



## MATHEMATICAL SCIENCES

# New statistical process control charts for overdispersed count data based on the Bell distribution

LAION L. BOAVENTURA, PAULO H. FERREIRA, ROSEMEIRE L. FIACCONE,  
PEDRO L. RAMOS & FRANCISCO LOUZADA

**Abstract:** Poisson distribution is a popular discrete model used to describe counting information, from which traditional control charts involving count data, such as the  $c$  and  $u$  charts, have been established in the literature. However, several studies recognize the need for alternative control charts that allow for data overdispersion, which can be encountered in many fields, including ecology, healthcare, industry, and others. The Bell distribution, recently proposed by Castellares et al. (2018), is a particular solution of a multiple Poisson process able to accommodate overdispersed data. It can be used as an alternative to the usual Poisson (which, although not nested in the Bell family, is approached for small values of the Bell distribution) Poisson, negative binomial, and COM-Poisson distributions for modeling count data in several areas. In this paper, we consider the Bell distribution to introduce two new exciting, and useful statistical control charts for counting processes, which are capable of monitoring count data with overdispersion. The performance of the so-called Bell charts, namely Bell- $c$  and Bell- $u$  charts, is evaluated by the average run length in numerical simulation. Some artificial and real data sets are used to illustrate the applicability of the proposed control charts.

**Key words:** average run length, Bell charts, count data, overdispersion.

## INTRODUCTION

Count data arise in many fields, including biology, ecology, healthcare, marketing, economics, and industry. Overdispersion (variance  $>$  mean) in count data is quite common (Hougaard et al. 1997, Okamura et al. 2012, Coly et al. 2016) and can occur for several reasons, including mechanisms that generate excessive zero counts or censoring (Avcı et al. 2015). In the presence of overdispersed count data, the Poisson model, which is the most popular distribution for analyzing count data and assumes equidispersion (variance = mean), may result in incorrect inference about parameter estimates, standard errors, tests, and confidence intervals (Avcı et al. 2015). Particularly in Statistical Process Control (SPC), a misspecified Poisson model ( $c$  or  $u$  chart) may increase the false alarm rate (Saghir & Lin 2015). Indeed, when dealing with the application of control charts to data that show excessive dispersion, several works (Spiegelhalter 2005, Mohammed & Laney 2006, Albers 2011) express concern about the use of such a structure. This concern is because increased variation may cause multiple data values to be falsely detected as out of control when they are merely false positive.

Sheaffer & Leavenworth (1976) and Kaminsky et al. (1992) considered the negative binomial distribution, while Sellers (2012) used the Conway–Maxwell–Poisson (COM-Poisson) model (*cm<sub>pc</sub>* and *cm<sub>pu</sub>* charts) as flexible alternatives to the Poisson control charts (*c* and *u* charts). In particular, Kaminsky et al. (1992) also developed control charts that plot the total or the average number of events based on the geometric distribution (*g* and *h* charts, respectively). However, as pointed out by Saghir & Lin (2015), there are still very few articles on monitoring dispersed count data. Other proposals involve using compound distributions, e.g., the Poisson-gamma mixture (Cheng & Yu 2013), the shifted (or zero-truncated) generalized Poisson distribution (Famoye 1994), among others.

Recently, Castellares et al. (2018) introduced a one-parameter discrete probability distribution named the Bell distribution. It is infinitely divisible and capable of modeling count data with overdispersion, and with many other attractive properties, e.g., it is a single parameter distribution that belongs to the exponential family of distributions. Although the Poisson distribution is not nested in the Bell family, the Bell distribution approaches the Poisson distribution for small values of the parameter. That is, the Poisson distribution is a limiting case of the Bell distribution which arises when the Bell parameter tends to zero. The Bell distribution motivated us to propose two new statistical control charts to describe the total and the average number of events per inspection unit. The so-called Bell charts, namely Bell-*c* and Bell-*u* charts, represent interesting and useful alternatives to the charts mentioned above when monitoring counting processes with overdispersed data. It is worth pointing out that, in a typical Shewhart control chart, such as the ones proposed in this paper, the main aim is to detect significantly massive shifts in the process parameter. This kind of control chart also ignores historical data, which is why it is called “without memory”.

The remainder of this paper is organized as follows. The “Bell distribution” section revises the Bell distribution and some of its basic properties. The “New control charts” section presents the new attribute control charts based on the Bell distribution. The “Performance evaluation” section provides simulation studies to assess the performance of the proposed Bell charts, even when compared to some traditional control charts. The “Applications” section illustrates the usefulness of the Bell charts through several examples. Finally, the “Final remarks” section concludes the paper with a few remarks and discussions on future works.

## BELL DISTRIBUTION

In this section, we present a brief review of the Bell distribution and some of its properties. We also discuss point estimation via the maximum likelihood estimator for its parameter.

Introduced by Castellares et al. (2018), a random variable  $Y$  is said to be Bell distributed with parameter  $\theta > 0$ , denoted by  $Y \sim \text{Bell}(\theta)$ , if its probability mass function (PMF) is given by

$$P(Y = y|\theta) = \frac{\theta^y e^{-\theta+1} B_y}{y!}, \quad \text{for } y = 0, 1, 2, \dots, \quad (1)$$

where  $B_y = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^y}{k!}$ ,  $y = 0, 1, 2, \dots$ , are the Bell numbers<sup>1</sup> (Bell 1934a, b), which can be computed via the *bell(.)* function of the “numbers” package (Borchers 2018) in R software (R Core Team 2018).

<sup>1</sup>It is worth noting that the Bell number  $B_y$  is the  $y$ -th moment of a Poisson distribution with parameter equal to one; see Remark 1 of Castellares et al. (2018).

If  $Y \sim Bell(\theta)$ , then the mean and variance of  $Y$  are given, respectively, by

$$E[Y] = \theta e^\theta \quad \text{and} \quad \text{Var}[Y] = \theta(1 + \theta)e^\theta. \quad (2)$$

Note that  $\text{Var}[Y]/E[Y] = 1 + \theta > 1, \forall \theta > 0$ . Therefore, the Bell distribution may be a suitable distribution for modeling count data with overdispersion, although it may not accommodate all possible forms of overdispersion; see Remark 5 of Castellares et al. (2018).

Regarding pseudo-random sample generation from  $Y \sim Bell(\theta)$ , we resort to Proposition 3 of Castellares et al. (2018), which stated that the random variable  $Y$  has the same distribution as the sum of  $N$  independent and identically distributed (IID) zero-truncated Poisson (ZTP) random variables with parameter  $\theta > 0$ , where  $N$  is Poisson distributed with parameter  $e^\theta - 1$ . In other words, if  $X_i \stackrel{\text{iid}}{\sim} ZTP(\theta)$ , for  $i = 1, \dots, N$ , with  $N \sim \text{Poisson}(e^\theta - 1)$  and independent of  $\{X_1, \dots, X_N\}$ , then  $Y = X_1 + \dots + X_N \sim Bell(\theta)$ . In R software, we can generate pseudo-random observations from the ZTP distribution by using the `rztpois(.)` function of the “actuar” package (Dutang et al. 2008).

We can easily estimate the parameter  $\theta$  via the maximum likelihood method. By considering the observed sample  $y = (y_1, y_2, \dots, y_n)$  of size  $n$  from  $Y \sim Bell(\theta)$ , we obtain the likelihood function

$$L(\theta|y) = \prod_{i=1}^n \frac{\theta^{y_i} e^{-e^\theta + 1} B_{y_i}}{y_i!} = \frac{\theta^{\sum_{i=1}^n y_i} e^{n(-e^\theta + 1)} \prod_{i=1}^n B_{y_i}}{\prod_{i=1}^n y_i!} \propto \theta^{\sum_{i=1}^n y_i} e^{-ne^\theta}. \quad (3)$$

Castellares et al. (2018) showed that the maximum likelihood estimator (MLE)  $\hat{\theta}$  of  $\theta$  has a closed-form expression and is given by

$$\hat{\theta} = W_0(\bar{Y}), \quad (4)$$

where  $W_0(\cdot)$  is the Lambert  $W$  function (Corless et al. 1996), which can be computed via the  $W(\cdot)$  function of the “LambertW” package (Goerg 2011, Goerg 2016) in R, and  $\bar{Y} = \sum_{i=1}^n Y_i/n$  is the sample mean.

Castellares et al. (2018) also presented an alternative and useful (mainly in a regression modeling framework, in which it is very common to model the mean of the response variable as a function of several other variables, also called explanatory variables or regressors) reparametrization of the Bell distribution, where  $\mu = E[Y] = \theta e^\theta$  and, hence,  $\theta = W_0(\mu)$ . In this case, the PMF, mean and variance of  $Y \sim Bell(\mu)$ ,  $\mu > 0$ , are written, respectively, as

$$P(Y = y|\mu) = \exp\{-e^{W_0(\mu)} + 1\} \frac{[W_0(\mu)]^y B_y}{y!}, \quad \text{for } y = 0, 1, 2, \dots,$$

$$E[Y] = \mu \quad \text{and} \quad \text{Var}[Y] = \mu [1 + W_0(\mu)].$$

Furthermore, it can be easily shown that the MLE  $\hat{\mu}$  of  $\mu$  is the sample mean, i.e.,  $\hat{\mu} = \bar{Y} = \sum_{i=1}^n Y_i/n$ .

## NEW CONTROL CHARTS

Suppose that a process (e.g., an industrial process) generates events (e.g., nonconformities or defects) according to a  $Bell(\theta)$  distribution. Letting  $Y$  denote the number of events per process unit, the PMF of  $Y$  is given by (1). Also, let  $Y_1, Y_2, \dots, Y_n$  be a random sample of size  $n$  from  $Y \sim Bell(\theta)$ , and consider  $T = \sum_{i=1}^n Y_i$  and  $\bar{Y} = T/n = \sum_{i=1}^n Y_i/n$ , which represent, respectively, the total number of events and

the average number of events per unit. Hence, it follows from Equation (2), using the properties of the expected value and variance operators, that the exact mean and variance of  $T$  and  $\bar{Y}$  are given by

$$\begin{aligned} E[T] &= n\theta e^\theta, & \text{Var}[T] &= n\theta(1 + \theta)e^\theta, \\ E[\bar{Y}] &= \theta e^\theta, & \text{Var}[\bar{Y}] &= \frac{\theta(1 + \theta)e^\theta}{n}. \end{aligned} \tag{5}$$

The above quantities can be used to construct the proposed control charts in the usual manner for Shewhart charts. Assuming that a standard value for  $\theta$  is available, the center lines and  $L\sigma$  control limits<sup>2</sup> for each chart - namely, the Bell- $c$  (total number of events per unit) and Bell- $u$  (average number of events per unit) charts - are shown in Table I. It is worth pointing out that since the Bell distribution approaches the Poisson distribution for small values of  $\theta$ , the control limits presented in Table I also approach the control limits for special cases of the  $c$  and  $u$  charts derived from the Poisson distribution. Moreover, similar to the  $c$  chart, the Bell- $c$  chart should not be used in the cases where the sample sizes are unequal, because both the center line and the control limits of the Bell- $c$  chart will vary with the sample size. Therefore, the Bell- $u$  chart is recommended over the Bell- $c$  chart for variable sample size as the former is easier to interpret for this scenario (the center line of the Bell- $u$  chart will not vary across samples).

**Table I. Control limits for the Bell- $c$  and Bell- $u$  charts (standards given). UCL = upper control limit, CL = center line, LCL = lower control limit.**

	Bell- $c$ chart	Bell- $u$ chart
UCL	$n\theta e^\theta + L\sqrt{n\theta(1 + \theta)e^\theta}$	$\theta e^\theta + L\sqrt{\frac{\theta(1 + \theta)e^\theta}{n}}$
CL	$n\theta e^\theta$	$\theta e^\theta$
LCL	$n\theta e^\theta - L\sqrt{n\theta(1 + \theta)e^\theta}$	$\theta e^\theta - L\sqrt{\frac{\theta(1 + \theta)e^\theta}{n}}$

In order to build the so-called Bell charts, considering the lack of knowledge about  $\theta$ , we can apply, for example, the MLE (4), using all the available  $m$  samples, in the place of  $\theta$  (plug-in approach) in the control limits shown in Table I. In this case, the obtained control limits are commonly treated as trial control limits (Montgomery 2013).

Since the  $\theta$  parameter contains little or no direct interpretation, it can be difficult for industry professionals, such as engineers and technicians, to provide a standard (or reference) value. In order to overcome this difficulty, we can use a convenient reparametrization of the Bell distribution in terms of the mean  $\mu = E[Y] = \theta e^\theta$ . This parametrization was described at the end of the previous section. The derived control limits for the proposed Bell charts, considering this alternative reparametrization, are displayed in Table II. When no standard is given, then  $\mu$  can be estimated by the overall sample mean, that is,  $\hat{\mu} = \bar{\bar{Y}} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n Y_{ij}$  (in the cases of the Bell- $c$  and Bell- $u$  charts with equal sample

<sup>2</sup>In the usual Six Sigma quality control program,  $L = 3$ .

sizes) or  $\hat{\mu} = \bar{Y} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} Y_{ij}}{\sum_{j=1}^m n_j}$  (in the case of the Bell-*u* chart with unequal sample sizes <sup>3</sup>), where  $Y_{ij}$  denotes the number of events for the *i*-th inspection unit of the *j*-th sample. Both estimators may be easily proved to be unbiased estimators for  $\mu$ .

Therefore, because of their practical advantage as well as simplicity, we shall hereafter consider the control limits provided in Table II.

**Table II. Control limits for the Bell-*c* and Bell-*u* charts, considering a different reparametrization of the Bell distribution (standards given).**

	Bell- <i>c</i> chart	Bell- <i>u</i> chart
UCL	$n\mu + L\sqrt{n\mu [1 + W_0(\mu)]}$	$\mu + L\sqrt{\frac{\mu [1 + W_0(\mu)]}{n}}$
CL	$n\mu$	$\mu$
LCL	$n\mu - L\sqrt{n\mu [1 + W_0(\mu)]}$	$\mu - L\sqrt{\frac{\mu [1 + W_0(\mu)]}{n}}$

## PERFORMANCE EVALUATION

In this section, we conduct Monte Carlo (MC) simulation studies to assess the performance of the proposed Bell charts, as well as of some existing/traditional control charts for count data (namely, the Poisson-based *c* and *u* charts, and the COM-Poisson-based *cmpc* and *cmpu* charts) when the true data-generating process is Bell distributed. All simulations and computations were performed using the R software version 3.6.1. Interested readers can email the authors for the corresponding R codes.

The average run length (ARL) is a measure commonly used to evaluate the performance of control charts. The in-control ARL, also denoted as  $ARL_0$ , is defined as the average number of samples (or monitoring points) before a signal is given (that is, a single point falls outside the control limits), assuming that the process is in control; while the out-of-control ARL (or  $ARL_1$ ) is the average number of samples that are taken until a mean shift is observed when the process is out of control (Saghir & Lin 2015).

Let us assume that  $Y \sim \text{Bell}(\mu)$  comes from a process with in-control average nonconformities, and let  $\mu_s$  be the shifted mean nonconformities parameter after a shift occurs in  $\mu$ , that is,  $Y \sim \text{Bell}(\mu_s)$ . For the proposed Bell charts, the  $ARL_0$  is defined as

$$ARL_0 = \frac{1}{\alpha},$$

<sup>3</sup>Note that, in this case, the control limits of the Bell-*u* chart will also vary with the sample size, since we should replace  $n$  by  $n_j$  in their expressions. But alternatively, Montgomery (2013) suggested to base the control limit calculations on an average sample size  $\bar{n}$ , which results in constant limits and is particularly helpful if the charts will be presented to management.

where  $\alpha = 1 - P(\text{LCL} < Y < \text{UCL} | \mu)$ . While the  $\text{ARL}_1$  is given by

$$\text{ARL}_1 = \frac{1}{1 - \beta},$$

with  $\beta = P(\text{LCL} < Y < \text{UCL} | \mu_s)$ .

For instance, in the usual Six Sigma programs,  $\alpha = 0.0027$  and, thus,  $\text{ARL}_0 = 1/0.0027 \approx 370$ . That is, even if the process is in control, an out-of-control signal will be given every 370 samples, on average (Montgomery 2013). On the other hand,  $\text{ARL}_1$  values near one are desirable, mainly for large-size shifts in a process mean (Shewhart control charts).

### In-control ARL

Without loss of generality, in this subsection we consider a Bell process with mean nonconformities rates:  $\mu = 3$  and 15, as well as two different values for the probability of false alarm:  $\alpha = 0.01$  and 0.1. These  $\alpha$  values correspond, respectively, to  $\text{ARL}_0 = 100$  and 10. We also assume three sample sizes for the process:  $n = 50, 100$  and 500 (or  $\bar{n} = 50, 100$  and 500, in the case of the Bell- $u$  chart with variable sample size <sup>4</sup>), even as processes with three different sample quantities in phase 1:  $m = 20, 50$  and 100.

Figures 1-4 and A1-A4 (see Appendix A) show the results obtained from 1,000 MC simulations (or replicates) with  $m^* = 5,000$  phase 2 samples each <sup>5</sup>, performed for each scenario studied. That is, by varying the number of phase 1 samples, the sample size, the Bell distribution parameter, and the probability of false alarm. In particular, the  $\text{ARL}_0$  results in Figures 1-4 concern the Bell- $c$  chart with equal sample sizes, while Figures A1-A4 contain the  $\text{ARL}_0$  values of the Bell- $u$  chart with unequal sample sizes. In all these figures, the dashed line represents the nominal  $\text{ARL}_0$  value, and the asterisk inside the boxplot indicates the average  $\text{ARL}_0$  estimate of the 1,000 simulations.

Despite some slight to moderate discrepancies from the target  $\text{ARL}_0$  value in some cases, which is indeed expected due to the parameter estimation effect <sup>6</sup>, the results presented in Figures 1-4 and A1-A4 seem to indicate the good performance of the proposed Bell charts. Note that, in general, the results improve, i.e., the  $\text{ARL}_0$  values approach the nominal one when both  $m$  and  $n$  (or  $\bar{n}$ ) increase.

### Out-of-control ARL

In this subsection, we evaluate the detection ability of the proposed Bell charts by means of  $\text{ARL}_1$ , for the same scenarios as in the "In-control ARL" subsection. Due to space constraints, we consider only shifts at three levels that represent percentage increases  $p$  in the nonconformities rate  $\mu$  of the Bell process. The hypothesized levels are as follows:  $p = 0.5\%$  ( $\mu_s = 3.015$  and 15.075),  $1\%$  ( $\mu_s = 3.03$  and 15.15) and  $10\%$  ( $\mu_s = 3.3$  and 16.5).

<sup>4</sup>Here, each sample size  $n_j$  is obtained by  $n_j = \bar{n} + k_j$ , for  $j = 1, \dots, m$ , where  $k_j$  is a random integer constant ranging over the interval  $[-\bar{n}/5; \bar{n}/5]$ .

<sup>5</sup>As pointed out by Montgomery (2013), the phases 1 and 2 of control chart application have different and distinct objectives. In phase 1, a set of process data is gathered and analyzed at once, constructing trial control limits to determine whether the process was in control when the initial  $m$  samples were collected (retrospective analysis). On the other hand, in phase 2, the chart built from a "clean" set of process data exhibiting control (reliable control limits), is used for monitoring future production (prospective process monitoring).

<sup>6</sup>For a further discussion on this relevant issue, see, e.g., the review paper by Jensen et al. (2006).

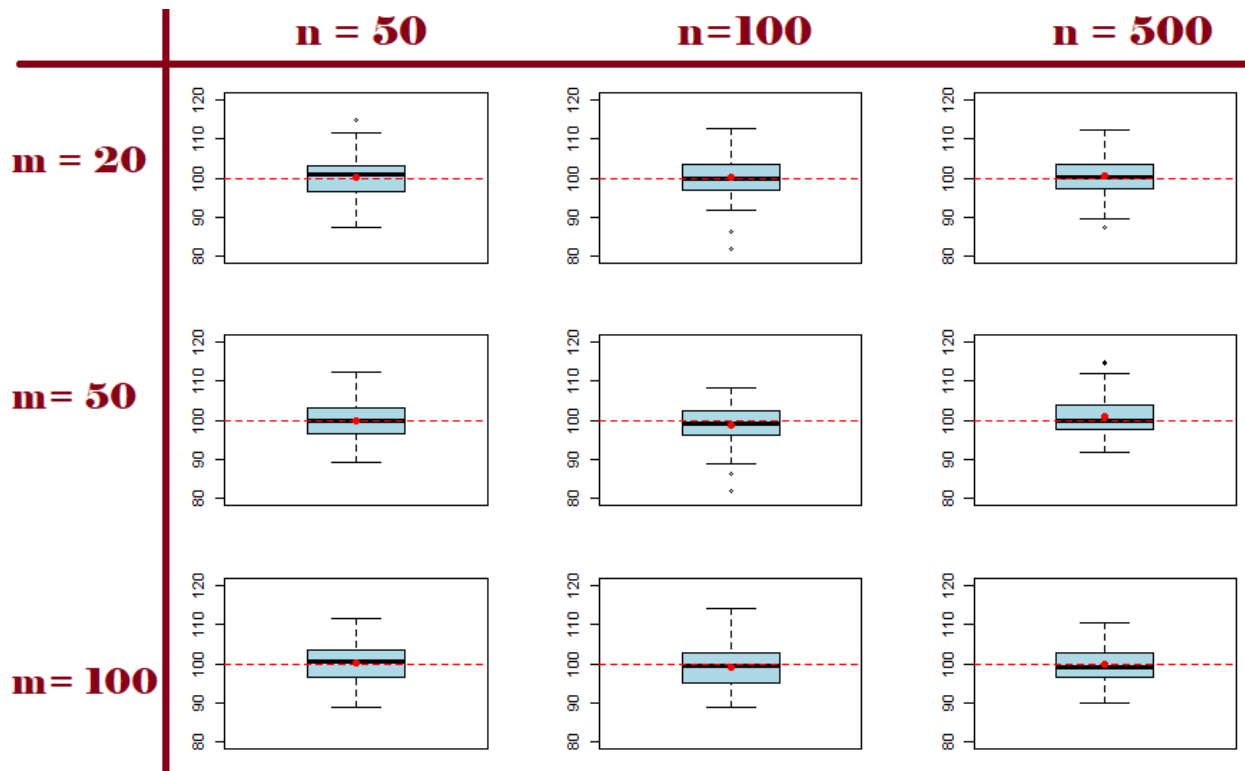


Figure 1. The  $ARL_0$  values of the Bell-c chart for various  $m$  and  $n$  ( $\mu = 3$  and  $\alpha = 0.01$ ).

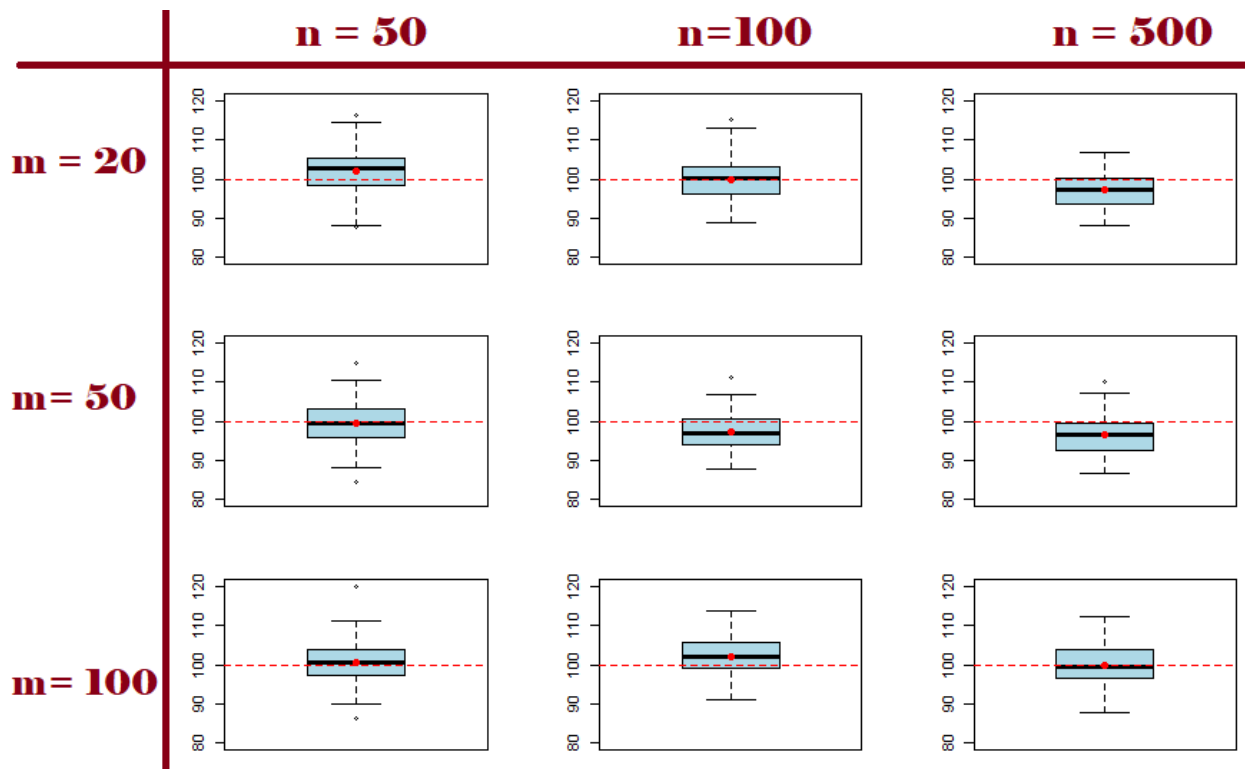


Figure 2. The  $ARL_0$  values of the Bell-c chart for various  $m$  and  $n$  ( $\mu = 15$  and  $\alpha = 0.01$ ).

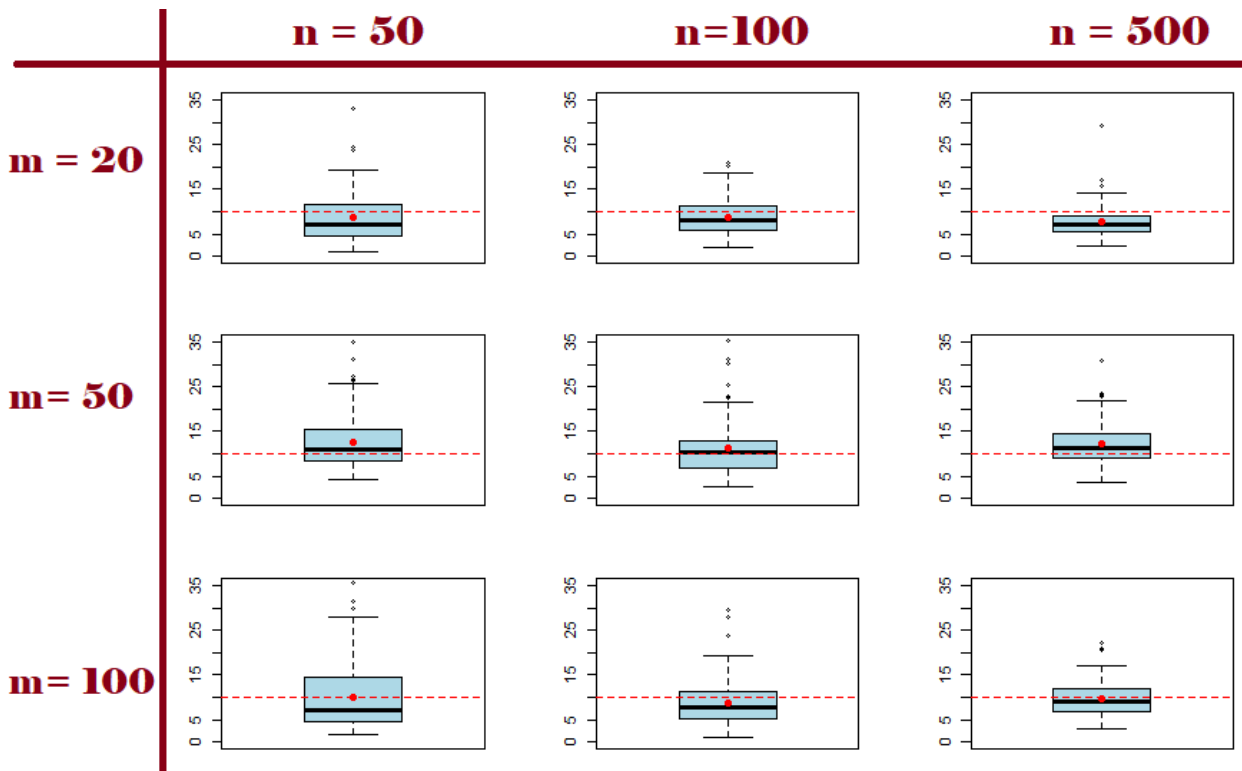


Figure 3. The  $ARL_0$  values of the Bell-c chart for various  $m$  and  $n$  ( $\mu = 3$  and  $\alpha = 0.1$ ).

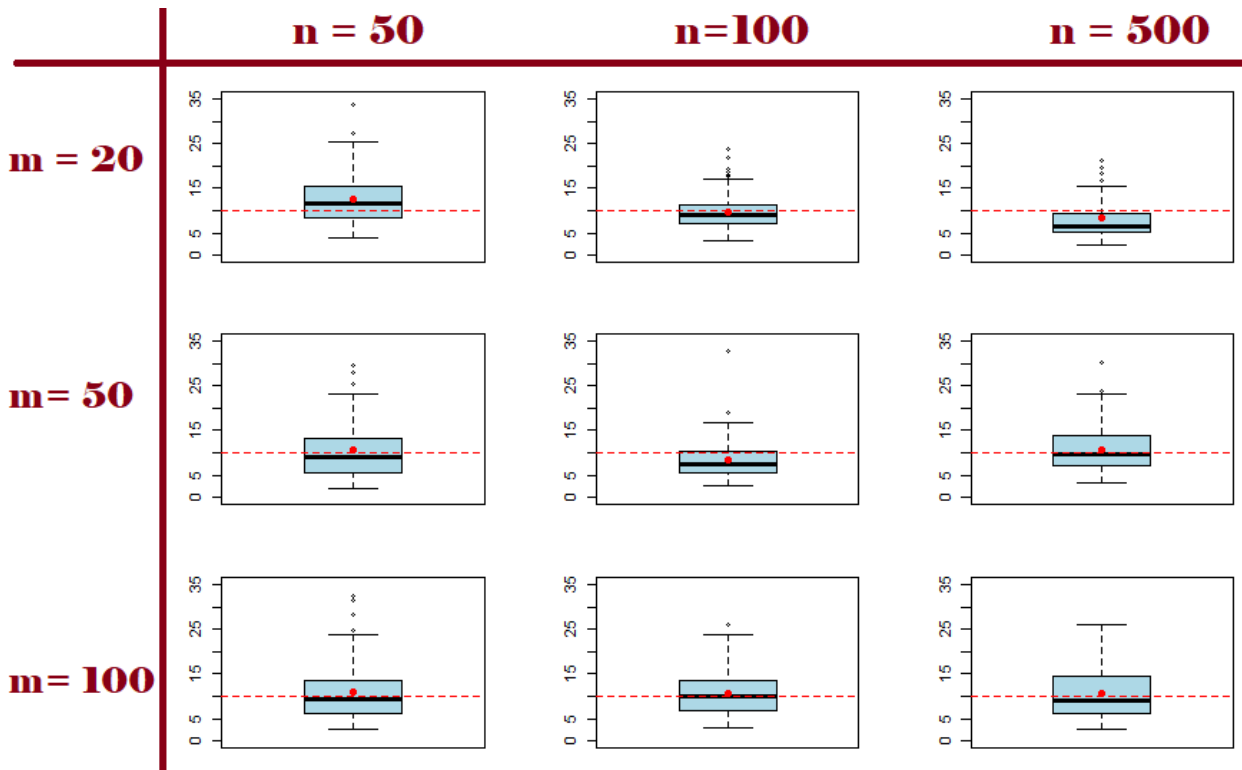


Figure 4. The  $ARL_0$  values of the Bell-c chart for various  $m$  and  $n$  ( $\mu = 15$  and  $\alpha = 0.1$ ).



The  $ARL_1$  values from the 1,000 MC simulations with  $m^* = 5,000$  samples each, are presented in Tables III and AI (see Appendix A) for the Bell- $c$  chart with equal sample sizes and the Bell- $u$  chart with unequal sample sizes, respectively. From these tables, it can be clearly seen that the  $ARL_1$  on average decreases with increasing  $p$ . Note also that, for both control charts based on the Bell distribution, the  $ARL_1$  values are quite close to one when  $p \geq 1\%$ , and regardless of the  $m, n$  (or  $\bar{n}$ ),  $\mu$  and  $\alpha$  values.

**Table III. Mean (standard deviation in parentheses) values of  $ARL_1$  of the Bell- $c$  chart, for the different scenarios studied.**

		$n = 50$			$n = 100$			$n = 500$			
		$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	
$m = 20$	$\mu = 3$	$\alpha = 0.1$	12.442 (0.112)	1.272 (0.011)	1.013 (0.023)	12.942 (0.132)	1.178 (0.011)	1.015 (0.024)	14.331 (0.111)	1.251 (0.012)	1.011 (0.011)
		$\alpha = 0.01$	117.212 (0.223)	1.244 (0.011)	1.021 (0.024)	112.732 (0.263)	1.162 (0.011)	1.016 (0.023)	110.213 (0.314)	1.267 (0.012)	1.013 (0.011)
	$\mu = 15$	$\alpha = 0.1$	17.131 (0.114)	1.151 (0.012)	1.063 (0.022)	12.622 (0.211)	1.265 (0.011)	1.011 (0.022)	11.972 (0.111)	1.145 (0.010)	1.023 (0.011)
		$\alpha = 0.01$	119.634 (0.411)	1.242 (0.010)	1.092 (0.021)	117.000 (0.410)	1.171 (0.011)	1.027 (0.023)	118.303 (0.382)	1.263 (0.012)	1.018 (0.014)
$m = 50$	$\mu = 3$	$\alpha = 0.1$	13.721 (0.110)	1.127 (0.011)	1.043 (0.023)	17.092 (0.111)	1.264 (0.014)	1.018 (0.023)	19.191 (0.112)	1.233 (0.011)	1.006 (0.012)
		$\alpha = 0.01$	123.232 (0.411)	1.257 (0.014)	1.021 (0.022)	117.500 (0.332)	1.265 (0.011)	1.016 (0.024)	112.621 (0.373)	1.237 (0.012)	1.002 (0.011)
	$\mu = 15$	$\alpha = 0.1$	12.202 (0.111)	1.171 (0.013)	1.033 (0.024)	12.210 (0.104)	1.227 (0.011)	1.018 (0.023)	11.993 (0.214)	1.221 (0.013)	1.011 (0.012)
		$\alpha = 0.01$	112.532 (0.311)	1.103 (0.014)	1.092 (0.023)	119.124 (0.282)	1.174 (0.011)	1.019 (0.024)	112.6 (0.423)	1.249 (0.012)	1.014 (0.010)
$m = 100$	$\mu = 3$	$\alpha = 0.1$	11.922 (0.111)	1.114 (0.014)	1.005 (0.023)	11.961 (0.102)	1.097 (0.012)	1.008 (0.024)	10.992 (0.113)	1.281 (0.012)	1.009 (0.011)
		$\alpha = 0.01$	112.124 (0.410)	1.194 (0.012)	1.032 (0.024)	101.812 (0.352)	1.267 (0.011)	1.018 (0.024)	101.924 (0.253)	1.152 (0.012)	1.011 (0.012)
	$\mu = 15$	$\alpha = 0.1$	12.304 (0.110)	1.113 (0.012)	1.042 (0.024)	11.994 (0.213)	1.154 (0.012)	1.009 (0.021)	12.261 (0.224)	1.219 (0.012)	1.004 (0.013)
		$\alpha = 0.01$	117.311 (0.182)	1.124 (0.013)	1.008 (0.024)	116.223 (0.213)	1.227 (0.012)	1.009 (0.022)	102.172 (0.214)	1.107 (0.014)	1.000 (0.012)

### The impact of Bell data on some standard control charts

The Bell control chart theory introduced here may be applied in practical situations as a useful and exciting alternative to the well-known  $c$  and  $u$  charts developed via the Poisson assumption, and the  $cm_{pc}$  and  $cm_{pu}$  charts derived from the COM-Poisson distribution, among others, when data are overdispersed. Hence, in-control count data with overdispersion can be modeled well via the proposed Bell charts.

In this subsection, we apply the above-mentioned Poisson- and COM-Poisson-based control charts to sample data generated from the Bell distribution. The aim is to investigate, employing simulations, the performance (in terms of  $ARL_0$  and  $ARL_1$ ) of these well-known control charts when they are applied to the Bell processes. For comparison purposes, we also use the Bell- $c$  and Bell- $u$

charts for the cases when the samples are of equal and unequal sizes. The simulations were conducted using the same settings as described in the previous subsections. However, due to space and time limitations, we consider only  $m = 100$  and  $n = 100$  (or  $\bar{n} = 100$ ). Of course, it may also be of interest to examine further the performance of the Bell-based control charts to analyze count data generated from other distributions, including the Poisson and COM-Poisson distributions with some level of dispersion. Although we provide some preliminary results (indeed, based on a single artificial sample only) in the “Poisson data” and “COM-Poisson data” subsections, this issue will be better addressed in our future work.

Tables IV and V display, respectively, the  $ARL_0$  and  $ARL_1$  results for the control charts with equal sample sizes (or  $c$ -type control charts). While the performance measures for the control charts with unequal sample sizes (or  $u$ -type control charts) are shown in Tables AII and AIII of Appendix A. From these tables, it can be seen that, regardless of the scenario configuration, the Poisson-based control charts produced poor results (i.e., gave many false alarms) for the in-control Bell samples. This result is in agreement with some authors, e.g., Saghir & Lin (2015). On the other hand, the COM-Poisson-based control charts generally provided reasonable to good results for both the in-control and out-of-control Bell samples (notice that such results are sometimes close to the ones obtained from the Bell-based control charts). This finding is somewhat expected as the COM-Poisson distribution has one extra parameter which adds some flexibility to the model and, consequently, the corresponding control chart.

Finally, it is worth pointing out that we obtained the CL and control limits (UCL and LCL) by finding the respective mean and standard deviation (SD) for each case, thus calculating the limits by the “Mean  $\pm L \times SD$ ” rule of thumb. For the Poisson and Bell distributions, the mean is easily determined. Meanwhile, the COM-Poisson summary statistics have complicated formulas, but can be readily computed using the “compoisson” package (Dunn 2012) in R.

**Table IV.** Mean (standard deviation in parentheses) values of  $ARL_0$  of the competitor  $c$ -type control charts when the true data-generating process is Bell( $\mu$ ) distributed, for some  $\mu$  and  $\alpha$  values ( $m=100$  and  $n=100$ ).

Control chart	$\mu = 3$		$\mu = 15$	
	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.1$
Bell-c	97.0814 (13.4512)	9.8157 (0.4083)	101.1592 (15.0990)	9.9066 (0.3851)
c	13.8551 (0.6888)	4.0038 (0.0965)	7.3268 (0.2633)	2.9126 (0.0558)
cmpc	117.8809 (17.9508)	10.6975 (0.4711)	105.1765 (16.4531)	9.4851 (0.3732)

**Table V.** Mean (standard deviation in parentheses) values of  $ARL_1$  of the competitor  $c$ -type control charts when the true data-generating process is Bell( $\mu$ ) distributed, for some  $\mu$ ,  $\alpha$  and  $p$  values ( $m=100$  and  $n=100$ ).

Control chart	$\mu = 3$					
	$\alpha = 0.01$			$\alpha = 0.1$		
	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$
Bell-c	91.2050 (12.6312)	74.9862 (9.7344)	12.6206 (0.6236)	9.7069 (0.3904)	9.2611 (0.3662)	2.8119 (0.0549)
$c$	13.5571 (0.6768)	12.4846 (0.6042)	4.1535 (0.1048)	3.9747 (0.0929)	3.8935 (0.0908)	1.8493 (0.0243)
$cmpc$	113.8549 (17.8464)	86.4678 (12.2304)	14.4834 (0.7670)	10.9720 (0.4775)	10.2203 (0.4228)	3.0474 (0.0633)
Control chart	$\mu = 15$					
	$\alpha = 0.01$			$\alpha = 0.1$		
	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$
Bell-c	91.9277 (12.4999)	50.3049 (5.2314)	2.6450 (0.0472)	9.8956 (0.4038)	8.3245 (0.3101)	1.5320 (0.0158)
$c$	7.1346 (0.2476)	5.5831 (0.1671)	1.3094 (0.0105)	2.9266 (0.0573)	2.7210 (0.0505)	1.1663 (0.0066)
$cmpc$	102.9636 (15.3876)	27.0492 (2.0920)	2.8850 (0.0543)	9.8275 (0.4035)	9.1967 (0.3691)	1.5444 (0.0160)

## APPLICATIONS

It is well-known that there are practical situations in which the nature of the production process allows an item or product to contain several nonconformities (defects) and not be classified as nonconforming. For example, the manufacture of personal computers might have one or more very minor flaws in the cabinet finish, but since these flaws do not seriously affect the unit's functional operation, it could be classified as conforming (Montgomery 2013). Thus, in this section, we apply SPC charts to inspect the total or average number of defects per sample, and determine if the process is in control.

### Simulated data examples

Here, we consider three artificial data sets containing some levels of overdispersion to illustrate the usefulness of the Bell- $c$  and Bell- $u$  charts, as well as their ability to produce bounds comparable to those established by the classical SPC theory.

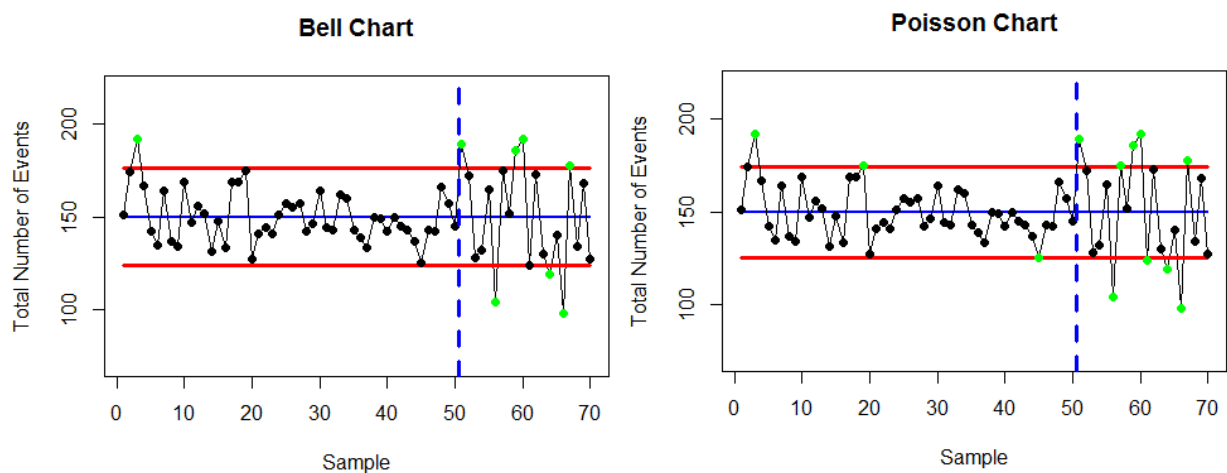
In the cases where the samples are of equal sizes, the control charts chosen for comparison were the traditional  $c$  chart and the  $cmpc$  chart. The  $c$  chart assumes that the Poisson distribution well

models the number of defects per inspection unit. However, according to Sellers (2012), the  $c$  chart, while having a better performance for large samples, does not work well in the cases of overdispersion. Thus, the author developed a control chart using a more general count distribution that relaxes the equidispersion assumption of the Poisson distribution. The so-called  $cmpc$  chart, based on the assumption of a COM-Poisson distribution, has shown excellent results when fitting overdispersed count data. Therefore, it can also be used as a reference tool to compare the performance of the Bell- $c$  chart in this work. For the cases when the samples are of unequal sizes, the control charts selected for the comparison were the  $u$  chart and the  $cmpu$  chart.

Similar to the “Performance evaluation” section, all the analyses were done using the R programming language, and it was assumed  $L = 3$ .

### Poisson data

In this application, we generated  $m = 50$  samples of size  $n = 100$  from a Poisson distribution with parameter  $\lambda = 3$ . The idea here was to compare the performance of the Bell- $c$  and  $c$  charts, mainly in phase 2. More specifically, in phase 1, we calculated the control limits of both charts from the aforementioned  $m = 50$  samples. While in phase 2 we simulated new  $m^* = 70$  samples of size  $n = 100$  from the same distribution (Poisson), but with the last 20 samples disturbed by  $\lambda_s = 3.3$ . Then, for both estimated control charts, the occurrence (or not) of false alarms in the first 50 samples and the number of runs until the first of the 20 nonconforming samples could be detected.



**Figure 5.** Performance comparison of Bell- $c$  and  $c$  charts, both constructed from Poisson distribution samples (phase 2).

Figure 5 illustrates this application, where the points (samples) to the right of the blue dashed vertical line are the disturbing observations (out-of-control samples). Furthermore, the red solid horizontal lines represent the UCL and LCL, the blue solid horizontal line indicates the CL, and the green points correspond to the out-of-control signals.

This figure points out for this particular artificial example that the Bell-based control chart produced better results than the Poisson-based control chart for the in-control samples, even though

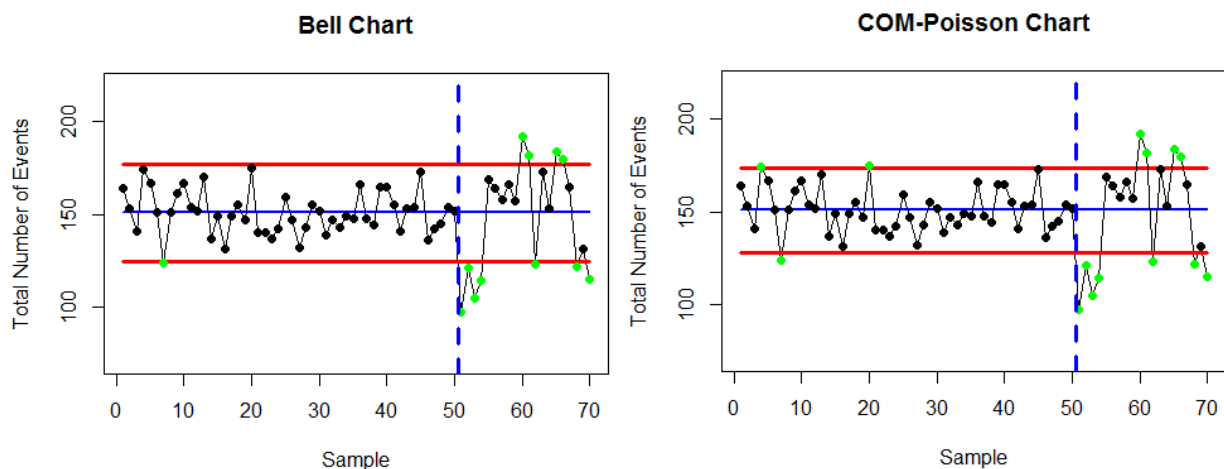
the true data-generating process is not Bell distributed. This was due to the fact that the Bell-c chart identified a single false alarm, while the  $c$  chart itself found three.

Finally, for the samples that are known to be out-of-control, both charts performed similarly, showing that the Bell-c chart may be a useful alternative to the traditional  $c$  chart.

### COM-Poisson data

As in the “Poisson data” subsection, in this application we generated  $m = 50$  samples of equal size  $n = 100$ , but this time, from a COM-Poisson distribution with parameters  $\lambda = 1.5$  and  $\nu = 0.5$ . Here, the idea also was to compare the performance of the Bell-c and  $cm_{pc}$  charts in phase 2 monitoring. Phases 1 and 2 analyses were carried out in the same way as before, with the exception that the last 20 samples (in a total of  $m^* = 70$  new samples) were perturbed by  $\nu_s = 0.75$ .

Figure 6 shows the estimated control charts applied to the phase 2 samples, from which it can be observed that similar to the previous subsection, the Bell-based control chart provided better results than the COM-Poisson-based control chart for the in-control samples. Although the COM-Poisson represents the true data-generating distribution, the Bell-c chart gave a single false alarm, while the  $cm_{pc}$  chart itself found three. Such excellent performance of the Bell-c chart is maintained when analyzing out-of-control samples, showing that the Bell-c chart can also be a useful, as well as simpler alternative to the  $cm_{pc}$  chart.



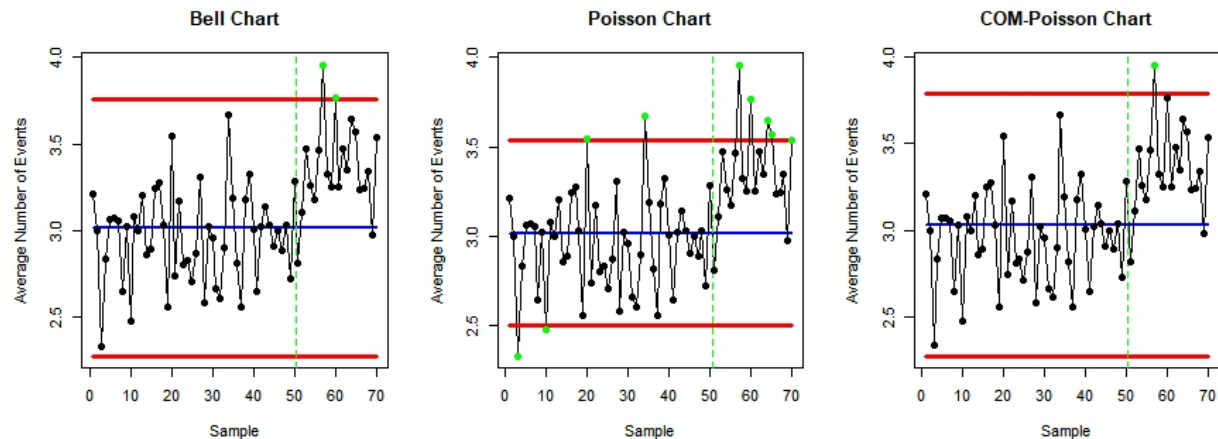
**Figure 6.** Performance comparison of Bell-c and  $cm_{pc}$  charts, both constructed from COM-Poisson distribution samples (phase 2).

### Bell data

In this last application, we simulated  $m = 50$  samples of unequal sizes (with  $\bar{n} = 100$ , and each sample size obtained by following the same approach described in footnote 4) from a Bell distribution with parameter  $\mu = 3$ , and compared the performance of the Bell- $u$  chart against the  $u$  chart and  $cm_{pu}$  chart. Once again, as in the previous subsections, we considered the application to phase 2 samples. The idea here was to evaluate both Poisson- and COM-Poisson-based control

charts, compared to the Bell-based control chart performance when the data are actually from a Bell distribution.

From the analysis of Figure 7, it can be noticed that, in the case of conforming (in-control) samples, the Bell- $u$  chart did not produce false alarms as the  $cm_{pu}$  chart, while the  $u$  chart identified four false alarms. However, the Bell-based control chart performed slightly better than the COM-Poisson-based control chart when in the presence of nonconforming samples, where the Bell out-of-control samples were generated with  $\mu_s = 3.3$ .



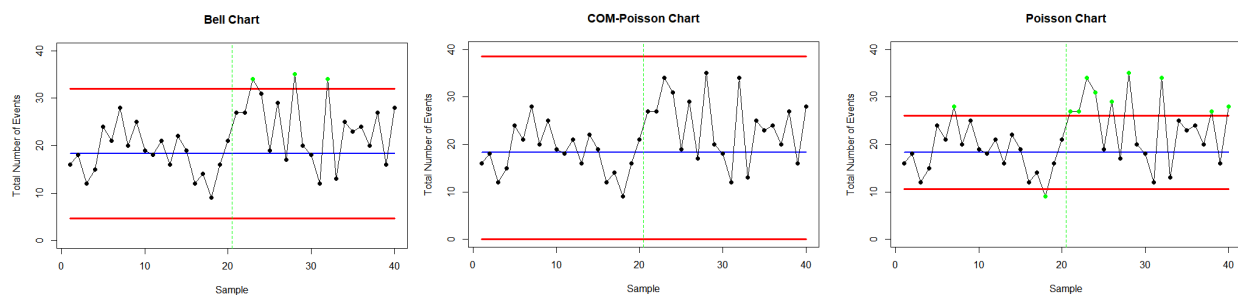
**Figure 7.** Performance comparison of Bell- $u$ ,  $u$  and  $cm_{pu}$  charts, all constructed from Bell distribution samples (phase 2).

### Real data example

In this subsection, we considered the Bell- $c$ ,  $cm_{pc}$  and  $c$  charts to analyze the real data set related to the number of nonconformities in samples of 100 printed circuit boards; see Chapter 7 of Montgomery (2013). The data analysis initially comprehends 20 samples with overdispersion (variance/mean  $\approx 1.21$ ), which suggests that the process can be described by a COM-Poisson distribution with  $\hat{\lambda} = 2.87$  and  $\hat{\nu} = 0.37$  (Sellers 2012). Further, Alevizakos & Koukouvinos (2022) considered 20 additional samples, also with overdispersion (variance/mean  $\approx 2.00$ ), simulated from a COM-Poisson distribution with  $\lambda = 3.09$  and  $\nu = 0.37$ , which implies a deterioration in the process; for more details, see Table 10 of Alevizakos & Koukouvinos (2022). Here, we checked the goodness-of-fit of the Bell distribution ( $\hat{\mu} \approx 20$ ) through the Pearson's  $\chi^2$  test application with the first 20 samples, where we found p-value = 0.57. We also used the same test procedure for the last 20 samples, in which we obtained a p-value = 0.98 ( $\hat{\mu} \approx 24$ ).

Figure 8 shows the estimated Bell- $c$ ,  $cm_{pc}$ , and  $c$  charts, applied to the printed circuit boards data set, with control limits calculated from the initial 20 samples. That is, the left part of the green dashed vertical line is precisely the monitoring of the first 20 samples (phase 1), while the right part of this line represents the monitoring of the remaining 20 samples, which did not participate in the construction of the control limits (phase 2).

It can be observed from Figure 8 that the Bell-c chart<sup>7</sup> detected the process mean shift upwards as soon as it occurred. This detection was unlike the *cm<sub>pc</sub>* chart, which did not identify any nonconformity in the process, neither on the left side nor on the right side of the green dashed vertical line. The *c* chart, as expected, produced false alarms just to the left of this line. This result is because, as we know, the *c* chart is not flexible enough to accommodate overdispersed data. Furthermore, it is essential to note that Alevizakos & Koukouvinos (2022) also considered the data set analyzed by Sellers (2012); they had already observed that the proposed chart did not point to declines over the UCL, which is undesirable. In our case, with the Bell-c chart, the first warning appeared on unit #22, with three units outside the UCL correctly showing process deterioration. This result confirms that the Bell-c chart is an excellent alternative to usual control charts, such as the *cm<sub>pc</sub>* and *c* charts.



**Figure 8.** Performance comparison of Bell-c, *cm<sub>pc</sub>* and *c* charts, constructed from the printed circuit boards data set.

## FINAL REMARKS

The one-parameter discrete Bell distribution, introduced by Castellares et al. (2018), has been established as a viable alternative to the Poisson distribution for analyzing count data with overdispersion. As demonstrated by the examples, we provided in the “Applications” section. The Bell-based control charts can be useful choices/alternatives to the traditional ones (e.g., the *c* and *u* charts derived from the Poisson distribution) for process monitoring of overdispersed count data. Notably, the proposed Bell charts provided satisfactory results (few false alarms and fast process change/shift detection) even when the sample data were generated from the two-parameter (and thus, more complex) COM-Poisson distribution.

In the “Performance evaluation” section, simulation results also demonstrated the excellent performance of the proposed Bell charts under different scenarios (i.e., by varying the number of phase 1 samples, the sample size, the Bell distribution parameter and the probability of false alarm). In this case, the ARL criterion was used to measure control chart performance.

In the applications provided in the “Applications” section, we determined the control limits based on the standard three-sigma rule (Six Sigma program). Nevertheless, alternatively, like in the “Performance evaluation” section, one can obtain the UCL and LCL via a specified type I error rate  $\alpha$ , so that the probability of finding sample points outside the control region is equal to  $\alpha$ . Finally, we considered a real data set related to the number of nonconformities in samples of 100 printed circuit

<sup>7</sup>The R code used for generating this Bell-c chart is available in Appendix B.

boards, which demonstrated that our proposed Bell-c chart outperformed the standard charts when showing process deterioration.

There are many extensions of the current work. For instance, following Braun (1999), Chakraborti & Human (2008), Castagliola & Wu (2012), and Castagliola et al. (2014), we intend to develop optimization designs in order to increase the performance of the Bell control charts when estimating the unknown process parameter  $\mu$  using phase 1 data.

## Acknowledgments

The authors thank the anonymous referees, the editor, and the associate editor for all their useful comments and suggestions which improved remarkably the quality of this article. Laion L. Boaventura acknowledges support from Bahia State Research Foundation (FAPESB, Proc. 19.573.201.5418). Paulo H. Ferreira acknowledges support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant no. 307221/2022-9). The research of Rosemeire L. Fiaccone has been supported by FAPESB (n. app 0071/2016). Francisco Louzada is supported by CNPq (grant no. 301976/2017-1) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, grant no. 2013/07375-0).

## REFERENCES

- ALBERS W. 2011. Control charts for health care monitoring under overdispersion. *Metrika* 74(1): 67-83.
- ALEVIZAKOS V & KOUKOUVINOS C. 2022. A progressive mean control chart for COM-Poisson distribution. *Communications in Statistics-Simulation and Computation* 51(3): 849-867.
- AVCI E, ALTÜRK S & SOYLU EN. 2015. Comparison count regression models for overdispersed alga data. *Int J Res Rev Appl Sci* 25(1): 1-5.
- BELL ET. 1934a. Exponential numbers. *The American Mathematical Monthly* 41(7): 411-419.
- BELL ET. 1934b. Exponential polynomials. *Ann Math* 35(2): 258-277.
- BRAUN WJ. 1999. Run length distributions for estimated attributes charts. *Metrika* 50(2): 121-129.
- CASTAGLIOLA P & WU S. 2012. Design of the c and np charts when the parameters are estimated. *Int J Reliab Qual Saf Eng* 19(02): 1250010.
- CASTAGLIOLA P, WU S, KHOO MBC & CHAKRABORTI S. 2014. Synthetic phase II Shewhart-type attributes control charts when process parameters are estimated. *Qual Reliab Eng Int* 30(3): 315-335.
- CASTELLARES F, FERRARI SLP & LEMONTE AJ. 2018. On the Bell distribution and its associated regression model for count data. *Applied Mathematical Modelling* 56: 172-185.
- CHAKRABORTI S & HUMAN SW. 2008. Properties and performance of the c-chart for attributes data. *J Appl Stat* 35(1): 89-100.
- CHENG SS & YU FJ. 2013. A CUSUM control chart to monitor wafer quality. In: *Proceedings of World Academy of Science, Engineering and Technology*. 78. p. 1300. Citeseer.
- COLY S, YAO AF, ABRIAL D & CHARRAS-GARRIDO M. 2016. Distributions to model overdispersed count data. *Journal de la Société Française de Statistique* 157(2): 39-64.
- CORLESS RM, GONNET GH, HARE DEG, JEFFREY DJ & KNUTH DE. 1996. On the LambertW function. *Adv Comput Math* 5(1): 329-359.
- DUNN J. 2012. *compoisson: Conway-Maxwell-Poisson Distribution*. URL <https://CRAN.R-project.org/package=compoisson>. R package version 0.3.
- DUTANG C, GOULET V & PIGEON M. 2008. *actuar: An R package for Actuarial Science*. *J Stat Soft* 25(7): 1-37.
- FAMOYE F. 1994. Statistical control charts for shifted generalized Poisson distribution. *J Ital Stat Soc* 3(3): 339-354.
- GOERG GM. 2011. Lambert W random variables - a new family of generalized skewed distributions with applications to risk estimation. *Ann Appl Stat* 5(3): 2197-2230.
- GOERG GM. 2016. *LambertW: An R package for Lambert W × F Random Variables*. URL <https://CRAN.R-project.org/package=LambertW>. R package version 0.6.4.
- HOUGAARD P, LEE MLT & WHITMORE GA. 1997. Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. *Biometrics* 53(4): 1225-1238.
- JENSEN WA, JONES-FARMER LA, CHAMP CW & WOODALL WH. 2006. Effects of parameter estimation on control chart



properties: a literature review. *J Qual Technol* 38(4): 349-364.

KAMINSKY FC, BENNEYAN JC, DAVIS RD & BURKE RJ. 1992. Statistical control charts based on a geometric distribution. *J Qual Technol* 24(2): 63-69.

MOHAMMED MA & LANEY D. 2006. Overdispersion in health care performance data: Laney's approach. *BMJ Quality & Safety* 15(5): 383-384.

MONTGOMERY DC. 2013. *Introduction to Statistical Quality Control*. 7<sup>th</sup> ed. John Wiley & Sons.

OKAMURA H, PUNT AE & AMANO T. 2012. A generalized model for overdispersed count data. *Pop Ecol* 54(3): 467-474.

R CORE TEAM. 2018. *R: A Language and Environment for Statistical Computing*. (Version 3.6.1). R Foundation for Statistical Computing. Vienna, Austria.

SAGHIR A & LIN Z. 2015. Control charts for dispersed count data: an overview. *Qual Reliab Eng Int* 31(5): 725-739.

SELLERS KF. 2012. A generalized statistical control chart for over- or under-dispersed data. *Qual Reliab Eng Int* 28(1): 59-65.

SHEAFFER RL & LEAVENWORTH RS. 1976. The negative binomial model for counts in units of varying size. *J Qual Technol* 8(3): 158-163.

SPIEGELHALTER DJ. 2005. Handling over-dispersion of performance indicators. *BMJ Quality & Safety* 14(5): 347-351.

#### How to cite

BOAVENTURA LL, FERREIRA PH, FIACCONE RL, RAMOS PL & LOUZADA F. 2023. New statistical process control charts for overdispersed count data based on the Bell distribution. *An Acad Bras Cienc* 95: e20200246. DOI 10.1590/0001-3765202320200246.

*Manuscript received on March 3, 2020;  
accepted for publication on August 30, 2020*

#### LAION L. BOAVENTURA<sup>1</sup>

<https://orcid.org/0000-0003-2095-4322>

#### PAULO H. FERREIRA<sup>1</sup>

<https://orcid.org/0000-0001-6312-6098>

#### ROSEMEIRE L. FIACCONE<sup>1</sup>

<https://orcid.org/0000-0001-5439-1551>

#### PEDRO L. RAMOS<sup>2</sup>

<https://orcid.org/0000-0002-5387-2457>

#### FRANCISCO LOUZADA<sup>3</sup>

<https://orcid.org/0000-0001-7815-9554>

<sup>1</sup>Departamento de Estatística, Universidade Federal da Bahia, Avenida Milton Santos, s/n, Campus de Ondina, 40170-110 Salvador, BA, Brazil

<sup>2</sup>Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Vicuña Mackenna, 4860, Macul, 0436 Santiago, Región Metropolitana, Chile

<sup>3</sup>Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Avenida Trabalhador São-Carlense, 400, Centro, 13566-590 São Carlos, SP, Brazil

Correspondence to: **Paulo H. Ferreira**

E-mail: [paulohenri@ufba.br](mailto:paulohenri@ufba.br)



APPENDIX A -

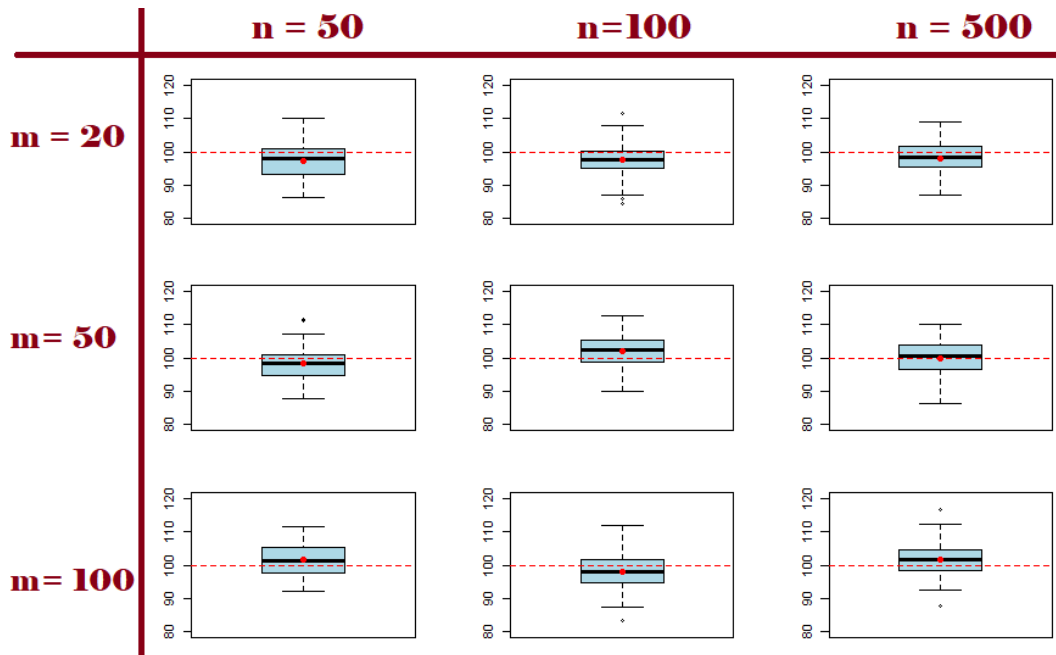


Figure A1. The  $ARL_0$  values of the Bell- $u$  chart for various  $m$  and  $\bar{n}$  ( $\mu = 3$  and  $\alpha = 0.01$ ).

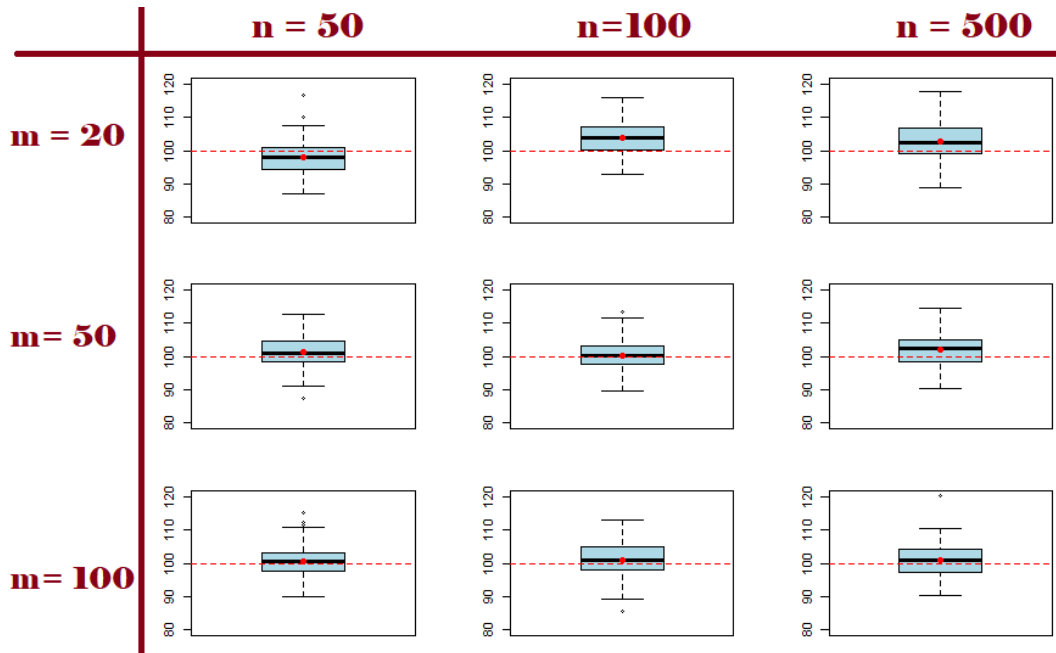


Figure A2. The  $ARL_0$  values of the Bell- $u$  chart for various  $m$  and  $\bar{n}$  ( $\mu = 15$  and  $\alpha = 0.01$ ).

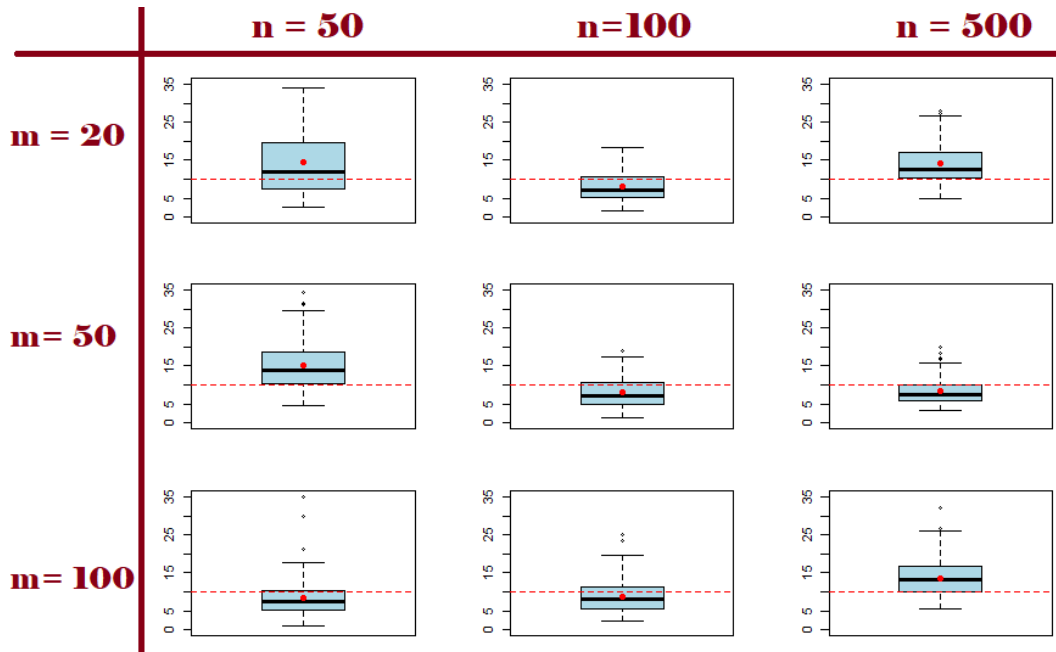


Figure A3. The  $ARL_0$  values of the Bell- $u$  chart for various  $m$  and  $\bar{n}$  ( $\mu = 3$  and  $\alpha = 0.1$ ).

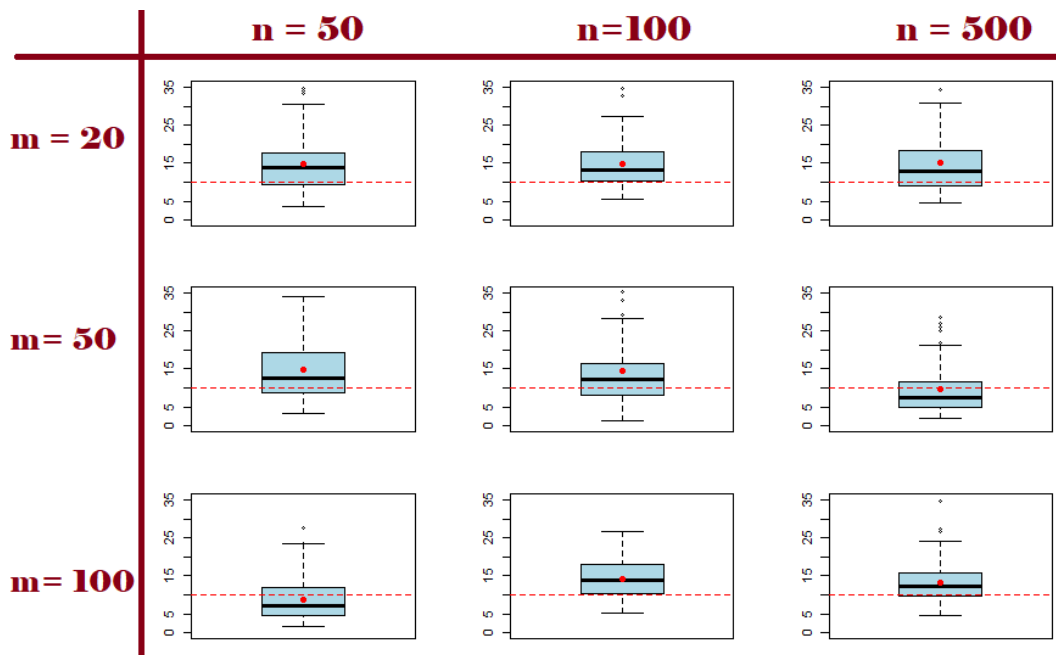


Figure A4. The  $ARL_0$  values of the Bell- $u$  chart for various  $m$  and  $\bar{n}$  ( $\mu = 15$  and  $\alpha = 0.1$ ).

**Table AI.** Mean (standard deviation in parentheses) values of  $ARL_1$  of the Bell- $u$  chart, for the different scenarios studied.

		$\bar{n} = 50$			$\bar{n} = 100$			$\bar{n} = 500$			
		$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	
$m = 20$	$\mu = 3$	$\alpha = 0.1$	14.233 (0.142)	1.172 (0.024)	1.013 (0.012)	13.242 (0.271)	1.228 (0.024)	1.015 (0.022)	15.233 (0.111)	1.151 (0.011)	1.011 (0.014)
		$\alpha = 0.01$	111.223 (0.282)	1.214 (0.022)	1.011 (0.024)	106.024 (0.482)	1.162 (0.021)	1.016 (0.024)	121.242 (0.261)	1.227 (0.022)	1.011 (0.014)
	$\mu = 15$	$\alpha = 0.1$	11.631 (0.124)	1.124 (0.013)	1.046 (0.022)	17.261 (0.113)	1.242 (0.023)	1.011 (0.022)	14.242 (0.111)	1.224 (0.014)	1.024 (0.010)
		$\alpha = 0.01$	113.224 (0.212)	1.174 (0.014)	1.072 (0.014)	112.424 (0.213)	1.161 (0.013)	1.024 (0.011)	115.222 (0.222)	1.262 (0.013)	1.012 (0.014)
$m = 50$	$\mu = 3$	$\alpha = 0.1$	12.274 (0.114)	1.156 (0.012)	1.022 (0.013)	13.292 (0.112)	1.245 (0.014)	1.018 (0.024)	12.262 (0.112)	1.232 (0.014)	1.005 (0.011)
		$\alpha = 0.01$	116.142 (0.211)	1.246 (0.013)	1.026 (0.014)	113.524 (0.222)	1.244 (0.011)	1.026 (0.022)	111.263 (0.203)	1.235 (0.014)	1.002 (0.012)
	$\mu = 15$	$\alpha = 0.1$	11.242 (0.091)	1.151 (0.011)	1.045 (0.014)	12.561 (0.113)	1.213 (0.012)	1.018 (0.023)	12.291 (0.102)	1.121 (0.012)	1.011 (0.014)
		$\alpha = 0.01$	118.524 (0.314)	1.193 (0.022)	1.032 (0.024)	112.100 (0.222)	1.171 (0.014)	1.018 (0.020)	114.100 (0.202)	1.159 (0.014)	1.004 (0.011)
$m = 100$	$\mu = 3$	$\alpha = 0.1$	12.522 (0.112)	1.104 (0.013)	1.005 (0.023)	16.120 (0.214)	1.197 (0.012)	1.007 (0.024)	13.692 (0.142)	1.126 (0.012)	1.009 (0.014)
		$\alpha = 0.01$	118.124 (0.292)	1.163 (0.014)	1.012 (0.012)	115.242 (0.262)	1.197 (0.014)	1.018 (0.014)	103.024 (0.412)	1.151 (0.022)	1.006 (0.014)
	$\mu = 15$	$\alpha = 0.1$	12.392 (0.114)	1.102 (0.010)	1.041 (0.023)	16.221 (0.210)	1.151 (0.012)	1.009 (0.024)	12.224 (0.113)	1.222 (0.012)	1.004 (0.011)
		$\alpha = 0.01$	115.224 (0.432)	1.183 (0.014)	1.008 (0.022)	116.524 (0.330)	1.126 (0.012)	1.008 (0.024)	119.503 (0.332)	1.206 (0.013)	1.000 (0.010)

**Table AII.** Mean (standard deviation in parentheses) values of  $ARL_0$  of the competitor  $u$ -type control charts when the true data-generating process is Bell( $\mu$ ) distributed, for some  $\mu$  and  $\alpha$  values ( $m=100$  and  $\bar{n}=100$ ).

Control chart	$\mu = 3$		$\mu = 15$	
	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.1$
Bell- $u$	96.3690 (13.5892)	10.2200 (0.4345)	96.5712 (13.6545)	9.7825 (0.4296)
$u$	13.7370 (0.6982)	4.0295 (0.0972)	7.2248 (0.2526)	2.8840 (0.0551)
$cm\mu$	112.8635 (17.3837)	10.6225 (0.4540)	76.3384 (9.3869)	7.1528 (0.2550)

**Table AIII.** Mean (standard deviation in parentheses) values of  $ARL_1$  of the competitor  $u$ -type control charts, when the true data-generating process is  $Bell(\mu)$  distributed, for some  $\mu$ ,  $\alpha$  and  $p$  values ( $m=100$  and  $\bar{n}=100$ ).

Control chart	$\mu = 3$					
	$\alpha = 0.01$			$\alpha = 0.1$		
	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$
Bell- $u$	90.8468 (12.3245)	84.1567 (11.6707)	13.4427 (0.6791)	10.1559 (0.4489)	9.9885 (0.4074)	3.0521 (0.0619)
$u$	13.5156 (0.6548)	13.1820 (0.6323)	4.2784 (0.1085)	4.0293 (0.0989)	3.9949 (0.0984)	1.9329 (0.0267)
$cm\mu$	108.1874 (16.5333)	89.3612 (12.5799)	15.3133 (0.8142)	10.6599 (0.4826)	10.5533 (0.4375)	3.2587 (0.0693)
Control chart	$\mu = 15$					
	$\alpha = 0.01$			$\alpha = 0.1$		
	$p = 0.5\%$	$p = 1\%$	$p = 10\%$	$p = 0.5\%$	$p = 1\%$	$p = 10\%$
Bell- $u$	93.6208 (12.9182)	83.0078 (10.8026)	3.0950 (0.0612)	10.0302 (0.4335)	9.1960 (0.3735)	1.4209 (0.0131)
$u$	7.2176 (0.2531)	6.9502 (0.2402)	1.3911 (0.0123)	2.9273 (0.0584)	2.8343 (0.0562)	1.1289 (0.0055)
$cm\mu$	69.3032 (8.0359)	82.7237 (10.6946)	3.0015 (0.0575)	7.4049 (0.2543)	9.7473 (0.4111)	1.4464 (0.0138)

**APPENDIX B**

```

# y = vector with the total number of defects per sample (including
#   phases 1 and 2 samples, in this order)
# m = number of phase 1 samples
# n = sample size (all the phases 1 and 2 samples are of the same size n)
# L = "distance" between the CL and the control limits (in the usual
#   Six Sigma quality control program, L = 3)

Bell.Chart.TOTAL = function(y, m, n, L){
  require(LambertW)
  mu = mean(y[1:m])/n
  UCL = n*mu+L*sqrt((n*mu*(1+W(mu))))
  CL = n*mu
  LCL = n*mu-L*sqrt((n*mu*(1+W(mu))))
  UCL = rep(UCL, length(y))
  CL = rep(CL, length(y))
  LCL = rep(LCL, length(y))
  plot(UCL, main="Bell Chart", xlab="Sample", ylab="Total Number of Events",
       ylim=c(min(LCL, y), max(UCL, y)), col="red", type="l", lwd=3)
  lines(CL, col="blue", lwd=2)
  lines(y, col="black", lwd=1.5)
  lines(LCL, col="red", lwd=3)
  abline(v=m+0.5, col="green", lty=2, lwd=1.5)
  points(y, col= ifelse(y>UCL | y<LCL, "green", "black"), pch=19)
  out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
  rownames(out) = c(" ")
  colnames(out) = c("UCL", "CL", "LCL")
  out
}

x = c(16,18,12,15,24,21,28,20,25,19,18,21,16,22,19,12,14,9,16,21,27,27,
      34,31,19,29,17,35,20,18,12,34,13,25,23,24,20,27,16,28)
Bell.Chart.TOTAL(y=x, m=20, n=100, L=3)

```