On ARL-unbiased $c-$charts for i.i.d. and INAR(1) Poisson counts

Sofia Raquel Bastos Benvindo Paulino
Instituto Superior Técnico - Universidade de Lisboa

April, 2015

Abstract

In Statistical Process Control (SPC) it is usual to assume that counts have a Poisson distribution. The non-negative, discrete and asymmetrical character of a control statistic with such distribution and the value of its target mean may prevent the quality control practitioner to deal with a $c-$chart with: a pre-specified in-control average run length (ARL); a positive lower control limit; the ability to control not only increases but also decreases in the mean of those counts in a timely fashion. Furthermore, as far as we have investigated, the $c-$charts proposed in the SPC literature tend not to be ARL-unbiased.[1]

In this paper, we explore the notions of unbiased, randomized and uniformly most powerful unbiased (UMPU) tests to correct the bias of the ARL function of the $c-$chart. We use the R statistical software to provide instructive illustrations of: ARL-unbiased $c-$charts for i.i.d. Poisson counts; "quasi" ARL-unbiased $c-$charts for the mean of first-order integer-valued autoregressive (INAR(1)) Poisson counts.

Keywords: SPC; Poisson INAR(1) processes; unbiased, randomized and UMPU tests.

1 Introduction

The $c-$chart enjoys widespread popularity as the chart of choice to control the mean of the number of nonconformities (White *et al.*, 1997) in a constant-size sample. However, $c-$charts may behave poorly when it comes to the detection of decreases in λ. Expectedly, several authors pursued other alternatives and tried to define charts that would be more sensitive to downward shifts in λ. Regrettably, all these charts still have low responsiveness of the lower control limit when dealing with small process means. Furthermore, they tend to be ARL-biased control charts.

Another important issue is related to the fact that count processes are often autocorrelated but we falsely assume serial independence, especially while planning a control chart to monitor the process mean. Unsurprisingly, this naïve approach has a severe impact on the ARL and this calls for the use of charts accounting for the autocorrelation structure such as the ones proposed by Weiss (2009, Chapter 20). Regretfully, these charts are not ARL-unbiased.

As far as we know, no ARL-unbiased chart has been proposed for count data, regardless of the fact that the detection of decreases in λ plays a crucial role in the effective implementation of any quality improvement program. Therefore, the aim of this paper is to propose and illustrate control charts for the mean of count processes, whether i.i.d. or autocorrelated, that allow us to pre-specify an in-control ARL value, say $ARL^{\star}$; control in a timely fashion not only increases, but also decreases in the process mean; deal with an unbiased (or "quasi" unbiased) ARL function. In order to achieve this, the paper is organized as follows.

In Section 2, we tackle i.i.d. Poisson counts. The $c-$chart is revisited, associated with both $3-\sigma$ and quantile based control limits, along with some variants introduced by other authors. Inspired by randomized tests (rarely used in practice in SPC) and, mostly, uniformly most powerful unbiased (UMPU) tests, we propose another variant of the c-chart, obtaining an ARL-unbiased $c-$chart with in-control ARL

[1] Pignatiello *et al.* (1995) invented the term *ARL-unbiased* and used it to coin any control chart for which all out-of-control ARL values are smaller than the in-control ARL.

exactly equal to $ARL^{\star}$. This variant is associated with quantile based control limits and randomization probabilities of triggering a signal when the value of the control statistic is equal to LCL and UCL, to bring the size of the test exactly to level $\alpha = (ARL^{\star})^{-1}$. We present some striking and instructive examples that show that the performance of this last variant of the $c-$chart supersedes the previous ones.

In Section 3, we motivate and introduce the definition of binomial thinning and first-order integer-valued autoregressive (INAR(1)) process with Poisson distributed marginals.

We start Section 4 by briefly illustrating the impact of falsely assuming i.i.d. Poisson counts in the performance of the $c-$chart. We revisit the $c-$chart introduced by Weiss (2007) to monitor the mean of a Poisson INAR(1) process. Two variants of this chart are introduced; one of them comprises randomization probabilities. We go on to compare the performance of all these charts.

2 Control charts for the mean of i.i.d. Poisson counts

The control statistic of the $c-$chart is X_t, the total number of defects in the $t-$th sample ($t \in \mathbb{N}$). We shall assume that the random variables (r.v.) X_t, $t \in \mathbb{N}$, are independent and identically distributed (i.i.d.) to the r.v. $X \sim \text{Poisson}(\lambda)$.

Let $\lambda = \lambda_0 + \delta$ be the process mean, where: λ_0 denotes the target mean and is frequently assumed to be known; and δ represents the magnitude of the shift in the process mean. The process is said to be in-control (resp. out-of-control) if $\delta = 0$ (resp. $\delta \in (-\lambda_0, 0) \cup (0, +\infty)$). If $\delta \in (-\lambda_0, 0)$ (resp. $\delta \in (0, +\infty)$) then a downward (resp. upward) shift has occurred and the practitioner ought to be alerted to this change.

The $c-$chart triggers an alarm at sample t if X_t is beyond the control limits. The $c-$chart makes use of the $k - \sigma$ control limits, defined for this integer-valued control statistic as the following ceiling and floor functions of λ_0: $LCL = \lceil \max\{0, \lambda_0 - k\sqrt{\lambda_0}\} \rceil$ and $UCL = \lfloor \lambda_0 + k\sqrt{\lambda_0} \rfloor$, where k is a positive constant, usually selected so that false alarms are rather infrequent, and is often taken as equal to 3.

The performance of this chart is generally evaluated using the number of samples taken until an alarm is triggered, the run length (RL) or, more specifically, the expected value of this number, the average run length (ARL). Ideally, it is desirable that the emission of false alarms occurs rarely (corresponding to large in-control ARL) and the emission of valid signals occurs as quickly as possible (corresponding to small out-of-control ARL).

The RL of this Shewhart-type chart has a geometric distribution and depends on the magnitude of the shift in the process mean, $\delta \in (-\lambda_0, +\infty)$: $RL(\delta) \sim \text{Geometric}(\xi(\delta))$, where $\xi(\delta) = P(\text{emission of a signal} \mid \delta) = P(X \notin [LCL,\, UCL] \mid \delta)$. Hence $ARL(\delta) = \frac{1}{\xi(\delta)}$.

The $c-$chart associated with $3-\sigma$ control limits has been historically used, implicitly assuming that the normal approximation to the Poisson distribution is adequate (Ryan and Schwertman, 1997). Adopting $3-\sigma$ limits, as Shewhart did, corresponds to a probability of a false alarm of approximately 2×0.00135 (Ramos, 2013, p. 3) if that approximation is reasonable. However, since the Poisson distribution is skewed to the right, the lower and upper tail areas tend to be uneven, leading to an in-control ARL that can substantially differ from $(2 \times 0.00135)^{-1} \simeq 370.4$. We might add that the $c-$chart with $3-\sigma$ control limits tends to be ARL-biased in the sense that some out-of-control ARL values are larger than the in-control ARL, i.e., it takes longer, in average, to detect some shifts in the process mean than to trigger a false alarm. This unacceptable performance leads us to state that there is plenty of room for improvement.

Aebtarm and Bouguila (2011) thoroughly described ten alternative charts to monitor the process mean of i.i.d. Poisson counts and the comparison study undertaken by these authors led to the conclusion that the *optimal control limits for the c−chart* showed outstanding results and is the *best defect chart* to replace the $c-$chart with $3 - \sigma$ limits. This chart follows the work of Ryan and Schwertman (1997) and can be described as follows: the number of defects, c, in a sample of fixed size is plotted against the

control limits $LCL = 1.5307 + 1.0212\bar{c} - 3.2197\sqrt{\bar{c}}$ and $UCL = 0.6182 + 0.9996\bar{c} + 3.0303\sqrt{\bar{c}}$, where $\bar{c}$ is an estimate of the target value λ_0. These *optimal control limits for the c−chart* were obtained by Aebtarm and Bouguila (2011) by linear regression based on a table of the "best" $c-$chart limits for several values of λ_0 considering a certain loss function.

Alternatively, it occurred to us that control limits based on quantiles of the Poisson distribution could be an appropriate alternative to the $3-\sigma$ limits,[2] and their adoption a possible remedy for the ARL-biased character of this standard chart. Moreover, since we are aware that the discrete character of this distribution makes it difficult to achieve a pre-specified probability of false alarm, say α, we suggest the use of the following control limits termed, from now on, *quantile based control limits*.

Definition 2.1. Quantile based control limits

Requiring that $P(X \notin [LCL, UCL] \mid \delta = 0) \leq \alpha$ and setting $\alpha = \alpha_{LCL} + \alpha_{UCL}$ leads to the obtention of a LCL and a UCL such that:

$$P(X < LCL \mid \delta = 0) \leq \alpha_{LCL}; \tag{1}$$

$$P(X > UCL \mid \delta = 0) \leq \alpha_{UCL}. \tag{2}$$

Thus, the quantile based LCL (resp. UCL) is the largest (resp. smallest) integer satisfying (1) (resp. (2)).

We now compare the ARL functions of the charts associated with: the $3-\sigma$ control limits; the Ryan & Schwertman control limits, as defined by Aebtarm and Bouguila (2011); the *quantile based control limits* with $\alpha = 0.0027$ and $\alpha_{LCL} = \alpha_{UCL}$. The comparisons between the ARL functions of these three charts shall be done with great care, because the charts are bound to have distinct in-control ARL values, which are in turn different from the desired in-control ARL value $\alpha^{-1} \simeq 370.4$.

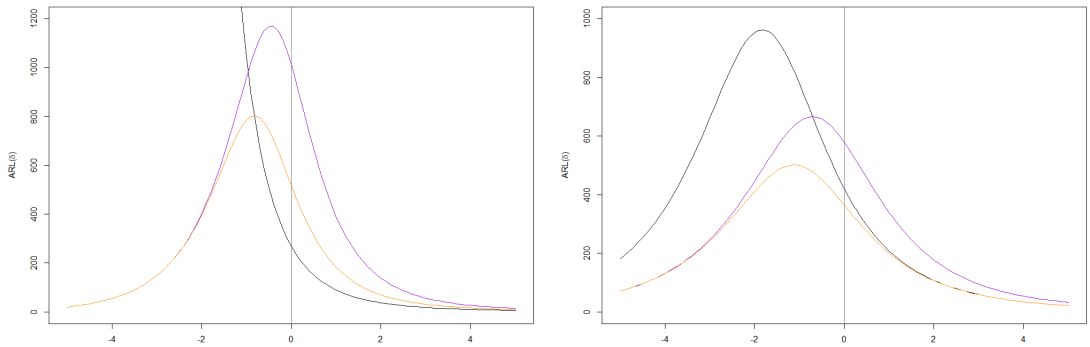

Figure 1: $ARL(\delta)$ curves, for $3-\sigma$ (curves in **black**), Ryan & Schwertman (curves in **orange**) and quantile based control limits (curves in **violet**) — $\lambda_0 = 8$ (left), 19 (right).

When $\lambda_0 = 8$ we compare an upper one-sided chart with $3-\sigma$ control limits with two charts with positive LCL, whereas when $\lambda_0 = 19$ we deal with three two-sided charts. Either way, we can state that the adoption of the Ryan & Schwertman and quantile based control limits results in an improved ARL function since the detection of increases and most decreases in λ requires in average a number of samples under the in-control ARL value. Moreover, in both cases, the quantile based control limits lead to the less biased ARL function (with a larger in-control ARL value). Since the major purpose of this paper is to find ARL-unbiased charts the chart with quantile based control limits is the one we should favour, despite the fact that it compares unfavorably to the two other charts (resp. chart with Ryan & Schwertman control limits) when it comes to the detection of increases (resp. most decreases) in the process mean, and it has an in-control ARL further away from the desired value of $\alpha^{-1} \simeq 370.4$.

[2] Coincidentally, Wetherill and Brown (1991, p. 215) also recommended the use of quantiles.

2.1 Towards an ARL-unbiased $c-$chart

Capitalizing on the close (yet controversial) resemblance between control charting and repeated hypothesis testing over time, and inspired by the somewhat promising results of the quantile based control limits and by UMPU tests (Lehmann, 1986, pp. 135–140), we decided to combine the use of quantile based control limits, LCL and UCL, and randomization probabilities, γ_{LCL} and γ_{UCL}, of triggering a signal when the value of the control statistic is equal to LCL and UCL. This approach allowed us to bring the in-control ARL to α^{-1} and also guarantee that we deal with an ARL-unbiased $c-$chart.

The next proposition provides a systematic way of obtaining a $c-$chart that was inspired by an UMPU test confronting $H_0 : \lambda = \lambda_0$ and $H_1 : \lambda \neq \lambda_0$, and as far as we have investigated has not been proposed in the SPC setting.

Proposition 2.2. ARL-unbiased $c-$chart

The ARL-unbiased $c-$chart, with in-control ARL value equal to α^{-1}, makes use of the quantile based control limits LCL and UCL, associated with $\alpha = \alpha_{LCL} + \alpha_{UCL}$. It triggers a signal with: probability one if the sample number of defects is beyond LCL or UCL; and probabilities γ_{LCL} and γ_{UCL}, obtained by solving the system of equations

$$\gamma_{LCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{LCL}}{LCL!} + \gamma_{UCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{UCL}}{UCL!} = \alpha - 1 + \sum_{x=LCL}^{UCL} \mathrm{e}^{-\lambda_0} \frac{\lambda_0^x}{x!} \tag{3}$$

$$LCL \times \gamma_{LCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{LCL}}{LCL!} + UCL \times \gamma_{UCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{UCL}}{UCL!} = \alpha \times \lambda_0 - \lambda_0 + \sum_{x=LCL}^{UCL} x\, \mathrm{e}^{-\lambda_0} \frac{\lambda_0^x}{x!}, \tag{4}$$

if the sample number of defects is equal to LCL and UCL, respectively. It is important to note that (4) simplifies into:

$$\gamma_{UCL} = \frac{\alpha - 1 + \sum_{x=0}^{UCL-1} \mathrm{e}^{-\lambda_0} \frac{\lambda_0^x}{x!}}{\mathrm{e}^{-\lambda_0} \frac{\lambda_0^{UCL-1}}{(UCL-1)!}} \tag{5}$$

if $LCL = 0$; and

$$\gamma_{LCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{LCL-1}}{(LCL-1)!} + \gamma_{UCL} \times \mathrm{e}^{-\lambda_0} \frac{\lambda_0^{UCL-1}}{(UCL-1)!} = \alpha - 1 + \sum_{x=LCL-1}^{UCL-1} \mathrm{e}^{-\lambda_0} \frac{\lambda_0^x}{x!}, \tag{6}$$

if $LCL > 0$.

The signal is triggered by the ARL-unbiased $c-$chart with probability $\xi(\delta) = E_\lambda[\phi(X)]$, where the critical function $\phi(X)$ is given by

$$\phi(x) = P(\text{Reject } H_0 | X = x) = \begin{cases} 1 & \text{if } x < LCL \text{ or } x > UCL \\ \gamma_{LCL} & \text{if } x = LCL \\ \gamma_{UCL} & \text{if } x = UCL \\ 0 & \text{if } LCL < x < UCL. \end{cases} \tag{7}$$

2.2 Illustrations

The results presented in this section refer to the ARL of the $c-$chart with quantile based control limits with $\alpha_{LCL} = \alpha_{UCL}$ and the ARL-unbiased $c-$chart. In any case, we took the target values $\lambda_0 = 8, 19$ and $\alpha = \alpha_{LCL} + \alpha_{UCL} = 0.0027$. However, feeling that considering $\alpha_{LCL} = \alpha_{UCL}$ could not be the best way of *distributing* the probability of a false alarm α when dealing with a control statistic with an asymmetrical distribution, the ARL-unbiased $c-$chart was tested with $\alpha_{LCL} = \left(1 - \frac{1}{m}\right) \alpha$ and $\alpha_{UCL} = \frac{1}{m}\, \alpha$, where $m = 2, \ldots, 50$. Interestingly, the best values of m do not exceed 4, in any of the cases we have considered.

In Table 1 we can find the set up of each chart: target values; control limits, value of m; randomization

probabilities (if any);[3] supremum of the ARL, its bias and in-control value.

Table 1: LCL, UCL, m, pair of randomization probabilities; supremum, bias and in-control value of ARL, for $\lambda_0 = 8, 19$ — listed in order corresponding to the $c-$chart with quantile based control limits and the ARL-unbiased $c-$chart.

λ_0	$[LCL, UCL]$	m	$(\gamma_{LCL}, \gamma_{UCL})$	Supremum of ARL	Bias of ARL	In-control ARL
8	$[1, 18]$	2	—	1170.5200	-0.446816	1014.3730
	$[1, 18]$	2	$(0.482414, 0.444451)$	370.3704	-5.7732×10^{-15}	370.3704
19	$[7, 33]$	2	—	666.4702	-0.707531	579.2475
	$[8, 34]$	3	$(0.003234, 0.951408)$	370.3704	-3.3751×10^{-14}	370.3704

Table 1 and Figure 2 confirm that the ARL-unbiased $c-$chart has indeed a pre-specified in-control ARL and that it requires, in average, less time to trigger a signal in the presence of any shift in λ than to trigger a false alarm. For these reasons this chart clearly outperforms the $c-$charts with $3-\sigma$ or quantile based control limits. The randomization of the emission of a signal not only allows the ARL-unbiased $c-$chart to have the in-control ARL we desire, but also to quickly detect decreases in the process mean even if we are dealing with a null lower control limit such as when $\lambda_0 = 8$.

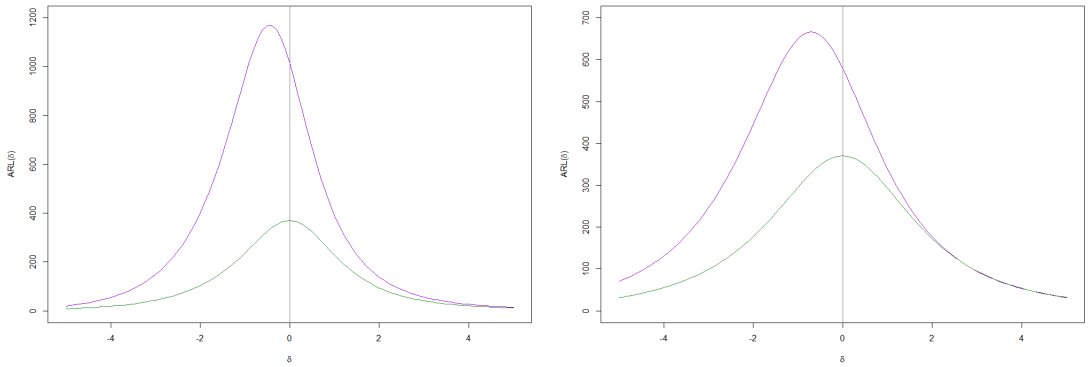

Figure 2: $ARL(\delta)$ curves, for the $c-$chart with quantile based control limits (curves in **violet**) and the ARL-unbiased $c-$chart (curves in **green**) — $\lambda_0 = 8$ (left), 19 (right).

3 INAR(1) Poisson counts

In the last three decades, there has been a growing interest in modelling discrete-valued time series, i.e., series taking values on a finite or countably infinite set (Silva, 2005, p. 21; Morais, 2012, p. 48).

When we are dealing with processes of counts with large values, the time series in question can be studied by using continuous-valued models (Silva, 2005, p. 21). However, when the observed values of the time series are small, processes such AR(p) models are of limited use because the simple procedure of multiplying an integer-valued r.v. by a real constant may lead to a non-integer r.v., as stated by Silva (2005, p. 22) (Morais, 2012, p. 49).

A possible way out is to replace the (scalar) multiplication by a random operation such as the thinning operation described below; it can be thought as the (scalar) multiplication counterpart in the integer-valued setting that preserves the integer structure of the original process (Morais, 2012, p. 49).

Definition 3.1. Binomial thinning operation (Steutel and Van Harn, 1979, p. 894)
Let X be a discrete r.v. with range $\mathbb{N}_0$ and β a scalar in $(0, 1)$. Then the binomial thinning operation on X results in another r.v. defined as $\beta \circ X = \sum_{i=1}^{X} Y_i$, where: $\circ$ represents the binomial thinning operator; $\{Y_i : i \in \mathbb{N}\}$ is a sequence of i.i.d. Bernoulli(β) r.v., independent of X. In this case we usually say that $\beta \circ X$ arises from X by binomial thinning.

[3] Please note that in a very few cases we were not able to obtain admissible values for the randomization probabilities associated with the ARL-unbiased $c-$chart (e.g., $(\lambda_0, m) = (8, 5), (19, 2)$). Moreover, since for this kind of chart we obtained the same bias in every admissible case, we selected the best value of m as the smallest m associated with admissible randomization probabilities.

The first-order integer-valued autoregressive, INAR(1), process was introduced by Mckenzie (1985), who discussed this model with Poisson or negative binomial marginals (Weiss, 2009, pp. 281–295).

Definition 3.2. INAR(1) process (Silva, 2005, p. 34; Weiss, 2009, p. 282; Paulino *et al.*, 2014)
$\{X_t : t \in \mathbb{Z}\}$ is said to be a first-order integer-valued autoregressive process if

$$X_t = \beta \circ X_{t-1} + \epsilon_t, \tag{8}$$

where: $\beta \in (0,1)$; $\circ$ represents the binomial thinning operator; $\{\epsilon_t : t \in \mathbb{Z}\}$ is a sequence of nonnegative integer-valued i.i.d. r.v., with mean μ_ϵ and variance σ^2_ϵ; ϵ_t and X_{t-1} are assumed to be independent r.v.; all thinning operations are performed independently of each other and of $\{\epsilon_t : t \in \mathbb{Z}\}$; the thinning operations at time t are independent of $\{\ldots, X_{t-2}, X_{t-1}\}$.

As for the marginal moments of the INAR(1) process, Morais (2012, p. 53) refers that, in case ϵ_t has a discrete self-decomposable distribution (Steutel and Van Harn, 1979) with $E(\epsilon_t) = \mu_\epsilon$ and $V(\epsilon_t) = \sigma^2_\epsilon < +\infty$, the INAR(1) process is stationary, has mean and variance given by $E(X_t) = \frac{\mu_\epsilon}{1-\beta}$ and $V(X_t) = \frac{\beta\mu_\epsilon + \sigma^2_\epsilon}{1-\beta^2}$. The class of discrete self-decomposable distributions contains the family of Poisson distributions (Silva, 2005, p. 35) and a Poisson INAR(1) model can be defined (Morais, 2012, p. 53).

Definition 3.3. Poisson INAR(1) process (Silva, 2005, p. 35; Weiss, 2009, pp. 284, 286–287)
If $\epsilon_t \overset{i.i.d.}{\sim} Poisson(\lambda)$, $t \in \mathbb{Z}$, then $\{X_t = \beta \circ X_{t-1} + \epsilon_t : t \in \mathbb{Z}\}$ is said to be a Poisson INAR(1) process.

Proposition 3.4. Properties of the INAR(1) process (Silva, 2005, p. 35-36; Weiss, 2009, p. 283, 284; Paulino *et al.*, 2014)
The Poisson INAR(1) process, $\{X_t : t \in \mathbb{Z}\}$, is stationary, and if $X_1 \sim Poisson(\frac{\lambda}{1-\beta})$ then $X_t \sim Poisson(\frac{\lambda}{1-\beta})$. Furthermore, this process is a (time-)homogeneous Markov chain with state space $\mathbb{N}_0$ and (one-step) transition probabilities $p_{i\,j} \equiv p_{i\,j}(\lambda, \beta)$ equal to

$$\sum_{m=0}^{\min\{i,j\}} \binom{i}{m} \beta^m (1-\beta)^{i-m} \times e^{-\lambda} \frac{\lambda^{j-m}}{(j-m)!}, \; i, j \in \mathbb{N}_0. \tag{9}$$

4 Control charts for the mean of a Poisson INAR(1) process

As noted by Ramos (2013, p. 15) one of the standard assumptions while designing a control chart is that the underlying output is independent. However, this assumption is not fulfilled for many real life data sets, it can be totally unrealistic and significantly affect the performance of standard control charts (Wieringa, 1999, pp. 2, 10; Morais, 2012, p. 70), as we illustrate below for an INAR(1) process with $\lambda_0 = 5$.

We are going to rely on Monte Carlo simulation to obtain estimates of ARL of this traditional chart, when $\lambda = \lambda_0 + \delta_\lambda$, with $\delta_\lambda = 0, 1$ and $\beta = 0(0.01)0.99$.

We ought to note that, when $\beta = 0$, the estimates of $ARL(0)$ and $ARL(1)$, 184.6222 and 48.9368, do not differ much from the exact values 183.3822 and 49.7711, respectively.

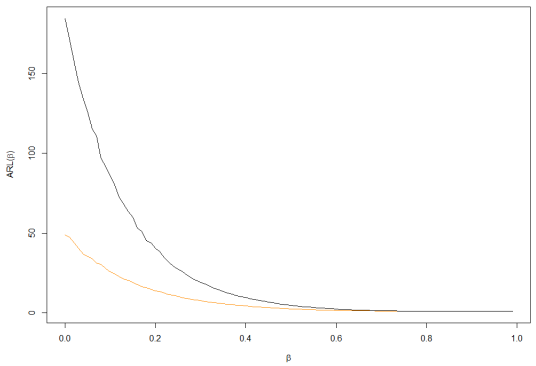

Figure 3: Estimates of in-control and out-of-control ARL — $\lambda_0 = 5$ and $\delta_\lambda = 0, 1$ (black and **orange** curves, resp.).

Figure 3 portrays quite vividly the fact that, when the autocorrelation structure is not recognized or it is ignored and a $c-$chart with $3-\sigma$ control limits is used, the true in-control or out-of-control ARL values are bound to be overestimated; in particular, overestimating ARL means a swifter detection of shifts in λ but regretfully more frequent false alarms. This overestimation becomes more severe as the thinning parameter β increases, confirming how inept is the traditional

Shewhart-type $c-$chart with $3-\sigma$ limits in the monitoring of the process mean, under the false assumption of independence when the output follows an INAR(1) process. Consequently, we need alternative charts to control the process mean and whose control limits take into account the dependence structure of the INAR(1) process.

4.1 $c-$chart for the mean of a Poisson INAR(1) process

Controlling the mean of an INAR(1) process is a relatively recent research subject and the detection of changes from the target value $\lambda_0/(1-\beta_0)$ to $\lambda/(1-\beta)$ can be done by using a control chart (Paulino *et al.*, 2014).

Throughout this section, we shall assume that, in the absence of assignable causes, $\lambda = \lambda_0$ and $\beta = \beta_0$, and the purpose of using a control chart is to monitor the mean of a Poisson INAR(1) process from its target value $\frac{\lambda_0}{1-\beta_0}$ to $\frac{\lambda}{1-\beta}$ due to a change from λ_0 to $\lambda = \lambda_0 + \delta_\lambda$ or from β_0 to $\beta = \beta_0 + \delta_\beta$, where $\delta_\lambda \in (-\lambda_0, +\infty)$ and $\delta_\beta \in [-\beta_0, 1-\beta_0)$.

Definition 4.1. $c-$chart for the mean of a Poisson INAR(1) process (Weiss, 2007; Weiss, 2009, p. 419; Morais, 2012, p. 99)
X_t, $t \in \mathbb{N}$, is the control statistic of the $c-$chart for the mean of a Poisson INAR(1) process. The $k-\sigma$ control limits of this chart are given by $LCL = \max\left\{0, \frac{\lambda_0}{1-\beta_0} - k\sqrt{\frac{\lambda_0}{1-\beta_0}}\right\}$ *and* $UCL = \frac{\lambda_0}{1-\beta_0} + k\sqrt{\frac{\lambda_0}{1-\beta_0}}$, *where k is a positive constant, usually selected by considering a reasonably large in-control ARL.*

Now, let us address the performance of this chart. Since the control statistics of this chart are dependent r.v., the RL distribution of this Shewhart-type chart is no longer geometric. In fact, if we consider that $\{X_t,\, t \in \mathbb{N}_0\}$ is a Poisson INAR(1) process governed by the transition matrix $\mathbf{P} = [p_{ij}]_{i,j\in\mathbb{N}_0}$, where $p_{ij} = p_{ij}(\lambda, \beta)$ are the transition probabilities defined in (9), $y = \lceil LCL \rceil$, $x = \lfloor UCL \rfloor$ and $RL^u(\lambda, \beta) = \min\{t : X_t < y \text{ or } X_t > x | X_0 = u\}$ is the RL of the $c-$chart, conditional on $X_0 = u$ ($u \in \{y, ..., x\}$) and on the values of λ and β (instead of the magnitude of the shifts on these two parameters), then $RL^u(\lambda, \beta)$ is the first passage time to the set of states $\{0, \ldots, y-1\} \cup \{x+1, \ldots\}$ by the discrete-time Markov chain $\{X_t,\, t \in \mathbb{N}_0\}$ (Morais, 2012, pp. 99–100).

The next proposition provides an expression for the mean of $RL^u(\lambda, \beta)$.

Proposition 4.2. RL of the $c-$chart for the mean of a Poisson INAR(1) process
For $u \in \{y, \ldots, x\}$, $RL^u(\lambda, \beta)$ has a discrete phase-type distribution and $ARL^u(\lambda, \beta) = \underline{\boldsymbol{e}}_u^\top \times [\boldsymbol{I} - \boldsymbol{Q}(\lambda, \beta)]^{-1} \times \underline{\mathbf{1}}$, *where $\underline{\boldsymbol{e}}_u^\top$ is the $(u-y+1)^{th}$ vector of the orthogonal basis for $\mathbb{R}^{(x-y+1)}$,* $\boldsymbol{Q}(\lambda, \beta) = [p_{lk}]_{l,k=y}^{x}$, *$\boldsymbol{I}$ represents an identity matrix with rank $(x-y+1)$, and $\underline{\mathbf{1}}$ is a column-vector with $(x-y+1)$ ones.*

Since the observed value of $X_0 \equiv X_0(\lambda, \beta) \sim \text{Poisson}\left(\frac{\lambda}{1-\beta}\right)$ is unknown and a value of $X_0(\lambda, \beta)$ beyond the control limits would lead to a null ARL, Weiss (2009, p. 422) recommended the use of what this author called *overall* ARL: $ARL(\lambda, \beta) = \sum_{u=y}^{x} ARL^u(\lambda, \beta) \times P[X_0(\lambda, \beta) = u]$.

Expectedly, the $c-$chart with $3-\sigma$ limits for the mean of an INAR(1) process has similar disadvantages as the one for i.i.d. output, after all: if $\frac{\lambda_0}{1-\beta_0} \leq 9$ then we deal with a null LCL and the chart is unable to detect decreases in the process mean in a timely fashion; the control limits disregard the asymmetrical character of the probability function of its control statistic $X_t \sim \text{Poisson}\left(\frac{\lambda_0}{1-\beta_0}\right)$, therefore we are bound to use an ARL-biased chart.

4.2 Towards a quasi ARL-unbiased $c-$chart

Since in the INAR case the control statistics are dependent r.v., the probability of triggering a signal at sample t is time-dependent. Thus, quantile based control limits (such as the previously used in the i.i.d.

case) are meaningless and we propose the following method to determine alternative control limits to get eventually closer to a pre-specified in-control overall ARL, without resorting (for now) to randomization.

1. Set a lower and an upper one-sided $c-$chart. The lower (resp. upper) one-sided control $c-$chart has integer limits $LCL^\star$ and $UCL = +\infty$ (resp. $LCL = 0$ and $UCL^\star$).
2. Fix the smallest acceptable in-control overall ARL value for each one-sided $c-$chart, say $2 \times ARL^\star$.
3. Search for the largest (resp. smallest) integer $LCL^\star$ (resp. $UCL^\star$) leading to a lower (resp. upper) one-sided $c-$chart whose in-control overall ARL is larger than $2 \times ARL^\star$.[4] [5]
4. Use the two control limits we found, $LCL^\star$ and $UCL^\star$, to set what we shall call from now on the unrandomized $c-$chart.

Proposition 4.3. RL of the unrandomized $c-$chart for the mean of a Poisson INAR(1) process
The characterization of the RL of the unrandomized $c-$chart follows from Proposition 4.2, by taking $y \equiv LCL^\star$ and $x \equiv UCL^\star$.

Adapting the randomization to suit the INAR(1) case is straightforward in practice. We termed the resulting chart randomized $c-$chart for the mean of a Poisson INAR(1) process. As for defining its overall RL we refer the reader to the following proposition.

Proposition 4.4. RL of the randomized $c-$chart for the mean of a Poisson INAR(1) process
The characterization of the RL of the randomized $c-$chart, $RL(\lambda, \beta, \gamma_{LCL^\star}, \gamma_{UCL^\star})$, also follows from Proposition 4.2, by taking $y \equiv LCL^\star$ and $x \equiv UCL^\star$ and considering the following sub-stochastic matrix:

$$\begin{aligned}\mathbf{Q} &= \mathbf{Q}(\lambda, \beta, \gamma_{LCL^\star}, \gamma_{UCL^\star}) \\ &= \begin{bmatrix} (1-\gamma_{LCL^\star})\, p_{y\,y} & (1-\gamma_{LCL^\star})\, p_{y\,y+1} & \cdots & (1-\gamma_{LCL^\star})\, p_{y\,x-1} & (1-\gamma_{LCL^\star})\, p_{y\,x} \\ p_{y+1\,y} & p_{y+1\,y+1} & \cdots & p_{y+1\,x-1} & p_{y+1\,x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{x-1\,y} & p_{x-1\,y+1} & \cdots & p_{x-1\,x-1} & p_{x-1\,x} \\ (1-\gamma_{UCL^\star})\, p_{x\,y} & (1-\gamma_{UCL^\star})\, p_{x\,y+1} & \cdots & (1-\gamma_{UCL^\star})\, p_{x\,x-1} & (1-\gamma_{UCL^\star})\, p_{x\,x} \end{bmatrix} \end{aligned} \quad (10)$$

The procedure to obtain the randomization probabilities is iterative, involves the bisection method (Quarteroni *et al.*, 2007, pp. 250–253) and comprises the following steps.

1. Set a lower one-sided $c-$chart. The control limits of this chart coincide with $LCL^\star$ and UCL, which were obtained using the procedure described above.
2. Apply the bisection method to obtain $\gamma_{LCL^\star}$ so that the in-control overall ARL of the chart is equal to $2 \times ARL^\star$.
3. Repeat the previous steps with an upper one-sided $c-$chart (whose control limits are 0 and $UCL^\star$) in order to obtain $\gamma_{UCL^\star}$.
4. Use the randomization probabilities $\gamma^\star_{LCL}$ and $\gamma^\star_{UCL}$ along with $[LCL^\star, UCL^\star]$ to set a $c-$chart.

Since the results in the next section suggest that the estimated in-control overall ARL function of the resulting $c-$chart tends to be very close to $ARL^\star$ and it is practically unbiased, Paulino *et al.* (2014) coined it "quasi" ARL-unbiased $c-$chart.

[4] It is important to notice that we are not able to calculate the overall ARL values for the lower one-sided control $c-$chart because $UCL = +\infty$. Therefore, we obtain an approximate overall ARL of the unrandomized $c-$chart, by considering as the value of UCL the smallest value such that $P(X \geq UCL) < 10^{-10}$, where $X \sim \text{Poisson}\left(\frac{\lambda_0}{1-\alpha_0}\right)$.

[5] The initial value of $LCL^\star$ (resp. $UCL^\star$) is 0 (resp. $\lfloor \lambda_0/(1-\beta_0) \rfloor + 1$). $LCL^\star$ (resp. $UCL^\star$) is searched in the set $\{0, 1, \ldots, \lceil \lambda_0/(1-\beta_0) \rceil - 1\}$ (resp. $\{\lfloor \lambda_0/(1-\beta_0) \rfloor + 1, \ldots, UCL\}$).

4.3 Illustrations

The results in this section refer to the overall ARL of the $c-$charts introduced so far to monitor the mean of an INAR(1) process: the $c-$chart with $3-\sigma$ control limits, the unrandomized $c-$chart and the randomized $c-$chart. In all cases, we considered the targets $(\lambda_0, \beta_0) = (3, 0.6), (10, 0.5)$ and $2 \times ARL^{\star} = 740.7407$.

In Table 2 we can find the parameters of each chart: target; control limits, randomization probabilities (if any); in-control overall ARL values.[6] All these performance measures are listed in order corresponding to: the $c-$chart with $3-\sigma$ control limits; the unrandomized $c-$chart; the randomized $c-$chart.

Table 2: LCL, UCL, pair of randomization probabilities, in-control ARL, for $(\lambda_0, \beta_0) = (3, 0.6), (10, 0.5)$ — listed in order corresponding to the $c-$chart with $3-\sigma$ control limits, the unrandomized $c-$chart and the randomized $c-$chart.

(λ_0, β_0)	$[LCL, UCL]$	$(\gamma_{LCL}, \gamma_{UCL})$	In-control ARL
$(3, 0.6)$	$[0, 15]$	—	274.0152
	$[1, 17]$	—	826.0381
	$[1, 17]$	$(0.215880, 0.691129)$	367.5809
$(10, 0.5)$	$[7, 33]$	—	375.3676
	$[8, 35]$	—	666.8522
	$[8, 35]$	$(0.494095, 0.981438)$	368.1995

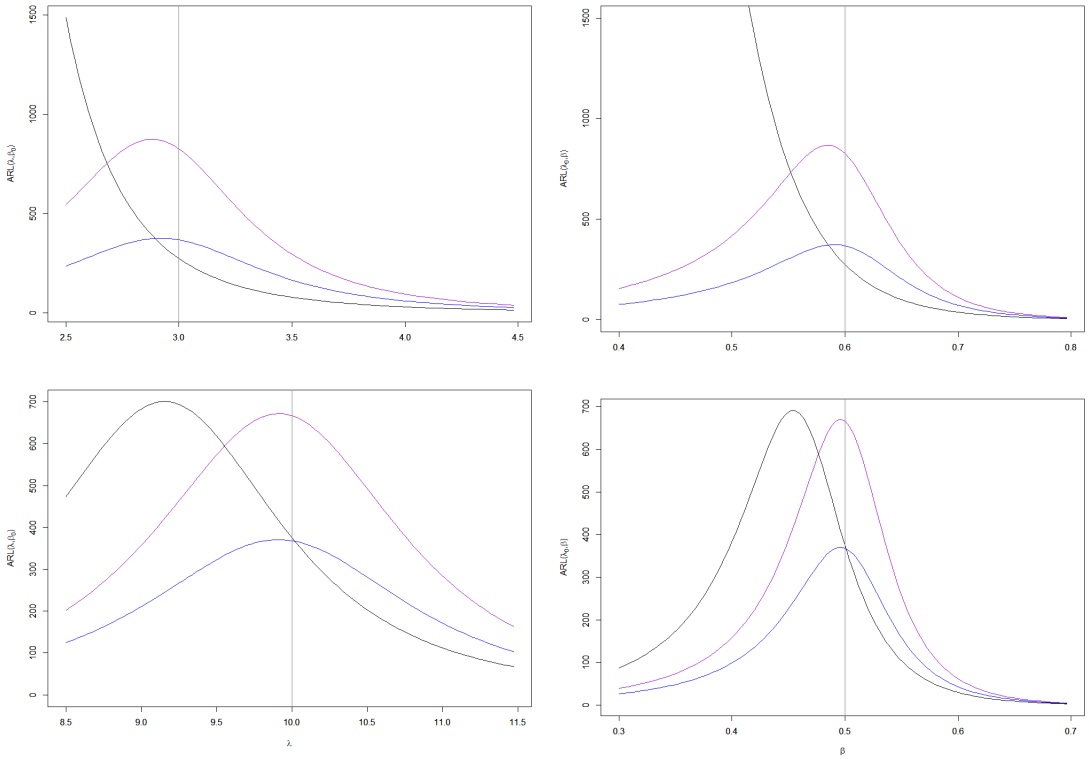

Figure 4: Overall ARL as function of λ (left) and β (right), for the $c-$chart with $3-\sigma$ control limits (curves in **black**), the unrandomized $c-$chart (curves in **violet**) and the randomized $c-$chart (curves in **blue**) — $(\lambda_0, \beta_0) = (3, 0.6), (10, 0.5)$ (top and bottom, resp.).

Table 2 and Figure 4 lead us to believe that: the in-control overall ARL of the unrandomized $c-$chart tends to be much larger than $ARL^{\star}$ and further away from this value than the in-control overall ARL of the $c-$chart with $3-\sigma$ control limits, however, the bias of the overall ARL of the unrandomized $c-$chart is incomparably smaller; the $c-$chart with $3-\sigma$ control limits has a higher detection speed than the other two charts when we are dealing with increases. They also attest that the adoption of randomization

[6] The complexity of these overall functions prevented us to determine the supremum and bias of these RL-related performance measures.

probabilities leads to a substantially improvement of the performance of the resulting $c-$chart: it has an in-control overall ARL much closer to its pre-specified and desired value $ARL^{\star}$ than any of the other two charts; it requires, in average, less time to trigger a signal in the presence of most shifts in either λ or β than to trigger a false alarm, i.e., it is practically (overall) ARL-unbiased, unlike the $c-$chart with $3-\sigma$ control limits and the randomized $c-$chart. For these reasons the use of the randomized $c-$chart should be favoured.

References

Aebtarm, S. and Bouguila, N. (2011). An empirical evaluation of attribute control charts for monitoring defects. *Expert Systems with Applications* **38**, 7869–7880.

Lehmann, E.L. (1986). *Testing Statistical Hypotheses*. Pacific Grove, California: Wadsworth & Brooks/-Cole Advanced Books & Software.

Mckenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* **21**, 645–650.

Morais, M.C. (2012). *Real- and Integer-valued Time Series and Quality Control Charts.* Unpublished manuscript. Departamento de Matemática, Instituto Superior Técnico.

Paulino, S., Morais, M.C., and Knoth, S. (2014). *On ARL-unbiased c-charts for i.i.d. and INAR(1) Poisson counts*. ERCIM 2014 (Session *ES144: Statistical methods for processes in engineering and industry*. Organizer: Christine Mueller). Pisa, December 8, 2014.

Pignatiello, J.J.Jr., Acosta, C.A., and Rao, B.V. (1995). The performance of control charts for monitoring process dispersion. *4th Industral Engineering Research Conference*, 320–328.

Quarteroni, A., Sacco, R., and Saleri, F. (2007). *Numerical Mathematics*. Second Edition. Berlin Heidelberg: Springer.

Ramos, P.A.A.C.F.P. (2013). *Performance analysis of simultaneous control schemes for the process mean (vector) and (co)variance (matrix)*. Ph.D. thesis. Instituto Superior Técnico, Universidade Técnica de Lisboa.

Ryan, T.P. and Schwertman, N.C. (1997). Optimal limits for attribute control charts. *Journal of Quality Techonology* **29**, 86–98.

Silva, I. (2005). *Contributions to the nalysis of Discrete-valued Time Series*. Ph.D. thesis. Faculdade de Ciências da Universidade do Porto, Portugal.

Steutel, F.W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability* **7**, 893–899.

Weiss, C.H. (2007). Controlling correlated processes of Poisson counts. *Quality and Reliability Engineering International* **23**, 741–754.

— (2009). *Categorical Time Series Analysis and Applications in Statistical Quality Control*. Ph.D. thesis. Fakultät fur Mathematik und Informatik der Universitat Würzburg. dissertation.de – Verlag im Internet GmbH.

Wetherill, G.B. and Brown, D.W. (1991). *Statistical Process Control: Theory and Pratice*. London: Chapman and Hall.

White, C.H., Keats, J.B., and Stanley, J. (1997). Poisson CUSUM versus c-chart for defect data. *Quality Engineering* **9**, 673–679.

Wieringa, J.E. (1999). *Statistical Process Control for Serially Correlated Data*. Ph.D. thesis. Faculty of Economics, University of Groningen.