, t \in \mathbb{R}$, where $\theta = \alpha / \sqrt{1 + \alpha^2} \in (-1, 1)$, and there exist finite moments of all orders.

Other classes of skew normal distributions, for the univariate and the multivariate case, together with the related classes of skew-t distributions, have been recently revisited and studied in the literature. For details see Fernandez and Steel (1998), Abtahi *et al.* (2011) and Jamalizadeb *et al.* (2011), among others. In this paper some control charts based on the skew-normal distribution are proposed. They still are parametric control charts, and should be compared with the so-called nonparametric or distribution-free control charts that require even less restrictive assumptions, a topic out of the scope of this paper. We merely mention that the nonparametric charts have the same in-control run-length distribution for every continuous distribution, and thus, are by definition robust. In the literature several Shewhart, CUSUM and EWMA type nonparametric control charts have been proposed. Most of them are devised to monitor the location and are based on well-known nonparametric test statistics. For a recent overview on the latest developments on nonparametric control charts, see Chakraborti *et al.* (2011) and references therein.

This paper is organized as follows. Section 2 provides some information about the family of skew-normal distributions, in what concerns properties, random sample generation and inference. Section 3 presents bootstrap control charts for skew-normal processes and some simulation results about their performance. Control charts based on specific statistics with a skew normal distribution are considered to monitor bivariate normal processes, and their properties evaluated. In Section 4, an application in the field of SQC is provided. The paper ends with some conclusions and recommendations in Section 5.

2. THE UNIVARIATE SKEW-NORMAL FAMILY OF DISTRI-BUTIONS

Without loss of generality, we are going to enhance some properties of this family of distributions by considering a standard skew-normal rv X, with pdf

$$(2.1) \qquad f(x;\alpha) = 2\,\phi(x)\,\Phi(\alpha x)\,, \qquad x \in \mathbb{R} \ \ (\alpha \in \mathbb{R})\ .$$

Note that, if $Y \sim SN(\lambda, \delta^2, \alpha)$ then $X = \dfrac{Y - \lambda}{\delta} \sim SN(\alpha)$.

2.1. An overview of some properties

In Figure 1 we illustrate the shape of the pdf of X for several values of α. We easily observe the shape parameter α controls the direction and the magnitude of the skewness exhibited by the pdf. As $\alpha \to \pm\infty$ the asymmetry of the pdf increases, and if the sign of α changes, the pdf is reflected on the opposite side of the vertical axis. For $\alpha > 0$ the pdf exhibits positive asymmetry, and for $\alpha < 0$ the asymmetry is negative.
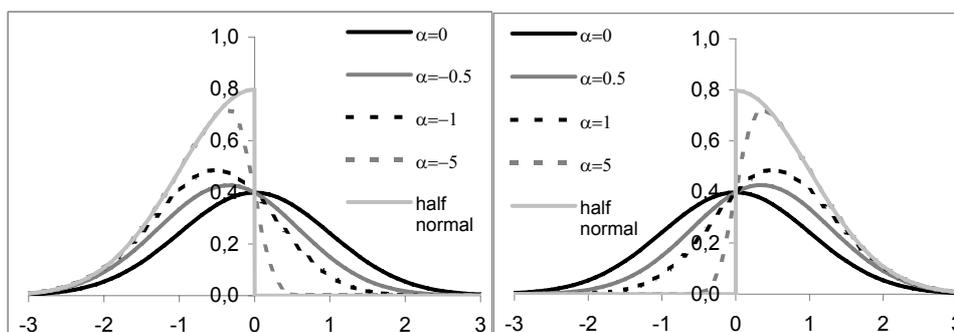
Figure 1: Density functions of standard skew-normal distributions with shape parameter α and the negative and positive half-normal pdf's.

From the Definition 2.1, we easily prove the following results:

Proposition 2.1. *As $\alpha \to \pm\infty$ the pdf of the rv X converges to a half-normal distribution. If $\alpha \to +\infty$, the pdf converges to $f(x) = 2\,\phi(x)$, $x \geq 0$, and if $\alpha \to -\infty$, the pdf converges to $f(x) = 2\,\phi(x)$, $x \leq 0$.*

Proposition 2.2. *If $X \sim SN(\alpha)$ then the rv $W = |X|$ has a half-normal distribution with pdf given by $f(w) = 2\,\phi(w)$, $w \geq 0$, and the rv $T = X^2$, the square of a half-normal distribution, has a pdf given by $f(t) = \frac{1}{\sqrt{2\pi}}\,t^{-1/2}\,e^{-t^2/2}$, $t \geq 0$, i.e., has a chi-square distribution with 1 degree of freedom.*

Denoting the usual sign function by $\mathrm{sign}(\cdot)$ and taking $\theta = \alpha/\sqrt{1+\alpha^2}$, the rv X with a standard skew-normal distribution $SN(\alpha)$ has mean value given by

$$\mathbb{E}(X) = \sqrt{\frac{2}{\pi}}\,\theta \underset{\alpha\to\pm\infty}{\longrightarrow} \mathrm{sign}(\alpha) \times 0.79788 \ ,$$

and variance equal to

$$\mathbb{V}(X) = 1 - \frac{2}{\pi}\,\theta^2 \underset{\alpha\to\pm\infty}{\longrightarrow} 0.36338 \ .$$

The Fisher coefficient of skewness is given by

$$\beta_1 = \frac{(4-\pi)\,\sqrt{2\,\theta^6/\pi^3}}{\sqrt{-8\,\theta^6/\pi^3 + 12\,\theta^4/\pi^2 - 6\theta^2/\pi + 1}} \underset{\alpha\to\pm\infty}{\longrightarrow} \mathrm{sign}(\alpha) \times 0.99527 \ .$$

From these expressions we easily observe that the mean value and the degree of skewness of the $SN(\alpha)$ distribution increases with $|\alpha|$ while the variance decreases, but they all converge to a finite value.

Taking into consideration the large asymmetry of the $SN(\alpha)$ distribution when $\alpha \to \pm\infty$, and the fact that the kurtosis coefficient expresses a balanced weight of the two-tails, we shall here evaluate separately the right-tail weight and the left-tail weight of the $SN(\alpha)$ distribution through the coefficients τ_R and τ_L defined by

$$\tau_\mathrm{R} := \left(\frac{F^{-1}(0.99) - F^{-1}(0.5)}{F^{-1}(0.75) - F^{-1}(0.5)}\right)\left(\frac{\Phi^{-1}(0.99) - \Phi^{-1}(0.5)}{\Phi^{-1}(0.75) - \Phi^{-1}(0.5)}\right)^{-1}$$

and

$$\tau_\mathrm{L} := \left(\frac{F^{-1}(0.5) - F^{-1}(0.01)}{F^{-1}(0.5) - F^{-1}(0.25)}\right)\left(\frac{\Phi^{-1}(0.5) - \Phi^{-1}(0.01)}{\Phi^{-1}(0.5) - \Phi^{-1}(0.25)}\right)^{-1} \ ,$$

where F^{-1} and Φ^{-1} denote the inverse functions of the cdf of the $SN(\alpha)$ and of the cdf of the standard normal distributions, respectively. These coefficients are based on the tail-weight coefficient τ defined in Hoaglin *et al.* (1983) for symmetric distributions. For the normal distribution, $\tau_\mathrm{L} = \tau_\mathrm{R} = 1$. If the distribution F has a right (left) tail heavier than the normal tails, $\tau_\mathrm{R} > 1$ ($\tau_\mathrm{L} > 1$), and if F has a right (left) tail thinner than the normal tails, $\tau_\mathrm{R} < 1$ ($\tau_\mathrm{L} < 1$).

Table 1 presents the mean value, the standard deviation, the median, the skewness coefficient, the left-tail weight and the right-tail weight of the $SN(\alpha)$ distribution for several values of $\alpha > 0$. From the values of Table 1 we notice that when α increases from 0 to $+\infty$, the mean value, the median and the coefficient of skewness increase, but the variance decreases, as expected. The $SN(\alpha)$ distribution has a right-tail heavier than the normal tail, and a left-tail thinner than the normal tail. Moreover, the right tail-weight of the $SN(\alpha)$ quickly converges to 1.1585, the right tail-weight of the half-normal distribution, while the left tail-weight of the $SN(\alpha)$ converges more slowly to the left tail-weight of the half-normal distribution, 0.5393, a value very smaller than the tail-weight of the normal distribution. When α decreases from 0 to $-\infty$ we easily obtain the values of these parameters (coefficients) from the values of this table, taking into consideration that if the sign of α changes, the pdf is reflected on the opposite side of the vertical axis.

Table 1: Mean value (μ), standard deviation (σ), median (μ_e), skewness coefficient (β_1), left-tail weight (τ_{L}) and right-tail weight (τ_{R}) of the $SN(\alpha)$ distribution.

α	μ	σ	μ_e	β_1	τ_{L}	τ_{R}
0	0	1	0	0	1	1
0.3	0.2293	0.9734	0.2284	0.0056	0.9986	1.0017
0.5	0.3568	0.9342	0.3531	0.0239	0.9946	1.0077
1	0.5642	0.8256	0.5450	0.1369	0.9718	1.0457
2	0.7136	0.7005	0.6554	0.4538	0.9008	1.1284
3	0.7569	0.6535	0.6720	0.6670	0.8291	1.1540
5	0.7824	0.6228	0.6748	0.8510	0.7222	1.1584
10	0.7939	0.6080	0.6745	0.9556	0.6124	1.1585
$+\infty$	0.7979	0.6028	0.6745	0.9953	0.5393	1.1585

2.2. Inference

Regarding the estimation of the parameters in the location-scale skew-normal family of distributions, $SN(\lambda, \delta^2, \alpha)$, we are only able to obtain numerical maximum likelihood estimates (MLE), and thus, a closed form for their sampling distribution is not available.

Let $(Y_1, ..., Y_n)$ be a sample of size n from a $SN(\lambda, \delta^2, \alpha)$ distribution. The likelihood function is given by

$$(2.2) \qquad L_{SN}(\lambda, \delta, \alpha) = \frac{2^n}{\delta^n} \prod_{i=1}^{n} \phi\left(\frac{y_i - \lambda}{\delta}\right) \prod_{i=1}^{n} \Phi\left(\alpha \frac{y_i - \lambda}{\delta}\right)$$

and the log-likelihood is given by

$$\ln L_{SN}(\lambda, \delta, \alpha) \;=\; n \ln 2 - n \ln \delta + \sum_{i=1}^{n} \ln \phi \left(\frac{y_i - \lambda}{\delta} \right) + \sum_{i=1}^{n} \ln \Phi \left(\alpha \, \frac{y_i - \lambda}{\delta} \right),$$

where $\ln(\cdot)$ denotes the natural logarithm function.

The MLE estimates of λ, δ and α, denoted $\widehat{\lambda}$, $\widehat{\delta}$ and $\widehat{\alpha}$, are the numerical solution of the system of equations

$$(2.3) \qquad \begin{cases} \delta^2 \;=\; \dfrac{1}{n} \sum_{i=1}^{n} (y_i - \lambda)^2 \;, \\[2em] \alpha \sum_{i=1}^{n} \dfrac{\phi \left(\alpha \, \frac{y_i - \lambda}{\delta} \right)}{\Phi \left(\alpha \, \frac{y_i - \lambda}{\delta} \right)} \;=\; \sum_{i=1}^{n} \dfrac{y_i - \lambda}{\delta} \;, \\[2em] \sum_{i=1}^{n} \dfrac{\frac{y_i - \lambda}{\delta} \, \phi \left(\alpha \, \frac{y_i - \lambda}{\delta} \right)}{\Phi \left(\alpha \, \frac{y_i - \lambda}{\delta} \right)} \;=\; 0 \;. \end{cases}$$

We may have some problems to obtain these estimates in the case of small-to-moderate values of the sample size n as well as for values of α close to zero. Note that if all the values of the sample are positive (negative), for fixed values of λ and δ, the log-likelihood function is an increasing (decreasing) function of α, producing therefore boundary estimates, and for $\alpha = 0$, the expected Fisher information matrix is singular.

Several authors have given important suggestions to find these estimates. For instance, for a fixed value of α, solve the last two equations of (2.3) for obtaining λ and δ, taking into account the first equation, and then, repeat these steps for a reasonable range of values of α. Another suggestion to get around these problems of estimation is to consider another re-parametrization for the skew-normal distributions $SN(\lambda, \delta^2, \alpha)$ in (1.1), in terms of the mean value μ, the standard deviation σ and the asymmetry coefficient β_1. For details in this topic see, for instance, Azzalini (1985), Azzalini and Capitanio (1999) and Azzalini and Regoli (2012), among others.

To decide between the use of a normal or a skew-normal distribution to fit the data, apart from the information given by the histogram associated to the data sample and the fitted pdf estimated by maximum likelihood, we can advance to the confirmatory phase with a likelihood ratio test.

To test the normal distribution against a skew-normal distribution, i.e., the hypotheses $H_0 \colon X \sim SN(\lambda, \delta^2, \alpha = 0)$ versus $H_1 \colon X \sim SN(\lambda, \delta^2, \alpha \neq 0)$, the

likelihood ratio statistic Λ is given by

$$(2.4) \qquad \Lambda = \frac{L_{SN}\left(\widehat{\lambda}, \widehat{\delta}, \alpha = 0\right)}{L_{SN}\left(\widehat{\lambda}, \widehat{\delta}, \widehat{\alpha}\right)} \,,$$

where $L_{SN}(\lambda, \delta, \alpha)$, given in (2.2), denotes the likelihood function for the $SN(\lambda, \delta^2, \alpha)$ distribution. Under the null hypothesis, $-2 \log \Lambda$ is distributed as a chi-square distribution with 1 degree of freedom. For a large observed value of $-2 \log \Lambda$, we reject the null hypothesis, i.e., there is a strong evidence that the $SN\left(\widehat{\lambda}, \widehat{\delta}^2, \widehat{\alpha}\right)$ distribution presents a better fit than the normal $N(\widehat{\mu}, \widehat{\sigma}^2)$ distribution to the data set under consideration.

2.3. Other stochastic results

Among other results valid for the skew-normal distribution, we shall refer the following ones:

Proposition 2.3. *If Z_1 and Z_2 are independent random variables with standard normal distribution, then $Z_1|_{Z_2 \leq \alpha Z_1} \sim SN(\alpha)$. Also,*

$$X := \begin{cases} Z_2 & \text{if } Z_1 < \alpha Z_2 \\ -Z_2 & \text{otherwise} \end{cases} \sim SN(\alpha) \,.$$

Proposition 2.3 allows us to write the following algorithm for the generation of random samples, $(Y_1, ..., Y_n)$, of size n, from a $SN(\lambda, \delta^2, \alpha)$ distribution.

Algorithm 2.1. Repeat Steps 1.–4. for $i = 1$ to n:

1. Generate two independent values, Z_1 and Z_2, from a $N(0,1)$ distribution;

2. Compute $T = \alpha Z_2$;

3. The value $X_i = \begin{cases} Z_2 & \text{if } Z_1 < T \\ -Z_2 & \text{otherwise} \end{cases}$ comes from a $SN(\alpha)$;

4. The value $Y_i = \lambda + \delta X_i$ comes from a $SN(\lambda, \delta^2, \alpha)$.

Figure 2 presents four histograms associated to samples of size one thousand generated from a $SN(\alpha)$ distribution with shape parameter $\alpha = 0, 1, 2, 3$, respectively, together with the pdf's of a normal and of a skew normal distribution fitted to the data by maximum likelihood. From Figure 2 we easily observe that as α increases the differences between the two estimated pdf's become larger,

and the normal fit is not the most appropriate to describe the data. Note that, even in potential normal processes, real data are not exactly normal and usually exhibit some level of asymmetry. Thus, in practice, we advise the use of the skew-normal distribution to model the data.
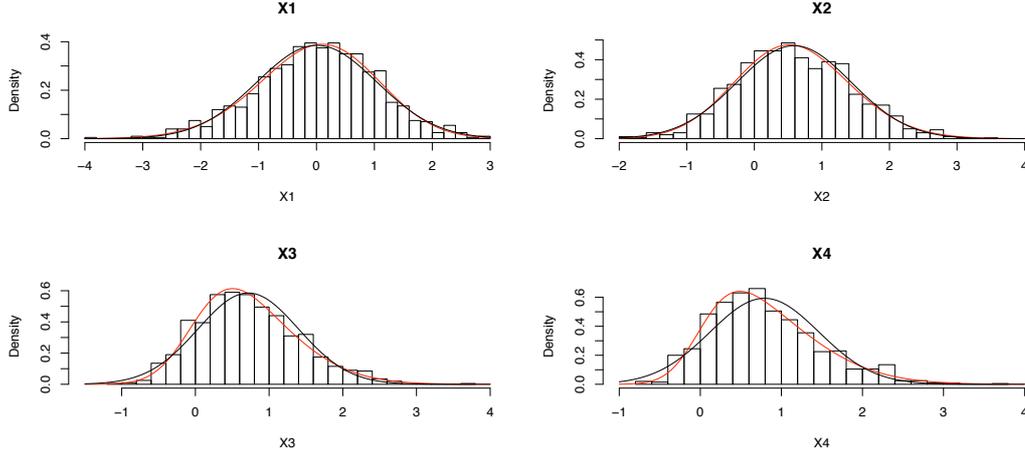
Figure 2: $X_1 \sim SN(0)$, $X_2 \sim SN(1)$, $X_3 \sim SN(2)$, $X_4 \sim SN(3)$.
Histograms and estimated pdf's, $SN(\hat{\lambda}, \hat{\delta}, \hat{\alpha})$ and $N(\hat{\mu}, \hat{\sigma})$.

Another result with high relevance for applications, which allows us to design, in Section 4, control charts to monitor specific bivariate normal processes, is the one presented in Proposition 2.4.

Proposition 2.4.   Let (Z_1, Z_2) be a bivariate normal variable, with $E(Z_1) = E(Z_2) = 0$, $V(Z_1) = V(Z_2) = 1$ and $\mathrm{corr}(Z_1, Z_2) = \rho$. Let $T_m = \min(Z_1, Z_2)$ and $T_M = \max(Z_1, Z_2)$, where $\min(\cdot)$ and $\max(\cdot)$ denote the minimum and the maximum operators, respectively.

  i.   If $\rho = 1$, T_m and T_M have a $N(0,1)$ distribution.

  ii.   If $\rho = -1$, T_m and T_M have half-normal distributions, being $T_m \leq 0$, $\forall m$ and $T_M \geq 0$, $\forall M$.

  iii.   If $|\rho| \neq 1$, $T_m \sim SN(-\alpha)$ and $T_M \sim SN(\alpha)$, with $\alpha = \sqrt{\dfrac{1-\rho}{1+\rho}}$.
   In particular, if Z_1 and Z_2 are independent variables, $\rho = 0$, and then, $T_m \sim SN(-1)$ and $T_M \sim SN(1)$.

3. CONTROL CHARTS BASED ON THE SKEW-NORMAL DISTRIBUTION

The most commonly used charts for monitoring industrial processes, or more precisely, a quality characteristic X at the targets μ_0 and σ_0, the desired mean value and standard deviation of X, respectively, are the Shewhart control charts with 3-sigma control limits. More precisely, the sample mean chart (M-chart), the sample standard deviation chart (S-chart) and the sample range chart (R-chart), which are usually developed under the assumptions of independent and normally distributed data. Additionally, the target values μ_0 and σ_0 are not usually fixed given values, and we have to estimate them, in order to obtain the control limits of the chart.

The ability of a control chart to detect process changes is usually measured by the expected number of samples taken before the chart signals, i.e., by its ARL (*average run length*), together with the *standard deviation of the run length* distribution, SDRL.

Whenever implementing a control chart, a practical advice is that 3-sigma control limits should be avoided whenever the distribution of the control statistic is very asymmetric. In such a case, it is preferable to fix the control limits of the chart at adequate probability quantiles of the control statistic distribution, in order to obtain a fixed ARL when the process is in-control, usually 200, 370.4, 500 or 1000, or equivalently, the desired FAR (*false alarm rate*), i.e., the probability that an observation is considered as out-of-control when the process is actually in-control, usually 0.005, 0.0027, 0.002 or 0.001. General details about Shewhart control charts can be found, for instance, in Montgomery (2005).

In the case of skew-normal processes we do not have explicit formulas for the MLE estimators of the location, scale and shape parameters, and thus, a closed-form for their sampling distribution is not available. The same happens for other statistics of interest, such as, the sample mean, the sample standard deviation, the sample range and the sample percentiles, among others. Thus, to monitor skew-normal processes, the bootstrap control charts are very useful, despite of the disadvantages of a highly time-consuming Phase I. Moreover, many papers, see for instance, Seppala *et al.* (1995), Liu and Tang (1996) and Jones and Woodall (1998), refer that for skewed distributions, bootstrap control charts have on average a better performance than the Shewhart control charts. Other details about the bootstrap methodology and bootstrap control charts can be found, for instance, in Efron and Tibshirani (1993), Bai and Choi (1995), Nichols and Padgett (2006) and Lio and Park (2008, 2010).

3.1. Bootstrap control charts for skew-normal processes

To construct a bootstrap control chart we only use the sample data to estimate the sampling distribution of the parameter estimator, and then, to obtain appropriate control limits. Thus, only the usual assumptions of Phase II of SPC are required: stable process and independent and identically distributed subgroup observations. The following Algorithm 3.1, similar to the ones proposed in Nichols and Padgett (2006) and Lio and Park (2008, 2010), can be used to implement bootstrap control charts for subgroup samples of size n, to monitor the process mean value and the process standard deviation of a skew-normal distribution, respectively. This algorithm can be easily modified in order to implement bootstrap control charts for other parameters of interest.

Algorithm 3.1.

Phase I: Estimation and computation of the control limits

1. From in-control and stable process, observe k, say 25 or 30, random samples of size n, assuming the observations are independent and come from a skew-normal distribution, $SN(\lambda, \delta^2, \alpha)$.

2. Compute the MLE estimates of λ, δ and α, using the pooled sample of size $k \times n$.

3. Generate a parametric bootstrap sample of size n, $(x_1^*, ..., x_n^*)$, from a skew-normal distribution and using the MLEs obtained in Step 2. as the distribution parameters.

4. Select the Step associated to the chart you want to implement:

    i. **Two-sided bootstrap M-chart** to monitor the process mean value μ: from the bootstrap subgroup sample obtained in Step 3., compute the sample mean, $\hat{\mu}^* = \overline{x}^*$.

    ii. **Upper one-sided bootstrap S-chart** to monitor the process standard deviation σ: from the bootstrap subgroup sample obtained in Step 3., compute the sample standard deviation, $\hat{\sigma}^* = s^*$.

5. Repeat Steps 3.–4., a large number of times, say $B = 10\,000$ times, obtaining B bootstrap estimates of the parameter of interest, in our case, the process mean value or the standard deviation.

6. Let γ be the desired false alarm rate (FAR) of the chart. Using the B bootstrap estimates obtained in Step 5.,

    i. Find the $100(\gamma/2)$th and $100(1-\gamma/2)$th quantiles of the distribution of $\hat{\mu}^*$, i.e., the lower control limit LCL and the upper control limit UCL for the bootstrap M-chart of FAR$=\gamma$, respectively.

    ii. Find the $100(1-\gamma)$th quantile of the distribution of $\hat{\sigma}^*$, i.e., the upper control limit UCL for the bootstrap S-chart of FAR$=\gamma$. The lower control limit LCL is placed at 0.

Phase II: Process monitoring

7. Take subgroup samples of size n from the process at regular time intervals. For each subgroup, compute the estimate $\overline{x}$ and s.

8. **Decision:**

   i. If $\overline{x}$ falls between LCL and UCL, the process is assumed to be in-control (targeting the nominal mean value); otherwise, i.e., if the estimate falls below the LCL or above the UCL, the chart signals that the process may be out-of-control.

   ii. If s falls below the UCL, the process is assumed to be in-control (targeting the nominal standard deviation); otherwise, the chart signals that the process may be out-of-control.

In order to get information about the robustness of the bootstrap control limits, we must repeat Steps 1.–6. of Algorithm 3.1 a large number of times, say $r = 1000$, and then, compute the average of the obtained control limits, UCL and LCL, and their associated variances. The simulations must be carried out with different subgroup sample sizes, n, and different levels of FAR, γ. From this simulation study one would expect that, when the subgroup sample size n increases, the control limits get closer together, and when FAR decreases, the limits become farther apart.

In this study, using Algorithm 3.1, we have implemented M and S bootstrap control charts for subgroups of size $n = 5$, to monitor the process mean value of a skew-normal process at a target μ_0, and the process standard deviation at a target σ_0. Without loss of generality we assume $\mu_0 = 0$, $\sigma_0 = 1$ and $\alpha = 0$. The main interest is to detect increases or decreases in μ and to detect increases in σ (and not decreases in σ). The FAR of the charts is equal to $\gamma = 0.0027$, which corresponds to an in-control ARL of approximately 370.4. In Phase I we have considered $k = 25$ subgroups of size $n = 5$.

The performance of these bootstrap control charts to detect changes in the process parameters is evaluated in terms of the ARL, for a few different magnitude changes. When the process changes from the in-control state to an out-of-control state we assume that $\mu = \mu_0 \rightarrow \mu_1 = \mu_0 + \delta\sigma_0$, $\delta \neq 0$ and/or $\sigma = \sigma_0 \rightarrow \sigma_1 = \theta\sigma_0$, $\theta > 1$. In this work we have repeated 30 times Steps 1.–6. of Algorithm 3.1, and then, we have chosen a pair of control limits that allow us to obtain an in-control ARL approximately equal to 370.4, discarding the most extreme upper and lower control limits. Our goal, although out of the scope of this paper, is to improve this algorithm in order to obtain more accurate control limits without replication.

Table 2 presents the ARL values of the bootstrap M-chart and S-chart, and the associated standard deviation SDRL. Indeed, as can be seen from Table 2, the bootstrap control charts present an interesting performance, even when we

consider small changes. As the magnitude of the change increases, the ARL values decrease fast. Despite of the fact that, in SPC, the classical M and S control charts are much more popular, these charts are good competitors, even for the case of normal data if we have to estimate the target process values.

Table 2: ARL and SDRL of the bootstrap M and S charts for subgroups of size $n = 5$. In-control, $\mu = \mu_0$ $(\delta = 0)$ and $\sigma = \sigma_0$ $(\theta = 1)$; when the process is out-of-control we assume either $\mu \to \mu_1 = \delta \neq 0$ or $\sigma \to \sigma_1 = \theta > 1$.

M-chart $(\mu \to \mu_1)$			S-chart $(\sigma \to \sigma_1)$		
δ	ARL	SDRL	θ	ARL	SDRL
0.0	370.5	(371.8)	1.0	370.7	(369.0)
0.1	371.7	(377.2)	1.1	112.8	(112.3)
0.3	168.3	(169.7)	1.2	45.1	(44.4)
0.5	61.5	(61.2)	1.3	22.5	(22.0)
1.0	8.4	(7.8)	1.4	12.9	(12.2)
1.5	2.4	(1.8)	1.5	8.4	(7.9)
2.0	1.3	(0.6)	1.6	6.1	(5.5)
2.5	1.0	(0.2)	1.7	4.6	(4.1)
−0.1	261.9	(261.4)	1.8	3.7	(3.2)
−0.3	90.7	(89.9)	1.9	3.1	(2.5)
−0.5	33.4	(32.4)	2.0	2.6	(2.1)
−1.0	5.0	(4.6)	2.5	1.6	(1.0)
−1.5	1.8	(1.2)			
−2.0	1.1	(0.4)			
−2.5	1.0	(0.1)			

3.2. Control charts for bivariate normal processes

Let (X_1, X_2) be a bivariate normal process and, without loss of generality, assume that the quality characteristics X_1 and X_2 are standard normal variables, possibly correlated, denoting ρ the correlation coefficient. The result presented in Proposition 2.4 allows us to design control charts based on the statistics $T_m = \min(X_1, X_2)$ and $T_M = \max(X_1, X_2)$ to monitor this bivariate normal process.

These univariate statistics permit the implementation of control charts, here denoted T_m-chart and T_M-chart, to monitor simultaneously two related quality characteristics, alternatives to the multivariate control charts based on the Hotelling (1947) statistic and its variants.

Moreover, these charts can be used when in each time of sampling we only have available one observation from each variable of interest, X_1 and X_2, but can be extended to other situations. For instance, when the distributions of X_1 and X_2 have different parameters, replacing X_1 and X_2 by standardized data, and also when we have samples of size greater than one from each of the variables X_1 and X_2, replacing the observations of the samples by the standardized sample means.

First we have implemented a two-sided T_M chart to detect changes in μ, from $\mu_0 = 0$ to $\mu_1 = \mu_0 + \delta\,\sigma_0$, $\delta \neq 0$, assuming that the standard deviation is kept at $\sigma_0 = 1$. We have considered different magnitude changes, and apart from independent data we have also considered correlated data with different levels of positive and negative correlation. The obtained ARL values are presented in Table 3.

Table 3: ARL of the two-sided T_M-chart. $X_i \sim N(\mu, \sigma)$, $i = 1, 2$, corr$(X_1, X_2) = \rho$. In-control: $\mu = \mu_0$ $(\delta = 0)$ and $\sigma = \sigma_0 = 1$; when the process is out-of-control, we assume that only $\mu \to \mu_1 = \delta \neq 0$.

δ \ ρ	0.0	0.1	0.25	0.5	0.9	1.0	-0.25	-0.5
0.0	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4
0.1	361.6	359.5	357.1	354.2	352.7	352.9	368.4	379.6
0.3	249.7	248.6	247.4	247.0	251.0	253.1	253.5	258.7
0.5	144.1	144.0	144.4	145.9	152.5	155.2	144.7	145.5
1.0	36.7	36.9	37.3	38.6	42.5	43.9	36.5	36.4
1.5	11.6	11.7	12.0	12.7	14.4	15.0	11.4	11.3
2.0	4.6	4.7	4.9	5.2	6.0	6.3	4.5	4.4
2.5	2.4	2.4	2.5	2.7	3.1	3.2	2.2	2.2
-0.1	330.8	334.7	339.6	345.9	352.1	352.9	318.2	298.2
-0.3	196.1	204.6	215.9	231.6	249.9	253.1	170.6	135.9
-0.5	100.8	107.9	117.9	132.6	151.5	155.2	80.6	56.8
-1.0	21.7	24.1	27.7	33.5	42.0	43.9	15.7	9.7
-1.5	6.7	7.5	8.8	10.9	14.2	15.0	4.8	3.1
-2.0	2.9	3.2	3.7	4.6	6.0	6.3	2.2	1.7
-2.5	1.7	1.9	2.1	2.4	3.1	3.2	1.4	1.2

From these values we observe that as the magnitude changes increases, the ARL decreases, as expected, and that reductions in μ are detected faster than increases. We easily observe that the level of correlation ρ does not have a great impact on the performance of the chart. However, if the quality characteristics, X_1 and X_2, are positively correlated, the ARL's become larger as the level of correlation increases, i.e., the chart becomes less efficient to detect the change.

Table 4: ARL of the upper one-sided T_M-chart. $X_i \sim N(\mu, \sigma)$, $i = 1, 2$, corr$(X_1, X_2) = \rho$.
In-control: $\mu = \mu_0$ $(\delta = 0)$ and $\sigma = \sigma_0$ $(\theta = 1)$; when the process is out-of-control, $\mu \rightarrow \mu_1 = \delta > 0$ and/or $\sigma \rightarrow \sigma_1 = \theta > 1$.

δ	θ	ρ 0.0	0.1	0.25	0.5	0.9	1.0	−0.25	−0.5
0.0	1.0	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4
	1.1	156.7	156.9	157.4	159.3	167.1	175.0	156.6	156.6
	1.5	22.2	22.4	22.8	23.8	27.6	31.4	22.0	22.0
	2.0	7.7	7.9	8.1	8.6	10.4	12.2	7.6	7.5
	2.5	4.6	4.7	4.9	5.2	6.4	7.5	4.5	4.4
0.1	1.0	268.0	268.1	268.4	269.3	272.3	273.4	268.0	268.0
	1.1	119.5	119.7	120.2	122.2	129.2	135.5	119.3	119.3
	1.5	19.0	19.2	19.5	20.5	23.8	27.1	18.8	18.8
	2.0	7.1	7.2	7.4	7.9	9.5	11.1	6.9	6.8
	2.5	4.3	4.4	4.6	4.9	6.0	7.1	4.2	4.1
0.3	1.0	144.4	144.5	145.0	146.6	151.4	153.1	144.2	144.2
	1.1	71.1	71.3	71.8	73.4	79.0	83.2	70.9	70.9
	1.5	14.2	14.3	14.6	15.4	18.0	20.4	14.0	13.9
	2.0	5.9	6.0	6.2	6.6	8.0	9.3	5.7	5.7
	2.5	3.8	3.9	4.1	4.4	5.3	6.2	3.7	3.6
0.5	1.0	80.7	80.9	81.4	82.9	87.4	89.0	80.0	80.5
	1.1	43.6	43.8	44.3	45.6	49.8	52.6	43.4	43.4
	1.5	10.7	10.8	11.1	11.7	13.8	15.6	10.5	10.5
	2.0	5.0	5.1	5.3	5.6	6.8	7.9	4.8	4.8
	2.5	3.4	3.5	3.6	3.9	4.7	5.5	3.3	3.2
1.0	1.0	22.2	22.4	22.7	23.6	26.0	26.8	22.0	22.0
	1.1	14.7	14.9	15.2	15.9	17.9	19.0	14.5	14.5
	1.5	5.7	5.8	6.0	6.4	7.6	8.5	5.6	5.5
	2.0	3.4	3.5	3.6	3.9	4.7	5.4	3.3	3.2
	2.5	2.6	2.7	2.8	3.0	3.6	4.2	2.5	2.4
1.5	1.0	7.7	7.8	8.1	8.5	9.6	10.0	7.6	7.5
	1.1	6.1	6.1	6.3	6.7	7.7	8.2	5.9	5.8
	1.5	3.4	3.5	3.6	3.9	4.6	5.1	3.3	3.2
	2.0	2.5	2.5	2.6	2.9	3.4	3.8	2.4	2.3
	2.5	2.1	2.2	2.2	2.4	2.9	3.3	2.0	1.9
2.0	1.0	3.4	3.5	3.6	3.9	4.4	4.6	3.3	3.2
	1.1	3.0	3.1	3.2	3.4	4.0	4.2	2.9	2.8
	1.5	2.3	2.3	2.4	2.6	3.0	3.3	2.1	2.1
	2.0	1.9	2.0	2.1	2.2	2.6	2.9	1.8	1.7
	2.5	1.8	1.8	1.9	2.0	2.4	2.7	1.7	1.6
2.5	1.0	1.9	2.0	2.0	2.2	2.5	2.6	1.8	1.7
	1.1	1.8	1.9	2.0	2.1	2.4	2.5	1.7	1.6
	1.5	1.7	1.7	1.8	1.9	2.2	2.4	1.6	1.5
	2.0	1.6	1.6	1.7	1.8	2.1	2.3	1.5	1.4
	2.5	1.5	1.5	1.6	1.7	2.0	2.2	1.4	1.3

On the other hand, the best performance of the chart is obtained when there is a decrease in the process mean value and the quality characteristics are negatively correlated. This control chart is ARL-biased, and maybe due to this fact, we have observed the chart is not appropriate to detect simultaneous changes in μ and σ. Then, we think sensible to implement an upper one-sided T_M-chart to detect changes in μ and/or σ.

From the ARL values presented in Table 4, we conclude that the upper one-sided T_M-chart presents an interesting performance to detect increases in one of the process' parameters, μ or σ, but also to detect simultaneous changes in these parameters. We observe again that the level of correlation, ρ, between the quality characteristics X_1 and X_2, has a small impact on the performance of the chart. Finally, the lower one-sided T_m-chart has a similar performance to detect changes from $\mu \to \mu_1 < 0$ and/or $\sigma \to \sigma_1 > 1$.

4.   AN APPLICATION IN THE FIELD OF SPC

In this section we consider an application to real data from a cork stopper's process production. The objective is modeling and monitoring the data from this process, for which we know the corks must have the following characteristics:

Table 5:   Technical specifications: cork stoppers caliber $45\,\text{mm} \times 24\,\text{mm}$.

Physical quality characteristic (mm)	Mean target	Tolerance interval
Length	45	45 ± 1
Diameter	24	24 ± 0.5

For this purpose we have collected from the process production a sample, of size $n = 1000$, of corks' lengths and diameters. First, we fitted a normal and a skew-normal distribution to the data set. Looking to the histograms obtained from the sample data, presented in Figure 3, both fits seem to be adequate, and the differences between the two pdf's are small.

Then, to test the underlying data distribution, we have used the Shapiro test of normality and the Kolmogorov–Smirnov (K-S) for testing the skew-normal distribution. Unexpectedly, although the fits seem to be similar, from these tests of goodness-of-fit the conclusions are different: the normality for the length's and diameter's data is rejected, for the usual levels of significance (5% and 1%), while
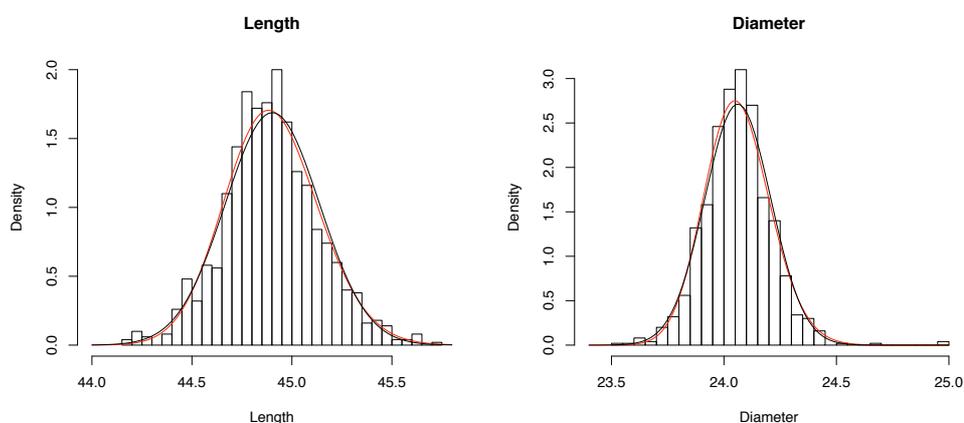
Figure 3:   Histograms and estimated pdf's of the normal and skew-normal
fit to the length and diameter data.

the skew-normal distribution is not rejected. The p-values for the Shapiro and
K-S tests are presented in Table 6. Looking to the maximum likelihood estimates
of some parameters of interest of the fitted distributions, presented in Table 7, we
observe that there exist some differences between the estimates obtained for the
mean value and the location, as well as between the estimates obtained for the
standard deviation and the scale. Moreover, the data exhibit some skewness and
the estimate of the shape parameter is not very close to zero, as it may happen
in the case of normal data.

Table 6:   P-value's of the Shapiro test of normality and of the
Kolmogorov–Smirnov (K-S) for testing a skew-normal.

	Length	Diameter	Decision
Shapiro	0.0018	0.0052	Normality rejected*
K-S	0.2376	0.2923	The skew-normal distribution is not rejected*

* Conclusion for a level of significance of 5% and 1%.

Table 7:   Maximum likelihood estimates of some parameters
of interest of the fitted distributions.

Data	Location	Scale	Shape	Mean	Standard deviation	Skewness
Length	44.7329	0.2907	1.0720	44.9025	0.2361	0.1591
Diameter	23.9526	0.1830	1.1358	24.0622	0.1466	0.1795

To confirm the conclusions obtained by the previous tests of goodness-of-fit we have used the likelihood ratio test presented in subsection 2.2. As we obtained an observed value $-2\ln\Lambda_{\mathrm{obs}} > 3.84$ (for length's and diameter's data), there is a strong evidence that the $SN\big(\widehat{\lambda},\widehat{\delta},\widehat{\alpha}\big)$ distribution presents a better fit than the normal $N(\widehat{\mu},\widehat{\sigma}^2)$ distribution, for a level of significance of 5%.

Finally, based on Algorithm 3.1, we illustrate the implementation of the M and S bootstrap control charts for subgroups of size $n = 10$ to monitor the process mean value and the process standard deviation of the corks' diameter. The Phase I data set consists of $k = 25$ subgroups of size $n = 10$, and we have been led to the following control limits: $\mathrm{LCL} = 23.936484$ and $\mathrm{UCL} = 24.215071$ for the M-chart, and $\mathrm{UCL} = 0.249708$ for the S-chart. From these subgroups we have also estimated the control limits of the corresponding Shewhart charts, assuming normality, here denoted by $\mathrm{LCL_{sh}}$ and $\mathrm{UCL_{sh}}$, and the center line, CL. We obtained $\mathrm{LCL_{sh}} = 23.947788$, $\mathrm{UCL_{sh}} = 24.200532$ and $\mathrm{LC} = 24.07416$ for the M-chart, and $\mathrm{UCL_{sh}} = 0.223152$ and $\mathrm{CL} = 0.129573$ for the S-chart.

In Figure 4 we picture the M and S bootstrap control charts together with the corresponding Shewhart charts with estimated control limits, for use in Phase II of process monitoring. We immediately observe that the bootstrap control limits, LCL and UCL, are set up farther apart than the control limits of the Shewhart M and S charts, $\mathrm{LCL_{sh}}$ and $\mathrm{UCL_{sh}}$.
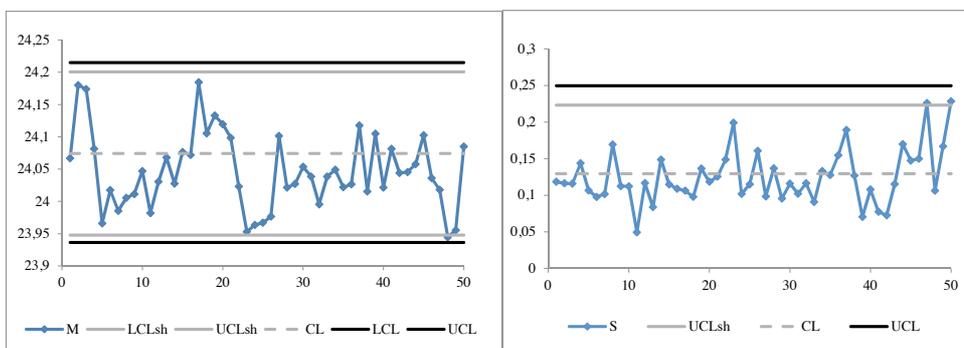
Figure 4: Bootstrap M and S charts together with the corresponding Shewhart charts with estimated control limits.

The Phase II data set used in this illustration consists of $m = 50$ subgroups of size $n = 10$, supposed to be in-control. We have computed the statistics $\overline{x}$ and s associated to these 50 subgroups, and we have plotted them in the charts (here denoted M and S). While the bootstrap charts do not signal changes in the process parameters, the Shewhart charts indicate that the process is out-of-control, due to changes in the process mean value and standard deviation.

5.   SUMMARY AND RECOMMENDATIONS

Designing a control chart under the assumption of skew-normal data and with control limits estimated via bootstrapping adds a relevant contribution to the SPC literature in what concerns the implementation of robust control charts. The use of this family of distributions, that includes the Gaussian as a particular member, allows more flexibility to accommodate uncontrollable disturbances in the data, such as some level of asymmetry or non-normal tail behavior. Moreover, despite of the fact that, in SPC, the classical M and S control charts are much more popular, these charts are good competitors, even for the case of normal data if we have to estimate the target process values.

In order to integrate it within a quality process control system, we can suggest, for instance, an a priori analysis of the process data. A simple boxplot representation with the Phase I data subgroups can anticipate an underlying data distribution that exhibits some level of asymmetry, possibly with some outliers, and in this case, we suggest the use of the proposed bootstrap control charts instead of the traditional Shewhart-type charts implemented for normal data.

Among other issues not addressed in this paper, the proposed control charts should be compared to the existing parametric and nonparametric control charts. Also important is to study the effect of increasing the Phase I sample on the performance of the chart, as well as the determination of the minimum number m of subgroups in Phase I, the sample size n and the number of replicates bootstrap r we must consider in order to have charts with the same performance for the scenarios of known and unknown process parameters. Finally, an exhaustive and comparative study about the performance of control charts based on the skew-normal and on the normal distributions must be carried out to have an idea about the range of values of the shape parameter α of the skew-normal distribution for which the performance of the two charts differ significantly. This will help a practitioner to make a decision on which control chart is preferable to suit his needs.

ACKNOWLEDGMENTS

REFERENCES

[1] ABTAHI, A.; TOWHIDI, M. and BEHBOODIAN, J. (2011). An appropriate empirical version of skew-normal density, *Statistical Papers*, **52**, 469–489.

[2] AZZALINI, A. (1985). A Class of distributions which includes the normal Ones, *Scandinavian J. of Statistics*, **12**, 171–178.

[3] AZZALINI, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, **XLVI**, 199–208.

[4] AZZALINI, A. (2005). The skew-normal distribution and related multivariate families, *Scandinavian J. of Statistics*, **32**, 159–188.

[5] AZZALINI, A. (2011). R package 'sn': The skew-normal and skew-t distributions (version 0.4-17). URL `http://azzalini.stat.unipd.it/SN`.

[6] AZZALINI, A. and CAPITANIO, A. (1999). Statistical applications of the multivariate skew normal distributions, *J. R. Stat. Soc.*, series B, **61**, 579–602.

[7] AZZALINI, A. and REGOLI, G. (2012). Some properties of skew-symmetric distributions, *Annals of the Institute of Statistical Mathematics*, **64**, 857–879.

[8] BAI, D.S. and CHOI, I.S. (1995). $\overline{X}$ and R control charts for skewed populations, *J. of Quality Technology*, **27**(2), 120–131.

[9] CHAKRABORTI, S.; HUMAN, S.W. and GRAHAM, M.A. (2011). *Nonparametric (distribution-free) quality control charts*. In "Methods and Applications of Statistics: Engineering, Quality Control, and Physical Sciences" (N. Balakrishnan, Ed.), 298–329.

[10] CHAKRABORTI, S.; HUMAN, S.W. and GRAHAM, M.A. (2009). Phase I statistical process control charts: an overview and some results, *Quality Engineering*, **21**(1), 52–62.

[11] EFRON, B. and TIBSHIRANI, R. (1993). *Introduction to the Bootstrap*, Chapman and Hall, New York.

[12] FERNANDEZ, C. and STEEL, M.F.J. (1998). On Bayesian modeling of fat tails and skewness, *J. Am. Stat. Assoc.*, **93**, 359–371.

[13] HOAGLIN, D.M.; MOSTELLER, F. and TUKEY, J.W. (1983). *Understanding Robust and Exploratory Data Analysis*, Wiley, New York.

[14] HOTELLING, H. (1947). *Multivariate quality control illustrated by air testing of sample bombsights*. In "Selected Techniques of Statistical Analysis" (C. Eisenhart, M.W. Hastay and W.A. Wallis, Eds.), pp. 111–184, McGraw-Hill, New York.

[15] HUMAN, S.W. and CHAKRABORTI, S. (2010). A unified approach for Shewhart-type phase I control charts for the mean, *International Journal of Reliability, Quality and Safety Engineering*, **17**(3), 199–208.

[16] JAMALIZADEB, A.; ARABPOUR, A.R. and BALAKRISHNAN, N. (2011). A generalized skew two-piece skew normal distribution, *Statistical Papers*, **52**, 431–446.

[17] JONES, L.A. and WOODALL, W.H. (1998). The performance of bootstrap control charts, *J. of Quality Technology*, **30**, 362–375.

[18] LIO, Y.L. AND PARK, C. (2008). A bootstrap control chart for Birnbaum–Saunders percentiles, *Qual. Reliab. Engng. Int.*, **24**, 585–600.

[19]   LIO, Y.L. and PARK, C. (2010). A bootstrap control chart for inverse Gaussian
        percentiles, *J. of Statistical Computation and Simulation*, **80**(3), 287–299.

[20]   LIU, R.Y. and TANG, J. (1996). Control charts for dependent and independent
        measurements based on the bootstrap, *JASA*, **91**, 1694–1700.

[21]   MONTGOMERY, D.C. (2005). *Introduction to Statistical Quality Control*, Wiley,
        New York.

[22]   NICHOLS, M.D. and PADGETT, W.J. (2006). A bootstrap control chart for
        Weibull percentiles, *Qual. Reliab. Engng. Int.*, **22**, 141–151.

[23]   O'HAGAN, A. and LEONARD, T. (1976). Bayes estimation subject to uncertainty
        about parameter constraints, *Biometrika*, **63**, 201–202.

[24]   OWEN, D.B. (1956). Tables for computing bivariate normal probabilities, *Ann.
        Math. Statist.*, **27**, 1075–1090.

[25]   SEPPALA, T.; MOSKOWITZ, H.; PLANTE, R. and TANG, J. (1995). Statistical
        process control via the subgroup bootstrap, *J. of Quality Technology*, **27**, 139–153.