Austrian Journal of Statistics
2025, Volume 54, 22–35.
http://www.ajs.or.at/
doi:10.17713/ajs.v54i4.2017

On a Control Chart for Monitoring Rates and Proportions Based on the Standard Two-sided Power Distribution

Artur J. Lemonte
Universidade Federal do Rio Grande do Norte

Abstract

This paper proposes a new control chart based on the two-parameter standard two-sided power distribution for monitoring rates and proportions, that is, when the quality characteristic of interest belongs to the unit interval $(0, 1)$. Control charts based on the well-known beta and Kumaraswamy distributions are usually considered to deal with this kind of data. The standard two-sided power distribution has many similarities to the beta and Kumaraswamy distributions and a number of advantages in terms of tractability. We evaluate and compare the performance of the new control chart with the beta and Kumaraswamy control charts through Monte Carlo simulation experiments. The simulation results reveal that the control chart based on the standard two-sided power distribution outperforms the beta and Kumaraswamy control charts in terms of run length analysis. An empirical application to a real data set is considered to illustrate the new control chart in practice, and comparisons with the two most traditional control charts for rates and proportions (beta and Kumaraswamy) are made.

Keywords: average run length, Beta distribution, Kumaraswamy distribution, statistical process control.

1. Introduction

The time-honored two-parameter beta family of distributions has been utilized extensively in statistical theory and practice, mainly when the interval used is the standard unit interval $(0, 1)$, since the data can be interpreted as rates or proportions. Its probability density function (PDF) is given by

$$f_{\mathrm{B}}(x) = \frac{x^{\lambda-1}(1-x)^{\xi-1}}{B(\lambda, \xi)}, \quad 0 < x < 1, \tag{1}$$

where $\lambda > 0$ and $\xi > 0$ are shape parameters, $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$ is the beta function, and $\Gamma(\cdot)$ denotes the complete gamma function. Kumaraswamy (1980) has introduced an alternative distribution to the beta distribution, which is known in the literature as Kumaraswamy ('Kw' for short) distribution. Its PDF has the form

$$f_{\mathrm{Kw}}(x) = \gamma\delta x^{\gamma-1}(1 - x^{\gamma})^{\delta-1}, \quad 0 < x < 1, \tag{2}$$

where $\gamma > 0$ and $\delta > 0$ are shape parameters. The literature about the beta and Kw distributions is vast, and these distributions have been extensively applied in several areas. Here, we refer the reader to the interesting study regarding the genesis, similarities and differences between the beta and Kw distributions provided by Jones (2009); see also the references cited therein.

The two-parameter standard two-sided power ('STSP') family of distributions was introduced in van Dorp and Kotz (2002), and its PDF is given by

$$f_{\mathrm{vDK}}(x) = \begin{cases} \eta \left(\dfrac{x}{\theta}\right)^{\eta-1}, & 0 < x \leq \theta, \\ \eta \left(\dfrac{1-x}{1-\theta}\right)^{\eta-1}, & \theta < x < 1, \end{cases} \tag{3}$$

where $0 \leq \theta \leq 1$ governs the location and skewness of the distribution (it is also called "turning point"), and $\eta > 0$ corresponds to the shape parameter. For $\eta = 1$, the STSP distribution simplifies to the standard uniform distribution on $(0, 1)$. For $\eta = 2$, it corresponds to a triangular distribution, and for $\theta = 1$, we have the power function distribution. van Dorp and Kotz (2002) derived various properties of the STSP distribution and discussed its flexibility as compared with the beta distribution. They also provided a novel application in the area of financial engineering. The STSP PDF can be unimodal, increasing, decreasing, U-shaped, symmetric, and left skewed or right skewed depending on the values of its parameters. In other words, the flexibility of the two-parameter STSP distribution is comparable to that of the well-known beta and Kw distributions. The reader is referred to van Dorp and Kotz (2002) and Kotz and van Dorp (2004, Cap. 3) for further details regarding the STSP family of distributions. It is worth stressing that the STSP distribution does not seem to be very familiar and has not received much attention in the statistical literature. In what follows, for the first time, we shall consider this distribution in a control chart framework to monitor rates and proportions.

An interesting area of research corresponds to Statistical Process Control, which is commonly used in monitoring and detecting shifts in the production processes through control charts. In particular, in many practical situations of control chart applications, there is a great interest in monitoring rates and proportions of the quality characteristic. It is well-known that the most used control chart when the data of interest correspond to fractions of integer values is the p control chart (Montgomery 2019), which assumes that the distribution of the nonconforming fraction has a binomial distribution and, hence, the control limits are computed by using the normal approximation (see, for example, Wang 2009). According to Wang (2009), the conventional p control chart constructed by the normal approximation for the binomial distribution may perform poorly, mainly when the characteristic of interest presents a proportion close to zero or one. In addition, another important disadvantage of the p chart is that the control limits computed from the normal approximation may not belong to the unit interval $(0, 1)$, and so the p chart analysis is compromised. To overcome the p chart flaws, some alternative control charts have been proposed in the literature. To mention a few, but not limited to, the reader is referred to Winterbottom (1993), Chen (1998), Sim and Lim (2008), Wang (2009), Chen (2013), Joekes and Barbosa (2013), Mukherjee and Chakraborti (2012) and Graham, Mukherjee, and Chakraborti (2017), but their control limits are still not restricted to the double bounded interval $(0, 1)$, and so the analysis based on these control charts is also compromised.

It is worth mentioning that there exist situations where rates or proportions are not results from Bernoulli experiments although the values belong to the unit interval $(0, 1)$. In addition, variables described as rates and proportions belong to the unit interval $(0, 1)$, which means that the usual assumption of normal approximation to obtain the control limits becomes inadequate for its modeling and monitoring. Consequently, the p chart is not suitable for such a situation. In this case, the quality characteristic has to be necessarily modeled using continuous distributions defined in the standard unit interval $(0, 1)$. The first step in this direction was provided by Sant'Anna and ten Caten (2012), who proposed a control chart to monitor rates and proportions based on the two-parameter beta distribution in (1). More recently, Lima-Filho and Bayer (2021) introduced a control chart for monitoring rates and proportions based on the two-parameter Kw distribution in (2). Obviously, due to the great flexibility and applicability of the beta and Kw distributions in several areas, the construction of control charts using these distributions becomes natural. It is worth stressing that the control limits obtained from the beta and Kw distributions will always belong to the unit interval $(0, 1)$, once these distributions are defined in this interval.

It is evident that the beta and Kw control charts have their merits and, hence, can be applied for monitoring rates and proportions in several areas. As we will later, the use of these charts depends on estimates of the unknown parameters, that is, their control limits depend on unknown parameters and they have to be replaced by estimates obtained from the data in practical applications. The maximum likelihood (ML) procedure is widely used in estimating these control charts, mainly due to the good properties of the ML estimates, such as consistency, asymptotic efficiency, and invariance principle. The ML estimates of the beta and Kw model parameters are obtained by solving nonlinear equations, which means that the ML estimates cannot be expressed in closed form, and so numerical methods need to be used to obtain them numerically as, for example, the Newton-Raphson algorithm. In this paper, similar to the works of Sant'Anna and ten Caten (2012) and Lima-Filho and Bayer (2021), we shall propose a new control chart for monitoring rates and proportions based on the two-parameter STSP distribution in (3). Besides the great flexibility of the STSP distribution mentioned earlier, the ML estimates of the STSP model parameters have closed form and, consequently, no numerical method is need to obtain them in practice.

It is worth emphasizing that there exist several real-world applications where the STSP, beta, and Kw control charts can be considered: (i) in a manufacturing process of frozen orange juice concentrate, where the percentage of nonconforming units is monitored (Montgomery 2019); (ii) in the study of contaminated peanuts by toxic substances, where the proportion of non-contaminated peanuts is monitored (Draper and Smith 1998); (iii) in a manufacturing process of nitric acid by ammonia oxidation collected, where the proportion of ammonia unconverted is monitored (Brownlee 1965); (iv) in the monitoring of the relative humidity given in terms of percentage, being the ratio of the partial pressure of water to the equilibrium vapor pressure of water (Lima-Filho and Bayer 2021); and (v) in a radial tire manufacturing process of a multinational company of rubber products, where the proportion of unconverted mass is monitored, being defined by the rate between the volume of raw material that was not converted into product and the total volume (Bayer, Tondolo, and Muller 2018); among others. The existence of real-world applications where these control charts can be applied enhances the practical significance of this paper.

The remaining of the paper is organized as follows. Section 2 presents a review of the beta and Kw control charts, and the estimation of the corresponding control limits of these charts is also briefly reviewed. We introduce the STSP control chart and its control limits estimation procedure considering individual measurements in Section 3. Monte Carlo simulation experiments to evaluate and compare the performance of the STSP control chart with the beta and Kw control charts are presented in Section 4. An empirical application to real data is presented and discussed in Section 5 for illustrative purposes. The paper closes up with some concluding remarks in Section 6.

2. Beta and Kw control charts

Here, we briefly review the beta and Kw control charts. Detailed descriptions of these charts are provided in Sant'Anna and ten Caten (2012) and Lima-Filho and Bayer (2021).

2.1. Beta control chart

Let X_{B} be a random variable beta-distributed, i.e. $X_{\text{B}} \sim \text{Beta}(\lambda, \xi)$. The mean and variance of X_{B} are

$$\mu_{\text{B}} := \mathbb{E}(X_{\text{B}}) = \frac{\lambda}{\lambda+\xi}, \quad \sigma^2_{\text{B}} := \mathbb{VAR}(X_{\text{B}}) = \frac{\lambda\xi}{(\lambda+\xi)^2(\lambda+\xi+1)}.$$

Let $Q_{\text{B}}(\omega; \lambda, \xi)$ denote the quantile function of the beta distribution, where $\omega \in (0,1)$. Given an in-control process and a false alarm probability, say $\alpha \in (0,1)$, the lower control limit (LCL), central line (CL) and upper control limit (UCL) of the beta control chart can be expressed as $\text{LCL}_{\text{B}} = Q_{\text{B}}(\alpha/2; \lambda, \xi)$, $\text{CL}_{\text{B}} = \lambda/(\lambda+\xi)$ and $\text{UCL}_{\text{B}} = Q_{\text{B}}(1-\alpha/2; \lambda, \xi)$, respectively, where $\text{Pr}_0(\text{LCL}_{\text{B}} \leq X_{\text{B}} \leq \text{UCL}_{\text{B}}) = 1-\alpha$ and, for simplicity, we consider $\text{Pr}_0(X_{\text{B}} \leq \text{LCL}_{\text{B}}) = \alpha/2$ and $\text{Pr}_0(X_{\text{B}} \geq \text{UCL}_{\text{B}}) = \alpha/2$; see Lima-Filho and Bayer (2021). Also, $\text{Pr}_0(\cdot)$ means that the probability is calculated under the in-control process parameters, and CL represents the mean value of the quality characteristic based on the in-control state.

In practical applications, the parameters λ and ξ are unknown and have to be estimated from the data at hand. The ML estimates $\widehat{\lambda}$ and $\widehat{\xi}$ of λ and ξ, respectively, are widely used in this case. Thus, we have the corresponding ML estimates

$$\widehat{\text{LCL}}_{\text{B}} = Q_{\text{B}}(\alpha/2; \widehat{\lambda}, \widehat{\xi}),$$
$$\widehat{\text{CL}}_{\text{B}} = \frac{\widehat{\lambda}}{\widehat{\lambda}+\widehat{\xi}},$$
$$\widehat{\text{UCL}}_{\text{B}} = Q_{\text{B}}(1-\alpha/2; \widehat{\lambda}, \widehat{\xi}).$$

Let $x_1, x_2, \ldots, x_n$ be an observed sample of size n from $X_{\text{B}} \sim \text{Beta}(\lambda, \xi)$. The ML estimates $\widehat{\lambda}$ and $\widehat{\xi}$ are obtained as solutions of the following nonlinear system of equations:

$$\begin{cases} \sum_{i=1}^{n} \ln(x_i) = n[\psi(\widehat{\lambda}) - \psi(\widehat{\lambda}+\widehat{\xi})], \\ \sum_{i=1}^{n} \ln(1-x_i) = n[\psi(\widehat{\xi}) - \psi(\widehat{\lambda}+\widehat{\xi})], \end{cases}$$

where $\psi(\cdot)$ is the digamma function. The ML estimates $\widehat{\lambda}$ and $\widehat{\xi}$ have no closed form and, hence, a numerical method needs to be used to obtain them numerically, such as the Newton-Raphson algorithm. There are several functions in the R programming language (R Core Team 2023) to compute the ML estimates of the beta distribution parameters. However, the R functions consider the above nonlinear system of equations to obtain $\widehat{\lambda}$ and $\widehat{\xi}$ numerically. Evidently, there is no guarantee of converge of the iteration process to compute $\widehat{\lambda}$ and $\widehat{\xi}$ in practice.

2.2. Kw control chart

The mean and variance of $X_{\text{Kw}} \sim \text{Kw}(\gamma, \delta)$ are

$$\mu_{\text{Kw}} := \mathbb{E}(X_{\text{Kw}}) = \gamma B\left(1+\frac{1}{\gamma}, \delta\right),$$

$$\sigma^2_{\text{Kw}} := \mathbb{VAR}(X_{\text{Kw}}) = \gamma B\left(1+\frac{2}{\gamma}, \delta\right) - \left[\gamma B\left(1+\frac{1}{\gamma}, \delta\right)\right]^2.$$

Let $Q_{\text{Kw}}(\omega; \gamma, \delta)$ be the Kw quantile function, for $\omega \in (0,1)$. From Lima-Filho and Bayer (2021), the LCL, CL and UCL of the Kw control chart are $\text{LCL}_{\text{Kw}} = Q_{\text{Kw}}(\alpha/2; \gamma, \delta)$, $\text{CL}_{\text{Kw}} =$

$\gamma B(1+1/\gamma,\delta)$ and $\mathrm{UCL_{Kw}} = Q_{\mathrm{Kw}}(1-\alpha/2;\gamma,\delta)$, respectively, where $\mathrm{Pr}_0(\mathrm{LCL_{Kw}} \leq X_{\mathrm{Kw}} \leq \mathrm{UCL_{Kw}}) = 1-\alpha$, $\mathrm{Pr}_0(X_{\mathrm{Kw}} \leq \mathrm{LCL_{Kw}}) = \alpha/2$ and $\mathrm{Pr}_0(X_{\mathrm{Kw}} \geq \mathrm{UCL_{Kw}}) = \alpha/2$.

In practice, after obtaining the ML estimates $\widehat{\gamma}$ and $\widehat{\delta}$ of γ and δ, respectively, by using the ML procedure, we obtain the following ML estimates

$$\begin{aligned}
\widehat{\mathrm{LCL}}_{\mathrm{Kw}} &= Q_{\mathrm{Kw}}(\alpha/2;\widehat{\gamma},\widehat{\delta}),\\
\widehat{\mathrm{CL}}_{\mathrm{Kw}} &= \widehat{\gamma}B\left(1+\frac{1}{\widehat{\gamma}},\widehat{\delta}\right),\\
\widehat{\mathrm{UCL}}_{\mathrm{Kw}} &= Q_{\mathrm{Kw}}(1-\alpha/2;\widehat{\gamma},\widehat{\delta}).
\end{aligned}$$

Let $x_1, x_2, \ldots, x_n$ denote an observed sample of size n from $X_{\mathrm{Kw}} \sim \mathrm{Kw}(\gamma,\delta)$. The ML estimates $\widehat{\gamma}$ and $\widehat{\delta}$ are obtained as solutions of the following nonlinear system of equations:

$$\begin{cases}
\dfrac{n}{\widehat{\gamma}} + \sum_{i=1}^{n}\ln(x_i) + (1-\widehat{\delta})\sum_{i=1}^{n}\dfrac{x_i^{\widehat{\gamma}}\ln(x_i)}{1-x_i^{\widehat{\gamma}}} = 0,\\
\dfrac{n}{\widehat{\delta}} + \sum_{i=1}^{n}\ln(1-x_i^{\widehat{\gamma}}) = 0.
\end{cases}$$

From the above equations, we have that $\widehat{\delta} = -n/\sum_{i=1}^{n}\ln(1-x_i^{\widehat{\gamma}})$, and the ML estimate $\widehat{\gamma}$ comes from the solution of the nonlinear equation

$$\frac{n}{\widehat{\gamma}}\left[1+T_1(\widehat{\gamma})+\frac{T_2(\widehat{\gamma})}{T_3(\widehat{\gamma})}\right] = 0,$$

where $T_1(\gamma) = (1/n)\sum_{i=1}^{n}\ln(y_i)/(1-y_i)$, $T_2(\gamma) = (1/n)\sum_{i=1}^{n}y_i\ln(y_i)/(1-y_i)$, $T_3(\gamma) = (1/n)\sum_{i=1}^{n}\ln(1-y_i)$, and $y_i = x_i^{\gamma}$ for $i = 1,2,\ldots,n$. The ML estimate $\widehat{\gamma}$ cannot be expressed in closed form and, hence, it has to be computed numerically via iterative techniques such as the Newton-Raphson algorithm. The R programming language also has some functions to find numerically the ML estimates of the Kw distribution parameters. In practice, similar to the beta distribution, there is no guarantee that the iteration process to compute $\widehat{\gamma}$ and $\widehat{\delta}$ numerically will converge.

3. STSP control chart

In the following, we introduce a new control chart to monitor rates and proportions based on the STSP distribution. Let X_{vDK} be a random variable STSP-distributed, say $X_{\mathrm{vDK}} \sim \mathrm{STSP}(\theta,\eta)$. Its cumulative distribution function has the form

$$F_{\mathrm{vDK}}(x) = \begin{cases}
\theta\left(\dfrac{x}{\theta}\right)^{\eta}, & 0 < x \leq \theta,\\
1-(1-\theta)\left(\dfrac{1-x}{1-\theta}\right)^{\eta}, & \theta < x < 1.
\end{cases}$$

The corresponding quantile function can be expressed in the form

$$Q_{\mathrm{vDK}}(\omega;\theta,\eta) = \begin{cases}
\theta^{(\eta-1)/\eta}\omega^{1/\eta}, & 0 < \omega \leq \theta,\\
1-(1-\theta)^{(\eta-1)/\eta}(1-\omega)^{1/\eta}, & \theta < \omega < 1,
\end{cases}$$

where $\omega \in (0,1)$. It is evident that the cumulative distribution function as well as the quantile function are very simple and need no special functions, which compares extremely favorably in terms of simplicity with the two-parameter beta distribution. For example, the cumulative distribution function of the beta distribution corresponds to an incomplete beta function ratio, while its quantile function the inverse thereof. We have that

$$\mu_{\mathrm{vDK}} := \mathbb{E}(X_{\mathrm{vDK}}) = \frac{(\eta-1)\theta+1}{\eta+1}, \quad \sigma^2_{\mathrm{vDK}} := \mathbb{VAR}(X_{\mathrm{vDK}}) = \frac{\eta-2(\eta-1)\theta(1-\theta)}{(\eta+2)(\eta+1)^2}.$$

We can express the mean in the form

$$\mu_{\text{vDK}} = \frac{1}{\eta+1}\,0 + \frac{(\eta-1)}{\eta+1}\,\theta + \frac{1}{\eta+1}\,1,$$

which corresponds to a weighted average of the lower bound 0, the parameter θ, and upper bound 1. Also, the parameter η determines the weights of the above weighted average. van Dorp and Kotz (2002, § 2.1) also provided interesting relationships between the mean and the parameters θ and η, which renders the STSP distribution of intuitive transparency. In short, the two-parameter STSP family of distributions is fairly tractable and, hence, it becomes an interesting alternative to the well-known beta and Kw distributions for modeling rates and proportions. Consequently, the use of this distribution in a control chart setup to monitor rates and proportions becomes natural.

The control limits of the STSP control chart are obtained in a similar way as the control limits of the beta and Kw control charts. Given an in-control process and a false alarm probability α, we have that $\text{Pr}_0(\text{LCL}_{\text{vDK}} \leq X_{\text{vDK}} \leq \text{UCL}_{\text{vDK}}) = 1-\alpha$. We also consider for simplicity that $\text{Pr}_0(X_{\text{vDK}} \leq \text{LCL}_{\text{vDK}}) = \alpha/2$ and $\text{Pr}_0(X_{\text{vDK}} \geq \text{UCL}_{\text{vDK}}) = \alpha/2$. Hence, the LCL, CL and UCL of the STSP control chart are given by $\text{LCL}_{\text{vDK}} = Q_{\text{vDK}}(\alpha/2;\theta,\eta)$, $\text{CL}_{\text{vDK}} = [(\eta-1)\theta+1]/(\eta+1)$ and $\text{UCL}_{\text{vDK}} = Q_{\text{vDK}}(1-\alpha/2;\theta,\eta)$, respectively. It is clear that the control limits depend also on the unknown parameters θ and η, and the ML procedure will be considered to estimate these parameters in practice. Let $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ be the corresponding order statistics of a sample $x_1, x_2, \ldots, x_n$ of size n from a STSP(θ,η) distribution. The ML estimates $\widehat{\theta}$ and $\widehat{\eta}$ of θ and η, respectively, were obtained by van Dorp and Kotz (2002) and are given by

$$\widehat{\theta} = x_{(r)}, \qquad \widehat{\eta} = -\frac{n}{\ln[M(r)]},$$

where $r = \arg\max_{s\in\{1,2,\ldots,n\}} M(s)$, and

$$M(s) = \prod_{i=1}^{s-1} \frac{x_{(i)}}{x_{(s)}} \prod_{i=s+1}^{n} \frac{1-x_{(i)}}{1-x_{(s)}}.$$

van Dorp and Kotz (2002, § 3) provided a simple procedure to compute $\widehat{\theta}$ and $\widehat{\eta}$. It is evident that the ML estimate $\widehat{\theta}$ is given by a specific order statistic, and the ML estimate $\widehat{\eta}$ has a simple analytical expression. Consequently, it is not necessary to consider nonlinear optimization algorithms to obtain these ML estimates in practice, as well as to estimate the corresponding control limits. Evidently, this tractability advantage of the STSP distribution regarding the ML estimates of the parameters θ and η is indeed of immense practical significance in statistics. The ML estimates of the control limits are

$$\begin{aligned}
\widehat{\text{LCL}}_{\text{vDK}} &= Q_{\text{vDK}}(\alpha/2;\widehat{\theta},\widehat{\eta}),\\
\widehat{\text{CL}}_{\text{vDK}} &= \frac{(\widehat{\eta}-1)\widehat{\theta}+1}{\widehat{\eta}+1},\\
\widehat{\text{UCL}}_{\text{vDK}} &= Q_{\text{vDK}}(1-\alpha/2;\widehat{\theta},\widehat{\eta}).
\end{aligned}$$

If the points x_i (for all $i = 1,2,\ldots,n$) are within the two control limits ($\widehat{\text{LCL}}_{\text{vDK}}$ and $\widehat{\text{UCL}}_{\text{vDK}}$), the process is considered to be in-control and so no action is needed in such a case. On the other hand, a point that is outside of the control limits revals out-of-control conditions and, therefore, it is necessary to consider corrective actions in the process.

4. Control chart performance

Here, we consider Monte Carlo simulation experiments to evaluate and compare the performance the STSP control chart for individual observations with the beta and Kw control charts.

To do so, we consider the average run length (ARL), which provides information on the average number of events until a detection of an out-of-control condition (Montgomery 2019). The ARL can be used from two perspectives (Montgomery 2019), namely: ARL_0 and ARL_1, where ARL_0 is considered to evaluate processes under control, and ARL_1 is considered to evaluate processes out-of-control. We have that $\text{ARL}_0 = 1/\alpha$, where $\alpha = 1 - \text{Pr}_0(\text{LCL} \leq X \leq \text{UCL})$ is the probability of an out-of-control point be detected, given that the process is under control (i.e. the probability of a false alarm occurring). Also, $\text{ARL}_1 = 1/(1 - \beta)$, where $\beta = \text{Pr}_1(\text{LCL} \leq X \leq \text{UCL})$ is the probability of an out-of-control point be detected, once the process is, indeed, out-of-control, and $\text{Pr}_1(\cdot)$ means that the probability is calculated under the out-of-control process. The numerical evaluation of the STSP, beta and Kw charts was carried out by considering $\alpha = 0.01$, and so for a process in-control we expect $\text{ARL}_0 = 100$. It is worth mentioning that other values for α are also commonly considered in the literature for evaluating control charts. The choice of α is related directly to the interval length given by the control limits of the chart; that is, the smaller the value of α, the greater the interval length. There is no standard value for α, but $\alpha = 0.01$ and $\alpha = 0.005$ are commonly considered in the literature; see, for example, Lima-Filho, Bayer, and da Silva (2021), Hossain and Riaz (2021) and Sagrillo, Guerra, Machado, and Bayer (2023), among others. In what follows, we evaluate and compare the performance of the STSP, beta and Kw control charts in terms of ARL (ARL_0 and ARL_1).

For the data-generating process, we investigate three different scenarios: the STSP distribution as the true data-generating process, the beta distribution as the true data-generating process, and the Kw distribution as the true data-generating process. Under these scenarios, different degrees of variability and kurtosis are considered, as well as symmetric and asymmetric PDFs. We consider 15000 Monte Carlo replications with 1000 occurrences of the $\text{STSP}(\theta, \eta)$, $\text{Beta}(\lambda, \xi)$ and $\text{Kw}(\gamma, \delta)$ distributions in each Monte Carlo replication. All simulations were performed considering the R programming language (R Core Team 2023). We consider the ML procedure to estimate the unknown parameters of the STSP, beta and Kw distributions in each Monte Carlo replication. Comparisons among the STSP, beta, and Kw control charts were carried out considering in-control and out-of-control simulations. A control chart with better performance corresponds to the one that most closely approximates to the nominal level in terms of ARL_0, but in a case of out-of-control low values of ARL_1 is preferable. To evaluate the ARL_1, we introduce an ε change in the process ($\mu_1 = \mu_0 + \varepsilon$), on what μ_0 is the in-control mean of the random variable considered (STSP, beta or Kw) and $\mu_1 = \mu_0 + \varepsilon$ is the out-of-control mean of the random variable considered (STSP, beta or Kw). Since ε is the induced change in the process, when $\varepsilon = 0$, the process is in-control. The control charts limits were adjusted to obtain ARL_0 equals to the specified nominal value of 100. The calibration of the control charts with the aim of equating the ARL_0 is suggested in the literature; see, for example, Moraes, Oliveira, Quinino, and Duczmal (2014).

Table 1 lists the simulation results for the in-control process. Under the STSP true data-generating process, it is evident that the best performance corresponds to the STSP control chart, since the empirical values of ARL_0 are very close to the respective nominal value 100 in all cases, while the beta and Kw control charts presented poor results. For example, when the true data-generating process is given by $\text{STSP}(\theta = 0.5, \eta = 4.0)$, the relative distortion of the empirical ARL_0 to the nominal value are 47.8% and 42.8% for the beta and Kw control charts, respectively; that is, the beta and Kw control charts obtained (approximately) averages of false alarms every 52 and 57 samples, respectively, thus compromising the performance of these control charts when the true data-generating process comes from the STSP distribution. On the other hand, under the beta and Kw true data-generating processes, the STSP control chart does not deliver good results. Interestingly, the beta and Kw control charts presented a similar performance in terms of ARL_0 in some cases considered, regardless of the true data-generating process. The similarity between the beta and Kw control charts was briefly reported in Lima-Filho and Bayer (2021) and confirmed by the simulations in Table 1.

Figure 1 displays the results of ARL_1 for the STSP, beta, and Kw control charts under the

Table 1: Performance of the STSP, beta and Kw control charts: ARL_0

STSP true data-generating process			
Scenario	STSP	Beta	Kw
STSP($\theta = 0.25, \eta = 2.0$)	98.54	78.60	82.29
STSP($\theta = 0.5, \eta = 4.0$)	98.66	52.23	57.15
STSP($\theta = 0.25, \eta = 1.5$)	98.28	89.20	91.11
STSP($\theta = 0.75, \eta = 2.5$)	98.63	72.74	74.97
STSP($\theta = 0.3, \eta = 4.0$)	98.78	56.55	59.23
Beta true data-generating process			
Scenario	STSP	Beta	Kw
Beta($\lambda = 3.0, \xi = 3.0$)	182.3	100.29	103.95
Beta($\lambda = 5.2, \xi = 2.0$)	63.45	98.73	103.23
Beta($\lambda = 2.2, \xi = 5.0$)	63.45	98.73	101.96
Beta($\lambda = 7.5, \xi = 2.5$)	71.05	97.74	104.15
Beta($\lambda = 3.2, \xi = 7.0$)	71.05	97.74	99.29
Kw true data-generating process			
Scenario	STSP	Beta	Kw
Kw($\gamma = 2.2, \delta = 3.0$)	130.2	93.90	98.75
Kw($\gamma = 2.0, \delta = 5.0$)	77.50	92.26	98.52
Kw($\gamma = 5.2, \delta = 3.0$)	70.66	90.84	98.75
Kw($\gamma = 1.5, \delta = 2.5$)	86.82	95.77	98.74
Kw($\gamma = 2.5, \delta = 2.5$)	145.4	93.37	98.74

STSP true data-generating process. Note that there is no uniform superiority of a control chart in relation to another one in terms of ARL_1; that is, in some cases the ARL_1 is smaller for the STSP control chart, smaller for the beta control chart, or smaller for the Kw control chart. (ARL_1 curves for the STSP, beta, and Kw control charts under the beta and Kw true data-generating processes showed a similar pattern. To save space, the results are not shown.) Basically, the numerical results regarding ARL_1 reveal that there is no uniform superiority of one control chart with respect to the others, thus providing to practitioners the importance of considering an appropriated distribution for monitoring variables that are rates and proportions to reduce false alarms and to ensure the power of the control chart.

In summary, the Monte Carlo simulation results provide important information. The performance of the beta and Kw control charts can be very poor when the data-generating process comes from the two-parameter STSP family of distributions. These results also reveal the importance of considering adequate distributions for rates and proportions to reduce false alarms and to ensure the chart's power to detect changes when the quality characteristic of interest belongs to the unit interval $(0, 1)$. In practice, however, the true data-generating process is unknown. Consequently, it is quite important to select a distribution that produces the best fit for the real data at hand before applying a control chart based on the chosen distribution. Assuming that the process is in control, the practitioner can proceed as follows: (i) Select the possible candidate distributions for rates and proportions as, for instance, the beta, Kw and STSP distributions; (ii) By using the ML method, estimate the unknown parameters of the distributions; (iii) Compute, for example, the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the nonparametric Kolmogorov–Smirnov (KS) statistic of all fitted distributions to the real data at hand; (iv) Select the distribution that delivered the lowest values of AIC, BIC and KS; (v) Assuming that the choice was adequate, compute the control limits and start monitoring the process.

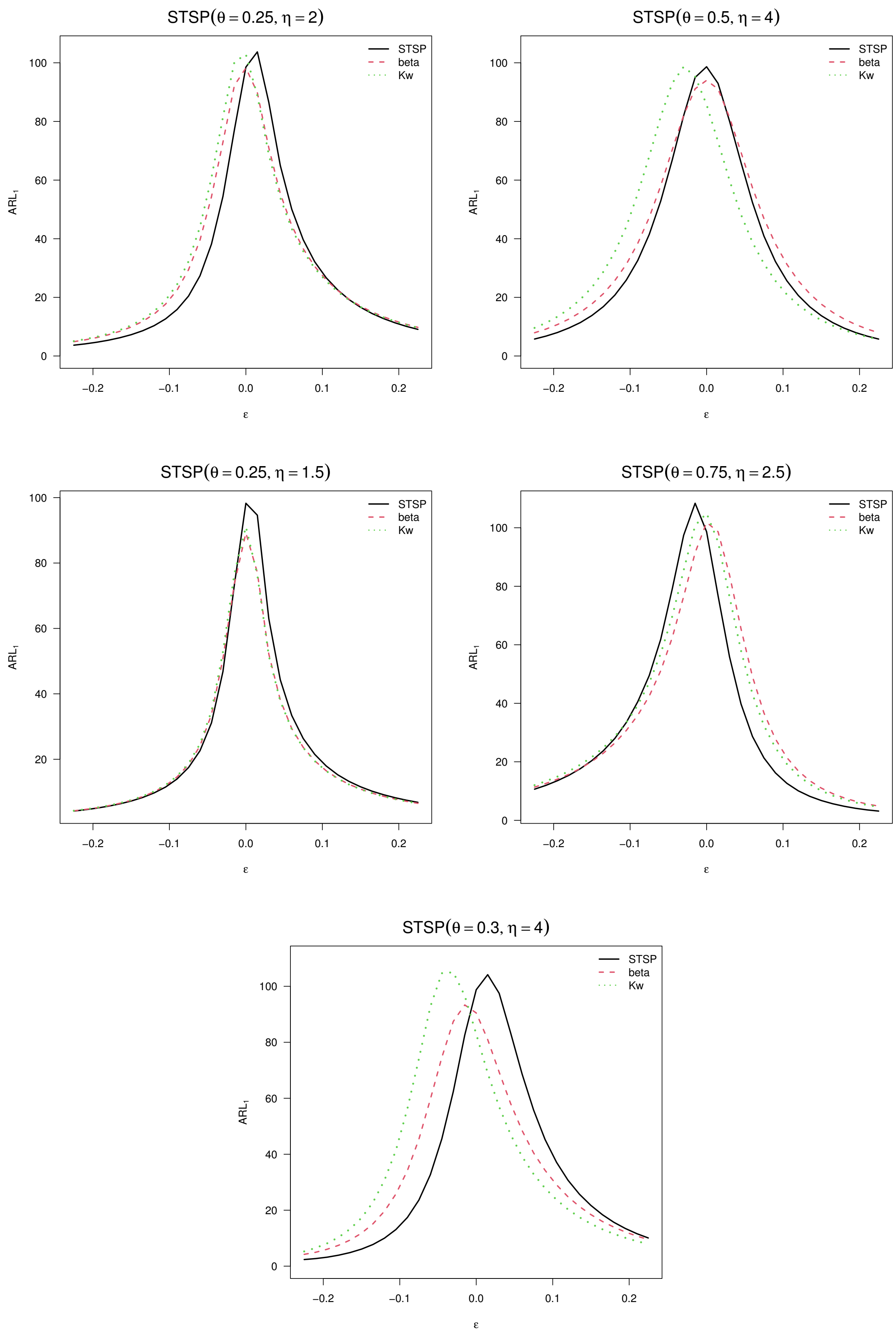

Figure 1: ARL_1 curves to evaluate the STSP, beta and Kw control charts for $\text{ARL}_0 = 100$. STSP true data-generating process

5. Real data illustration

In the following, we consider the STSP control chart in a real data application for illustrative purposes. We also consider the beta and Kw control charts for the sake of comparison. The real data used here correspond to a study of contaminated peanut by toxic substances in 34 batches of 120 pounds (Sant'Anna and ten Caten 2012). We have that the quality characteristic monitored corresponds to the proportion of non-contaminated peanuts, which is defined by the ratio between the volume of non-conforming raw material and the total volume produced; that is, the variable monitored is the ratio between continuous numbers. The data are listed in Table 2. For the STSP, beta and Kw control charts application, the real dataset was split into two groups, namely: the first 20 observations were considered only for control limits estimation (phase I), and the remaining 14 observations were used in phase II (monitoring phase).

Table 2: Proportion of non-contaminated peanuts

0.971	0.979	0.982	0.971	0.957	0.961	0.956	0.972	0.889
0.961	0.982	0.975	0.942	0.932	0.908	0.970	0.985	0.933
0.858	0.987	0.958	0.909	0.859	0.863	0.811	0.877	0.798
0.855	0.788	0.821	0.830	0.718	0.642			

It is evident that the STSP(θ, η), Beta(λ, ξ) and Kw(γ, δ) distributions can be used to model the variable monitored. The ML estimates of the unknown parameters based on the first 20 observations (phase I) are listed in Table 3. The values of AIC, BIC and KS of the fitted distributions are also presented in this table. According to the values of the selection criteria, the STSP distribution outperforms the beta and Kw distributions and, hence, should be chosen to model these data. The p-values of the KS test statistic are 0.986 (STSP distribution), 0.681 (beta distribution) and 0.681 (Kw distribution), which unquestionably favors the STSP distribution. Note also that the AIC, BIC and KS (p-value) of the beta and Kw distributions are quite near, which means that these two distributions are equivalent to model these data. The STSP distribution, on the other hand, provides a clear improvement over these distributions in the modeling of these data according to these selection criteria. Following the procedure described at the end of Section 4, the STSP distribution was the most appropriate (smaller AIC, BIC and KS values), and so the corresponding STSP control chart seems adequate to monitor these data. In addition, the histogram of the data and the estimated STSP, beta and Kw PDFs are displayed in Figure 2, which indicates that the STSP distribution yields a good fit to the data.

Table 3: ML estimates, and selection criteria

Distribution	ML estimates	AIC	BIC	KS (p-value)
STSP(θ, η)	$\widehat{\theta} = 0.987$, $\widehat{\eta} = 28.472$	-91.362	-89.371	0.102 (0.986)
Beta(λ, ξ)	$\widehat{\lambda} = 46.656$, $\widehat{\xi} = 2.280$	-85.456	-83.464	0.161 (0.681)
Kw(γ, δ)	$\widehat{\gamma} = 37.078$, $\widehat{\delta} = 2.765$	-86.103	-84.111	0.160 (0.681)

The STSP, beta and Kw control charts are displayed in Figure 3. We consider a type I error probability of $\alpha = 0.01$ to yield an $\text{ARL}_0 = 100$. Note that the control limits of the beta and Kw control charts are near. Also, the control charts do not trigger an alarm in phase I, and so any contradictions against the models are not obtained in this phase. In phase II (monitoring phase), the beta and Kw control charts indicated an out-of-control point at sample 5. On the other hand, the STSP control chart indicated an out-of-control point at sample 12. It is noteworthy that these results are in agreement with the Monte Carlo results presented in the

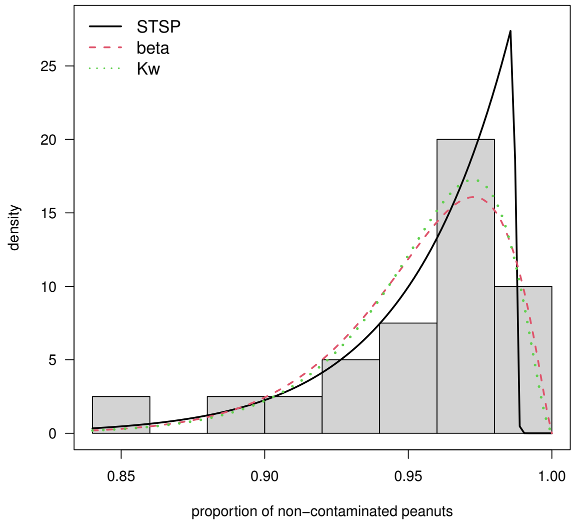

Figure 2: Histogram and estimated STSP, beta and Kw PDF's

previous section, where was verified that the beta and Kw control charts trigger a false alarm much earlier than the STSP control chart when the true data-generating process comes from the STSP distribution, that is, a lower ARL_0 indicates accruing more false alarms. Finally, since the STSP distribution is superior to the beta and Kw distributions in terms of model fitting, the beta and Kw control charts have to be used with some care by practitioners to avoid identifying an out-of-control point earlier, i.e. to prevent using equivocated control limits.

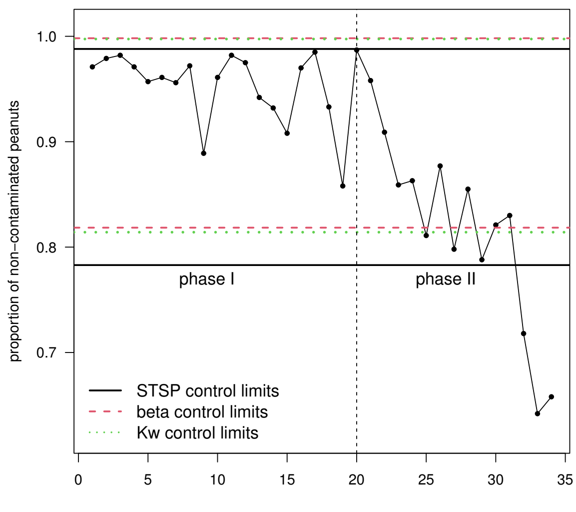

Figure 3: STSP, beta and Kw control charts for the proportion of non-contaminated peanuts

6. Concluding remarks

In this paper, we formally propose a new control chart for rates and proportions (not obtained from Bernoulli experiments) restricted to the range $(0, 1)$ based on the two-parameter standard two-sided power distribution. This distribution, besides being very flexible in modeling rates and proportions, is fairly tractable. The performance of the new control chart was evaluated and compared with the two most traditional control charts using Monte Carlo simulation experiments, namely: beta and Kumaraswamy control charts. The numerical results were quite promising and revealed that the proposed control chart is superior to the beta and Kumaraswamy control charts in terms of ARL, indicating that the use of equivocated control limits may provoke anticipation of false alarms based on the beta and Kumaraswamy control charts when the true data-generating process comes from the two-parameter standard two-sided power distribution. A real data application was also considered to show the flexibility of the new control chart in practice. Based on the results found, we verified that is quite important to identify which distribution fits the data better, and so to use its respective control limits in monitoring the quality characteristic of interest.

Finally, an anonymous reviewer has indicated an interesting topic of future research, namely: control charts for monitoring fractions and proportions if control variables (covariates) affect the process; that is, regression control charts to monitor variables of interest that are related to control variables. There are interesting proposals in the literature and the reader is referred to Bayer *et al.* (2018) and Ali, Asghar, and Shah (2024), among others. So, one can introduce a regression control chart based on the two-parameter standard two-sided power distribution. An in-depth investigation of this is beyond the scope of the current paper but certainly corresponds to a very interesting topic for future research.

Acknowledgments

The author deeply thanks the anonymous reviewer for the comments and suggestions, which have improved considerably the current work. The author also acknowledges the financial support of the Brazilian agency *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq grant 303554/2022–3).

References

Ali S, Asghar M, Shah I (2024). "Generalized Linear Model Based Gamma Control Chart." *Quality and Reliability Engineering International*, **40**, 699–711. doi:https://doi.org/10.1002/qre.3437.

Bayer FM, Tondolo CM, Muller FM (2018). "Beta Regression Control Chart for Monitoring Fractions and Proportions." *Computers and Industrial Engineering*, **119**, 416–426. doi:https://10.1016/j.cie.2018.04.006.

Brownlee KA (1965). *Statistical Theory and Methodology in Science and Engineering*. John Wiley and Sons, New York. ISBN: 978-047-111-355-3.

Chen G (1998). "An Improved *p* Chart through Simple Adjustments." *Journal of Quality Technology*, **30**, 142–151. doi:https://doi.org/10.1080/00224065.1998.11979833.

Chen JT (2013). "Design of Cumulative Count of Conforming Charts for High Yield Processes Based on Average Number of Items Inspected." *International Journal of Quality and Reliability Management*, **30**, 942–957. doi:https://doi.org/10.1108/IJQRM-01-2011-0014.

Draper NR, Smith H (1998). *Applied Regression Analysis*. John Wiley and Sons, New York. doi:https://doi.org/10.1002/9781118625590. ISBN: 978-047-117-082-2.

Graham MA, Mukherjee A, Chakraborti A (2017). "Design and Implementation Issues for a Class of Distribution-Free Phase II EWMA Exceedance Control Charts." *International Journal of Production Research*, **55**, 2397–2430. doi:https://doi.org/10.1080/00207543.2016.1249428.

Hossain MP, Riaz M (2021). "On Designing a New VEWMA Control Chart for Efficient Process Monitoring." *Computers and Industrial Engineering*, **162**, 107751. doi:https://doi.org/10.1016/j.cie.2021.107751.

Joekes S, Barbosa EP (2013). "An Improved Attribute Control Chart for Monitoring Non-Conforming Proportion in High Quality Processes." *Control Engineering Practice*, **21**, 407–412. doi:https://doi.org/10.1016/j.conengprac.2012.12.005.

Jones MC (2009). "Kumaraswamy's Distribution: A Beta-Type Distribution with Some Tractability Advantages." *Statistical Methodology*, **6**, 70–91. doi:https://doi.org/10.1016/j.stamet.2008.04.001.

Kotz S, van Dorp JR (2004). *Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications.* World Scientific Press, Singapore. doi:https://doi.org/10.1142/5720. ISBN: 978-981-256-115-2.

Kumaraswamy P (1980). "Generalized Probability Density-Function for Double-Bounded Random-Processes." *Journal of Hydrology*, **46**, 79–88. doi:https://doi.org/10.1016/0022-1694(80)90036-0.

Lima-Filho LMA, Bayer FM (2021). "Kumaraswamy Control Chart for Monitoring Double Bounded Environmental Data." *Communications in Statistics – Simulation and Computation*, **50**, 2513–2528. doi:https://doi.org/10.1080/03610918.2019.1635159.

Lima-Filho LMA, Bayer FM, da Silva AM (2021). "Control Chart to Monitor Circular Data." *Quality and Reliability Engineering International*, **37**, 966–983. doi:https://doi.org/10.1002/qre.2776.

Montgomery DC (2019). *Introduction to Statistical Quality Control, 8th ed.* John Wiley and Sons, New York. ISBN: 978-1-119-39930-8.

Moraes DAO, Oliveira FLP, Quinino RC, Duczmal LH (2014). "Self-Oriented Control Charts for Efficient Monitoring of Mean Vectors." *Computers and Industrial Engineering*, **75**, 102–115. doi:https://doi.org/10.1016/j.cie.2014.06.008.

Mukherjee A, Chakraborti S (2012). "A Distribution-Free Control Chart for the Joint Monitoring of Location and Scale." *Quality and Reliability Engineering International*, **28**, 335–352. doi:https://doi.org/10.1002/qre.1249.

R Core Team (2023). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sagrillo M, Guerra RR, Machado R, Bayer FM (2023). "A Generalized Control Chart for Anomaly Detection in SAR Imagery." *Computers and Industrial Engineering*, **177**, 109030. doi:https://doi.org/10.1016/j.cie.2023.109030.

Sant'Anna AMO, ten Caten CS (2012). "Beta Control Charts for Monitoring Fraction Data." *Expert Systems with Applications*, **39**, 10236–10243. doi:https://doi.org/10.1016/j.eswa.2012.02.146.

Sim CH, Lim MH (2008). "Attribute Charts for Zero-Inflated Process." *Communications in Statistics – Simulation and Computation*, **37**, 1440–1452. doi:https://doi.org/10.1080/03610910801983145.

van Dorp JR, Kotz S (2002). "The Standard Two-Sided Power Distribution and Its Properties: With Applications in Financial Engineering." *The American Statistician*, **59**, 90–99. URL https://www.jstor.org/stable/3087377.

Wang H (2009). "Comparison of p Control Charts for Low Defective Rate." *Computational Statistics and Data Analysis*, **53**, 4210–4220. doi:https://doi.org/10.1016/j.csda.2009.05.024.

Winterbottom A (1993). "Simple Adjustments to Improve Control Limits on Attribute Charts." *Quality and Reliability Engineering International*, **9**, 105–109. doi:https://doi.org/10.1002/qre.4680090207.

Affiliation:

Artur J. Lemonte
Departamento de Estatística, CCET
Universidade Federal do Rio Grande do Norte
Lagoa Nova, Natal/RN 59078970, Brazil
E-mail: arturlemonte@gmail.com

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 54 *Submitted:* 2024-08-02
2025 *Accepted:* 2025-03-03