REVSTAT – Statistical Journal
Volume 14, Number 4, October 2016, 455–479

ON HITTING TIMES FOR MARKOV TIME SERIES OF COUNTS WITH APPLICATIONS TO QUALITY CONTROL

Authors: Manuel Cabral Morais
– CEMAT and Department of Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Portugal
maj@math.ist.utl.pt

António Pacheco
– CEMAT and Department of Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Portugal
apacheco@math.ist.utl.pt

Received: July 2014 Revised: March 2015 Accepted: March 2015

Abstract:

- Examples of time series of counts arise in several areas, for instance in epidemiology, industry, insurance and network analysis. Several time series models for these counts have been proposed and some are based on the binomial thinning operation, namely the integer-valued autoregressive (INAR) model, which mimics the structure and the autocorrelation function of the autoregressive (AR) model.

  The detection of shifts in the mean of an INAR process is a recent research subject and it can be done by using quality control charts. Underlying the performance analysis of these charts, there is an indisputable popular measure: the run length (RL), the number of samples until a signal is triggered by the chart. Since a signal is given as soon as the control statistic falls outside the control limits, the RL is nothing but a hitting time.

  In this paper, we use stochastic ordering to assess:

  - the ageing properties of the RL of charts for the process mean of Poisson INAR(1) output;
  - the impact of shifts in model parameters on this RL.

  We also explore the implications of all these properties, thus casting interesting light on this hitting time for a Markov time series of counts.

Key-Words:

- *statistical process control; run length; phase-type distributions; stochastic ordering.*

AMS Subject Classification:

- 60E15, 62P30.

1. THE INAR(1) PROCESS

Time series of counts become apparent in areas such as:

- epidemiology — the number of new cases of some infectious and notifiable diseases is monitored monthly to assess and surveil the incidence of acute viral infections such as poliomyelitis, as reported by Zeger (1988) and Silva (2005, pp. 145–147);
- industry — the monthly number of accidents in a manufacturing plant (Silva *et al.*, 2009), the number of defects per sample (Weiss, 2009a) and the number non-conforming units within a sample of finite size counts (Weiss, 2009b,c) have to be controlled;
- insurance — modelling the number of claim counts is an extremely important part of insurance pricing (Boucher *et al.*, 2008);
- network analysis — the number of intrusions on computers and network systems (Weiss, 2009d, p. 11) also requires surveillance.

In some cases, the integer values of the time series are large and continuous-valued models can be (and are frequently) used. However, when the time series consists only of small integer numbers, ARMA processes are of limited use for modelling purposes, namely because the multiplication of an integer-valued random variable (r.v.) by a real constant may lead to a non-integer r.v. (Silva, 2005, p. 22).

A possible way out is to replace the scalar multiplication by a random operation, such as the binomial thinning operation. This operation can be thought as the scalar multiplication counterpart in the integer-valued setting which preserves the integer structure of the process, it is due to Steutel and Van Harn (1979) and may be stated as follows.

Definition 1.1. Let X be a discrete r.v. with range $\mathbb{N}_0 = \{0, 1, ...\}$ and α a scalar in $[0, 1]$. Then the binomial thinning operation on X results in the following r.v.:

$$\alpha \circ X = \sum_{i=1}^{X} Y_i \,, \tag{1.1}$$

where $\{Y_i : i \in \mathbb{N}\}$ is a sequence of i.i.d. Bernoulli(α) r.v. independent of X.

In this case $\alpha \circ X$ emerges from X by *binomial thinning*, and $\circ$ represents the *binomial thinning operator*. Furthermore, according to Steutel and Van Harn (1979), $\alpha \circ X$ also takes values in $\mathbb{N}_0$ and: $1 \circ X = X$; $0 \circ X = 0$; $\alpha \circ X \mid X =$

$x \sim \text{Bin}(x,\alpha)$; $E(\alpha \circ X) = \alpha \times E(X)$, as in scalar multiplication; $V(\alpha \circ X) = \alpha^2 \times V(X) + \alpha(1-\alpha) \times E(X)$, unlike in scalar multiplication.

Several models for time series of counts have been proposed based on the binomial thinning operation. These models are in general obtained as discrete analogues of the standard linear time series models. For example, the first-order integer-valued autoregressive (INAR(1)) model, introduced by McKenzie (1985) and Al-Osh and Alzaid (1987), mimics the structure and the autocorrelation function of the real-valued first-order autoregressive AR(1) model.

Definition 1.2. Let $\{\epsilon_t : t \in \mathbb{Z}\}$ be a sequence of nonnegative integer-valued independent and identically distributed (i.i.d.) r.v. with range $\mathbb{N}_0$, mean μ_ϵ and variance σ_ϵ^2, and α a scalar in $(0,1)$. Then $\{X_t : t \in \mathbb{Z}\}$ is said to be a INAR(1) process if it satisfies the recursion

$$X_t = \alpha \circ X_{t-1} + \epsilon_t\,, \tag{1.2}$$

where: $\circ$ represents the binomial thinning operator; 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 $\{..., X_{t-2}, X_{t-1}\}$; and ϵ_t and X_{t-1} are assumed to be independent r.v.

Besides taking only nonnegative integer values, the INAR(1) model also differs from the real-valued AR(1) model because the r.v. ϵ_t in this latter are usually interpreted as random noise, whereas in the INAR(1) model they introduce innovation to the process by *keeping the system alive* (Weiss, 2009d, p. 283) with *arrivals*.

Furthermore, the marginal distribution of the INAR(1) process can be expressed in terms of the r.v. ϵ_t (Silva, 2005, p. 35):

$$X_t \overset{d}{=} \sum_{j=0}^{+\infty} \alpha^j \circ \epsilon_{t-j} \tag{1.3}$$

(Al-Osh and Alzaid, 1987), the analogue of the moving average representation of the real-valued AR(1) model.

In the INAR(1) model setting, choosing an adequate family of distributions for the r.v. ϵ_t, say $\mathcal{F}$, so that X_t has a distribution that also belongs to $\mathcal{F}$, leads us to the class of discrete self-decomposable distributions defined by Steutel and Van Harn (1979): the r.v. X, with range $\mathbb{N}_0$, is said to have a discrete self-decomposable distribution if $X = \alpha \circ X' + X_\alpha$, where $\alpha \circ X'$ and X_α are independent, and X' is distributed as X.

It is worth mentioning that if ϵ_t has a discrete self-decomposable distribution such that $E(\epsilon_t) = \mu_\epsilon$ and $V(\epsilon_t) = \sigma_\epsilon^2 < +\infty$ then the INAR(1) process is second order weakly stationary with constant mean and variance function given by $E(X_t) = \frac{\mu_\epsilon}{1-\alpha}$ and $V(X_t) = \frac{\alpha\,\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2}$, respectively (Weiss, 2009d, p. 283). Moreover, the INAR(1) and the AR(1) processes have similar autocorrelation function: $\rho_k = corr(X_t, X_{t-k}) = \alpha^{|k|}$, $k \in \mathbb{Z}$ (Weiss, 2009d, p. 285).

The class of discrete self-decomposable distributions contains the family of Poisson distributions (Silva, 2005, p. 35) and the Poisson INAR(1) process can be defined and characterised.

Definition 1.3. If $\epsilon_t \overset{i.i.d.}{\sim}$ Poisson(λ), $t \in \mathbb{Z}$, then $\{X_t = \alpha \circ X_{t-1} + \epsilon_t : t \in \mathbb{Z}\}$ is said to be a Poisson INAR(1) process.

The Poisson INAR(1) process is a second order weakly stationary process with marginal distribution

$$X_t \sim \text{Poisson}\left(\frac{\lambda}{1-\alpha}\right), \qquad t \in \mathbb{Z}\,, \tag{1.4}$$

and can be characterized as follows, according to Weiss (2009d, p. 283).

Proposition 1.1. *The Poisson INAR*(1) *process is a (time-)homogeneous Markov chain, with state space* $\mathbb{N}_0$ *and one-step transition probability matrix (TPM)* $\mathbf{P}$, *which depends on the values of* λ *and* α *and whose entries are given by*

$$\begin{aligned} p_{i\,j} &\equiv p_{i\,j}(\lambda, \alpha) \\ &= P(X_t = j \mid X_{t-1} = i) \\ &= \sum_{m=0}^{i} P(\alpha \circ X_{t-1} = m \mid X_{t-1} = i) \times P(\epsilon_t = j - m) \\ &= \sum_{m=0}^{\min\{i,j\}} \binom{i}{m} \alpha^m\,(1-\alpha)^{i-m} \times e^{-\lambda}\frac{\lambda^{j-m}}{(j-m)!}\,, \qquad i, j \in \mathbb{N}_0\,. \end{aligned} \tag{1.5}$$

The calculation of $\mathbf{P}$, for a few values of λ and α, led us to believe that no particular features are apparent in this matrix. For instance, even though X_t is a nonnegative r.v., $\mathbf{P}$ has no triangular block of zeros in the lower left hand corner or equal values along a line parallel to the main diagonal, such as the TPM Brook and Evans (1972) or Morais (2002) dealt with in a quality control setting.

Nevertheless, we managed to identify a peculiar and important feature of the TPM associated with a Poisson INAR(1) process: $\mathbf{P}$ is totally positive of order 2 (TP_2), i.e., it is a nonnegative matrix whose 2×2 minors are all nonnegative —

$$p_{i\,j} \times p_{i'\,j'} \;\geq\; p_{i'\,j} \times p_{i\,j'}\,, \qquad i < i', \;\; j < j' \tag{1.6}$$

—, as proved in the next section.

2. DISTINCTIVE FEATURES OF THE POISSON INAR(1) PROCESS

It is well known that the Poisson and binomial probability functions (p.f.),

- $f_{Poi(\lambda)}(x) = e^{-\lambda} \frac{\lambda^x}{x!}$, $x \in \mathbb{N}_0$,
- $f_{Bin(n,p)}(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, ..., n$,

are log-concave in the sense that the likelihood ratio functions

$$\frac{f_{Poi(\lambda)}(x)}{f_{Poi(\lambda)}(x+1)} = \frac{x+1}{\lambda}$$

$$\frac{f_{Bin(n,p)}(x)}{f_{Bin(n,p)}(x+1)} = \frac{(x+1)(1-p)}{(n-x)p}$$

are nondecreasing functions of x over the supports of these p.f. That is to say, the Poisson and binomial distributions have what is also termed the Pólya frequency of order 2 (PF_2) property (Li and Shaked, 1997) or an increasing likelihood ratio (ILR), [1] the strongest ageing property that we consider here.

Furthermore, according to Casella and Berger (2002, p. 391), the families of Poisson and binomial p.f. have monotone likelihood ratio, in particular the following ones:

- $\{f_{Poi(\xi)}(x) : \xi > 0\}$;
- $\{f_{Bin(\xi,p)}(x) : \xi \in \mathbb{N}\}$ (here p is held fixed in $(0,1)$);
- $\{f_{Bin(n,\xi)}(x) : \xi \in (0,1)\}$ (n is fixed in $\mathbb{N}$).

For example, for $\xi_1 \leq \xi_2$,

$$\frac{f_{Poi(\xi_1)}(x)}{f_{Poi(\xi_2)}(x)} = e^{-(\xi_1-\xi_2)} (\xi_1/\xi_2)^x , \qquad x \in \mathbb{N}_0 , \tag{2.1}$$

is a monotone — in this case nonincreasing — function of x. Interestingly enough, if we consider $P(x,\xi) \equiv f_{Poi(\xi)}(x)$ (or $\equiv f_{Bin(\xi,p)}(x)$, $f_{Bin(n,\xi)}(x)$) then $P(x,\xi)$ is a TP_2 function in x and ξ, i.e., the determinant

$$\begin{vmatrix} P(x_1,\xi_1) & P(x_1,\xi_2) \\ P(x_2,\xi_1) & P(x_2,\xi_2) \end{vmatrix} \geq 0 , \qquad x_1 < x_2, \quad \xi_1 < \xi_2 \tag{2.2}$$

(Karlin and Proschan, 1960).

[1] If you define the likelihood ratio function as $\frac{P(X=x+1)}{P(X=x)}$ instead, like Kijima (1997, p. 114) did, then the PF_2 property is equivalent to a decreasing likelihood ratio (DLR), as noted by Kijima (1997, p. 115).

Incidentally, the monotone likelihood ratio character — or TP_2 property — of a family of p.f. is related to the notion of stochastically smaller in the likelihood ratio sense (Ross, 1983, p. 281) stated below.

Definition 2.1. Let X and Y be two discrete r.v. with p.f. $P(X=x)$ and $P(Y=x)$. Then X is said to be stochastically smaller than Y in the likelihood ratio sense — denoted by $X \leq_{lr} Y$ — iff $\frac{P(X=x)}{P(Y=x)}$ is a nonincreasing function of x over the union of the supports of the r.v. X and Y (Shaked and Shanthikumar, 1994, pp. 27–28).

Expectedly, if a family of p.f. has monotone nonincreasing (resp. nondecreasing) likelihood ratio then the associated r.v. stochastically increase (resp. decrease) in the likelihood ratio sense — i.e., if $\xi_1 \leq \xi_2$ then $X(\xi_1) \leq_{lr} X(\xi_2)$ (resp. $X(\xi_1) \geq_{lr} X(\xi_2)$), in short $X(\xi) \uparrow_{lr}$ with ξ (resp. $X(\xi) \downarrow_{lr}$ with ξ). For the families of Poisson and binomial p.f. we have considered:

- $X(\xi) \sim \mathrm{Poi}(\xi) \uparrow_{lr}$ with ξ $(\xi > 0)$;
- $X(\xi) \sim \mathrm{Bin}(\xi, \alpha) \uparrow_{lr}$ with ξ ($\xi \in \mathbb{N}$, here α is held fixed in $(0,1)$);
- $X(\xi) \sim \mathrm{Bin}(n, \xi) \uparrow_{lr}$ with ξ (n fixed in $\mathbb{N}$, $\xi \in (0,1)$).

After these preliminary notions we can state that $X_t \equiv X_t(\lambda,\alpha) \sim \mathrm{Poi}\left(\frac{\lambda}{1-\alpha}\right)$ has the PF_2 property and

$$X_t \equiv X_t(\lambda,\alpha) \uparrow_{lr} \quad \text{with } \lambda,\ \alpha\,. \tag{2.3}$$

But what can be said about the Poisson INAR(1) process $\{X_t \equiv X_t(\lambda,\alpha) : t \in \mathbb{Z}\}$?

- *Is the PF_2 (resp. TP_2) property of the (resp. families of) Poisson and binomial distributions somehow inherited by a Poisson INAR(1) process (resp. a family of Poisson INAR(1) processes)?*
- *If that is the case what are the consequences?*

Proper replies to these queries are provided in this and the following sections.

Proposition 2.1. *The Poisson INAR(1) process $\{X_t : t \in \mathbb{Z}\}$ satisfies*

$$(X_t \mid X_{t-1} = i) \leq_{lr} (X_t \mid X_{t-1} = m)\,, \qquad i \leq m\,, \tag{2.4}$$

for any $t \in \mathbb{Z}$. Equivalently, $(X_t \mid X_{t-1} = i) \uparrow_{lr}$ with i, for $t \in \mathbb{Z}$, and we write

$$\{X_t : t \in \mathbb{Z}\} \in \mathcal{M}_{lr}\,, \tag{2.5}$$

where $\mathcal{M}_{lr}$ stands for the class of stochastic processes that are stochastically monotone in the likelihood ratio sense.

We defer the proof of Proposition 2.1 until a few remarks are made.

$\{X_t : t \in \mathbb{Z}\} \in \mathcal{M}_{lr}$ can be written as $\mathbf{P} \in \mathcal{M}_{lr}$, where $\mathbf{P}$ obviously denotes the TPM of this Markov chain. This feature of $\mathbf{P}$ obviously means that, if we associate a p.f. of a discrete r.v. to one of its rows, the corresponding r.v. stochastically increase in the likelihood ratio sense as we progress along the rows of this stochastic matrix. It also means that

$$\mathbf{P} \in \mathrm{TP}_2 \,, \tag{2.6}$$

as mentioned by Kijima (1997, p. 129, Definition 3.11).

Bearing in mind that the i^{th} row of $\mathbf{P}$ corresponds to the probability (row vector) of the r.v. $(X_{t+1} \mid X_t = i)$ and taking advantage of $\leq_{lr}$ ordering, we are tempted to investigate whether $\mathbf{P} \in \mathcal{M}_{lr}$ by checking if $\frac{p_{i\,j}}{p_{i+1\,j}} \downarrow_j$, over $\mathbb{N}_0$, for any fixed $i \in \mathbb{N}_0$; another possibility of proving Proposition 2.1 would be to check whether $\mathbf{P} \in \mathrm{TP}_2$.

This is not the easiest way of proving that $\mathbf{P} \in \mathrm{TP}_2$, thus the proof of Proposition 2.1 relies on a different reasoning.

Proof: Let us first note that, for $i \in \mathbb{N}_0$, $(X_t \mid X_{t-1} = i) \stackrel{st}{=} Z(i) + \epsilon_t$, where: $Z(0) \stackrel{st}{=} 0$; $Z(i) \sim \mathrm{Bin}(i, \alpha)$, $i \in \mathbb{N}$; $\epsilon_t \sim \mathrm{Poisson}(\lambda)$; $Z(i)$ and ϵ_t are independent r.v.

Now, capitalizing not only on the fact that, for $i \leq m$ $(i, m \in \mathbb{N}_0)$ and α (held fixed in the interval $(0,1)$), $Z(i) \leq_{lr} Z(m)$, but also on the log-concave (or PF_2) character of the p.f. of the summands $Z(i)$ and the independence between $Z(n)$ and ϵ_t $(n = i, m)$, we can invoke the basic decomposition formula (Karlin, 1968, p. 17) or the closure of the stochastic order $\leq_{lr}$ under the sum of independent r.v. with log-concave densities (Shaked and Shanthikumar, 1994, p. 30) [2] to conclude that

$$Z(i) + \epsilon_t \leq_{lr} Z(m) + \epsilon_t, \; i \leq m,$$

thus proving the result. □

The stochastic ordering result in the next proposition may be thought as an extension of the notion of monotone likelihood ratio to the family of Poisson INAR(1) processes, $\{\{X_t \equiv X_t(\lambda, \alpha) : t \in \mathbb{Z}\} : (\lambda, \alpha) \in \mathbb{R}^+ \times (0,1)\}$.

Proposition 2.2. *Let $\{X_t(\lambda, \alpha) : t \in \mathbb{Z}\}$ be a Poisson INAR(1) process such that $X_t(\lambda, \alpha) = \alpha \circ X_{t-1}(\lambda, \alpha) + \epsilon_t(\lambda)$, for $(\lambda, \alpha) \in \mathbb{R}^+ \times (0,1)$. Then*

$$(X_t(\lambda_1, \alpha_1) \mid X_{t-1}(\lambda_1, \alpha_1) = i) \;\leq_{lr}\; (X_t(\lambda_2, \alpha_2) \mid X_{t-1}(\lambda_2, \alpha_2) = m)\,, \quad i \leq m\,, \tag{2.7}$$

for any $0 < \lambda_1 \leq \lambda_2$, $0 < \alpha_1 \leq \alpha_2 < 1$ and $t \in \mathbb{Z}$.

[2] For a slightly stronger result, please refer to Shaked and Shanthikumar (1994, p. 30, Theorem 1.C.5).

Proof: This proposition can be proved in a similar fashion to Proposition 2.1. Thus, let us consider $(X_t(\lambda,\alpha) \mid X_{t-1}(\lambda,\alpha) = i) \stackrel{st}{=} Z(i,\alpha) + \epsilon_t(\lambda)$, where: $Z(0,\alpha) \stackrel{st}{=} 0$; $Z(i,\alpha) \sim \text{Bin}(i,\alpha)$, $i \in \mathbb{N}$; $\epsilon_t(\lambda) \sim \text{Poisson}(\lambda)$; $Z(i,\alpha)$ and $\epsilon_t(\lambda)$ are independent r.v.

By taking into account the monotone likelihood ratio of the Poisson and binomial families, we can add that, for $0 < \lambda_1 \leq \lambda_2$, $0 < \alpha_1 \leq \alpha_2 < 1$: $Z(i,\alpha_1) \leq_{lr} Z(i,\alpha_2)$; $\epsilon_t(\lambda_1) \leq_{lr} \epsilon_t(\lambda_2)$.

If we add to these stochastic ordering results the PF_2 character of all the summands involved and the independence between $Z(i,\alpha_j)$ and $\epsilon_t(\lambda_j)$, for $j = 1,2$, we can use once again the closure of $\leq_{lr}$ under the sum of independent r.v. with log-concave densities to assert that

$$(X_t(\lambda_1,\alpha_1) \mid X_{t-1}(\lambda_1,\alpha_1) = i) \;\leq_{lr}\; (X_t(\lambda_2,\alpha_2) \mid X_{t-1}(\lambda_2,\alpha_2) = i)\;.$$

Finally, take notice that $\{X_t(\lambda_j,\alpha_j) : t \in \mathbb{Z}\} \in \mathcal{M}_{lr}$, for $j = 1,2$, as a consequence, $(X_t(\lambda_2,\alpha_2) \mid X_{t-1}(\lambda_2,\alpha_2) = i) \leq_{lr} (X_t(\lambda_2,\alpha_2) \mid X_{t-1}(\lambda_2,\alpha_2) = m)$, for $i \leq m$, and

$$(X_t(\lambda_1,\alpha_1) \mid X_{t-1}(\lambda_1,\alpha_1) = i) \;\leq_{lr}\; (X_t(\lambda_2,\alpha_2) \mid X_{t-1}(\lambda_2,\alpha_2) = m)\,, \qquad i \leq m\,.$$

This ends the proof. □

Corollary 2.1. *Let $P(n,\lambda,\alpha) \equiv P[X_t(\lambda,\alpha) = n \mid X_{t-1}(\lambda,\alpha) = i]$. Then $P(n,\lambda,\alpha)$ is TP_2 both as a function of n and λ (with α held fixed) and as a function of n and α (for fixed λ).*

As for the implications of propositions 2.1 and 2.2 — in particular on what the random time the Poisson INAR(1) process needs to exceed a critical level x is concerned — we refer the reader to the next sections.

3. VITAL PROPERTIES OF THE HITTING TIMES FOR POISSON INAR(1) PROCESSES

Hitting times (HT) arise naturally in level-crossing problems in several areas:

- in reliability theory, HT of appropriate stochastic processes often represent the time to failure of a device subjected to shocks (and wear), which fails when its damage level crosses a critical value (Li and Shaked, 1995);

- in queueing systems, the identity of the first customer whose waiting time exceeds a critical threshold is a HT and a relevant performance measure (Greenberg, 1997);
- HT become also apparent while dealing with the problem of the first detection of words in random sequences of letters from a finite alphabet (De Santis and Spizzichino, 2014).

Considering the applications of Poisson INAR(1) processes, studying the HT of these stochastic processes is surely of vital importance.

More than on the distribution of HT, in this section we are interested in assessing the ageing properties of the HT and the impact of an increase in

- the critical level,
- the initial state, and
- the parameters λ and α

on the associated HT. Needless to say that dealing with a stochastic process with a TP_2 TPM will play a major role in the derivation of all the results.

The conditions under which HT possess specific ageing properties have been extensively studied by many authors (see e.g.: Brown and Chaganty, 1983; Assaf *et al.*, 1985; Karasu and Özekici, 1989), and rigorously reported by Li and Shaked (1997). Furthermore, these conditions are closely related to the stochastic monotonicity character of the underlying process, as noted by Li and Shaked (1995).

The next result can be translated as follows in our specific setting: the PF_2 property of the Poisson and binomial distributions is shared with a particular HT. It is a consequence of an important result that can be traced back to Karlin (1964).

Proposition 3.1. *Let:* $\{X_t : t \in \mathbb{N}_0\}$ *be a Poisson INAR(1) process with initial state* $X_0 = 0$*;* $HT^0 = \min\{t \in \mathbb{N} : X_t > x \mid X_0 = 0\}$ *be the random number of transitions needed to exceed the critical level* x $(x \in \mathbb{N}_0)$ *starting from the initial state* 0*. Then* $HT^0 \in PF_2$.

Proof: This proposition follows from Theorem 3.1 by Assaf *et al.* (1985), who pointed out that their result was essentially proved by Karlin (1964, pp. 93–94). □

Let us remind the reader that, since $HT^0 \in \text{PF}_2$, $\{X_t : t \in \mathbb{N}_0\}$ is said to be a PF_2 process (Shaked and Li, 1997, p. 12). We ought to also mention that Proposition 3.1 can be referred to as the *PF_2 Theorem* (Shaked and Li, 1997, p. 12) for the Poisson INAR(1) process.

The next result translates the stochastic impact of an increase of the critical value x.

Proposition 3.2. *Let:* $\{X_t : t \in \mathbb{N}\}$ *be a Poisson INAR(1) process, where* $X_1 \sim$ *Poisson* $\left(\frac{\lambda}{1-\alpha}\right)$; $HT_x = \min\{t \in \mathbb{N} : X_t > x\}$ *be the random time at which the process exceeds the critical level* x ($x \in \mathbb{N}_0$)*. Then* $HT_x \uparrow_{lr}$ *with* x.

Proof: Since X_t can be written as a sum of r.v. with PF_2 p.f.,

$$X_t = \sum_{j=0}^{t-1} \alpha^j \circ \epsilon_{t-j} + \alpha^t \circ X_1 \ , \qquad t \in \mathbb{N}\backslash\{1\} \ ,$$

we can apply Theorem 2 from Karlin and Proschan (1960) and conclude that the p.f. of HT_x,

$$P(n,x) \ \equiv \ P(HT_x = n) \ = \ P(X_n > x;\, X_j \le x,\, j = 1,2,...,n-1) \ , \qquad n \in \mathbb{N} \ ,$$

as a function of n and x, is TP_2. Consequently, $HT_x \uparrow_{lr}$ with x. □

The next proposition shows how the TP_2 character of the Poisson INAR(1) process is crucial to guarantee a specific stochastic decrease of the HT with respect to the initial value of this process.

Proposition 3.3. *Let:* $\{X_t : t \in \mathbb{N}_0\}$ *be a Poisson INAR(1) process with initial state* $X_0 = i$ ($i \in \mathbb{N}_0$)*;* $HT^i = \min\{t \in \mathbb{N} : X_t > x \mid X_0 = i\}$ *be the random number of transitions needed to exceed the critical level* x ($x \in \mathbb{N}_0$) *starting from the initial state* i*. Then* $HT^i \downarrow_{lr}$ *with* i.

Proof: Since $\mathbf{P} \in \mathrm{TP}_2$, we are allowed to invoke Theorem 2.1 from Karlin (1964, pp. 42–43) and assert that the p.f. of HT^i,

$$\begin{aligned} P(n,i) \ &\equiv \ P(HT^i = n) \\ &= \ P(X_n > x; X_j \le x,\, j = 1,2,...,n-1) \ , \qquad n \in \mathbb{N} \ , \end{aligned}$$

as a function of n and i, is sign reverse rule of order 2 (RR_2), i.e.,

$$P(n,i) \times P(n',i') \ \le \ P(n',i) \times P(n,i') \ , \qquad n \le n', \ \ i \le i' \ .$$

This inequality is equivalent to

$$\frac{P(HT^i = n)}{P(HT^{i'} = n)} \le \frac{P(HT^i = n')}{P(HT^{i'} = n')} \ , \qquad n < n', \ \ i < i' \ ,$$

that is, $HT^i \ge_{lr} HT^{i'}$, for $i \le i'$. □

So far we were not able to prove the following conjecture regarding a stochastic implication of an increase in parameter λ.

- Let $\{X_t(\lambda,\alpha) : t \in \mathbb{N}_0\}$ be a Poisson INAR(1) process with initial state $X_0(\lambda,\alpha) = 0$ and $HT^0(\lambda,\alpha) = \min\{t \in \mathbb{N} : X_t(\lambda,\alpha) > x \mid X_0(\lambda,\alpha) = 0\}$. Then $HT^0(\lambda,\alpha) \downarrow_{lr}$ with λ.

Morais (2002, p. 47) discusses thoroughly the problems that arise when we try to prove results such as the one stated in previous conjecture while dealing with hitting times for discrete-time Markov chains arising in quality control. As a consequence we have to content ourselves with further — yet weaker — stochastic ordering results; they are stated in Section 5 and are particularly relevant in the performance analysis of a quality control chart, described in Section 4 and meant to detect changes in the mean of a Poisson INAR(1) process.

4. CONTROLLING THE MEAN OF A POISSON INAR(1) PROCESS

Although quality has long been considered absolutely relevant, we have to leap to the beginning of the 20th century to meet the founder of Statistical Process Control (SPC) (Ramos, 2013, p. 2). When Walter A. Shewhart joined the Western Electric Company, industrial quality exclusively relied on the inspection of end products and the removal of defective items; however, this physicist, engineer and statistician soon realized that it was important to control not only the finished product but also the process responsible for its production (ASQ, n.d.).

Shewhart essentially suggested that we should monitor a (production) process by:

- choosing a measurable characteristic, say X, of this process;
- selecting a relevant parameter;
- collecting data on a regular basis;
- plotting the observed value of a control statistic against time and comparing it with appropriate control limits;
- triggering a signal if the observed value of the statistic is beyond these control limits.

The resulting graphic device is called a quality control chart, undoubtedly one of the most important tools of SPC.

Control charts are used with the purpose of establishing whether the process is operating within its limits of expected variation (Nelson, 1982, p. 176), and to detect changes in process parameters which may indicate a deterioration in quality. The control chart should be set in such way that a change in the parameter is detected as fast as possible without triggering false alarms too frequently.

The detection of changes in the mean of an i.i.d. process of Poisson counts can be done by making use of quality control charts such as the *c*-chart pioneered by Shewhart (Montgomery, 2009, p. 309), the CUSUM chart (Brook and Evans, 1972; Gan, 1993) or the EWMA chart (Gan, 1990). However, autocorrelation often arises, severely changing the performance of all quality control charts relying on the assumption that the observations refer to i.i.d. r.v., hence the use of the charts such like the ones proposed by Weiss (2009, Chap. 20).

Throughout the remainder of this paper, we assume that:

- the target value of the process mean is $\frac{\lambda_0}{1-\alpha_0}$;
- the purpose of using an upper one-sided control chart is to detect an aggravation in the mean number of defects, from its target value $\frac{\lambda_0}{1-\alpha_0}$ to $\frac{\lambda}{1-\alpha}$, due to a change either from λ_0 to λ or from α_0 to α.

Consequently, we proceed to describe the upper one-sided version of the *c*-control chart found in Weiss (2007) and Weiss (2009d, p. 419).

Definition 4.1. Let $\{X_t \equiv X_t(\lambda,\alpha) : t \in \mathbb{N}_0\}$ be a Poisson INAR(1) process, where denotes the number of defects in sample t, for $t \in \mathbb{N}$, given that the process mean is at level $\frac{\lambda}{1-\alpha}$. Then $x_t \equiv x_t(\lambda,\alpha)$ is the observed value of the control statistic of the upper one-sided *c*-chart for the mean of this process and this chart triggers a signal at time t ($t \in \mathbb{N}$) if

$$x_t \,>\, UCL \,=\, \frac{\lambda_0}{1-\alpha_0} + k \times \sqrt{\frac{\lambda_0}{1-\alpha_0}}\,, \tag{4.1}$$

where k is a positive constant chosen in such way that increases in the process mean $\frac{\lambda}{1-\alpha}$ are detected as quickly as possible and false alarms are rather unfrequent.

Since points lying above the upper control limit (UCL) indicate a potential increase in the process mean that should be investigated and eliminated, the performance of this control chart is unsurprisingly assessed by making use of the run length (RL), the random number of samples collected before a signal (either false or a valid alarm) is triggered by the chart. Hence the RL coincides with the following HT for the Poisson INAR(1) process

$$HT(\lambda,\alpha) \,=\, \min\{t \in \mathbb{N} : X_t(\lambda,\alpha) > x\}\,, \tag{4.2}$$

where $x = \lfloor UCL \rfloor$ is the integer part of the upper control limit defined in (4.1).

In the next section we shall start by addressing a few weaker ageing notions with more tangible interpretations/implications than the PF_2 property of hitting times such as the RL of the upper one-sided c-chart for the mean of a Poisson INAR(1) process.

5. OTHER PROPERTIES OF THE HITTING TIMES FOR THE POISSON INAR(1) PROCESS

Let us start this section by reminding the reader of the ageing notions of increasing failure rate (IFR), new better than used (NBU) and new better than used in expectation (NBUE).

Definition 5.1. The nonnegative integer valued r.v. Y is said to be:

- increasing failure rate — $Y \in \text{IFR}$ — if $h_Y(m) = \frac{P(Y=m)}{P(Y\geq m)} \uparrow_{m\in\mathbb{N}}$;
- new better than used — $Y \in \text{NBU}$ — if $P(Y > j) \geq P(Y - m > j \mid Y > m)$, $m, j \in \mathbb{N}_0$;
- new better than used in expectation — $Y \in \text{NBUE}$ — if $E(Y) \geq E(Y - m \mid Y > m)$, $m \in \mathbb{N}_0$.

Please note that $Y \in \text{PF}_2 \Longrightarrow Y \in \text{IFR} \Longrightarrow Y \in \text{NBU}$ (Kijima, 1997, p. 118), and, clearly, $Y \in \text{NBU} \Longrightarrow Y \in \text{NBUE}$.

By capitalizing on the TP_2 character of the TPM of the Poisson INAR(1) process and on the fact that $Y \in \text{PF}_2 \Longrightarrow Y \in \text{IFR}$, we can immediately conclude that the RL of the upper one-sided c-chart starting with a zero value,

$$HT^0(\lambda,\alpha) \equiv \min\{t \in \mathbb{N} : X_t(\lambda,\alpha) > x \mid X_0(\lambda,\alpha) = 0\}\,, \tag{5.1}$$

has PF_2 character and therefore

$$HT^0(\lambda,\alpha) \in \text{IFR}\,, \tag{5.2}$$

as illustrated by Example 5.1.

Note, however, that, according to the *IFR Theorem* (Shaked and Li, 1997, p. 12), this property is ensured by a weaker condition than $\mathbf{P} \in \text{TP}_2$. In fact, if we let $\mathbf{Q} = [q_{ij}]_{i,j} \equiv [\sum_{k\leq j} p_{ik}]_{i,j}$ denote the matrix of left partial sums of $\mathbf{P}$ then $\mathbf{Q} \in \text{TP}_2$ would have been sufficient to have $HT^0(\lambda,\alpha) \in \text{IFR}$.

Example 5.1. Assume the number of defects in the t^{th} random sample of fixed size (say n) is modelled by a Poisson INAR(1) process $\{X_t \equiv X_t(\lambda,\alpha) : t \in \mathbb{N}_0\}$.

Consider that the detection of increases in the expected value of $X_t(\lambda, \alpha)$, $\mu = \frac{\lambda}{1-\alpha}$, is done by means of the upper one-sided c-chart Poisson chart in Definition 4.1.

The performance of this chart is measured via the HT

$$HT^i(\lambda, \alpha) \equiv \min\{t \in \mathbb{N} : X_t(\lambda, \alpha) > x \mid X_0(\lambda, \alpha) = i\} , \tag{5.3}$$

where i $(i = 0, 1, ..., x)$ is the fixed value assigned to $X_0(\lambda, \alpha)$ by the quality practitioner. If $i = 0$ (resp. $i > 0$) no head start (resp. a head start) has been given to the chart.

Moreover, assume the constant k in the expression of the upper control limit in (4.1) was set in such way that the average run length (ARL) when the values of λ and α are on-target, $E[HT^i(\lambda_0, \alpha_0)]$, is reasonably large, say larger than 100 samples.

It is well known that, for each x, $HT^i(\lambda, \alpha)$ has exactly the same distribution as the time to absorption of a Markov chain with state space $\{0, 1, ..., x+1\}$ and TPM represented in partitioned form,

$$\begin{bmatrix} \mathbf{Q} & (\mathbf{I}-\mathbf{Q})\,\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix} , \tag{5.4}$$

where:

- $\mathbf{Q} \equiv \mathbf{Q}(\lambda, \alpha) = [p_{ij}(\lambda, \alpha)]_{i,j=0}^{x}$;
- $\mathbf{I}$ is the identity matrix with rank $x + 1$;
- $\underline{\mathbf{1}}$ (resp. $\underline{\mathbf{0}}^\top$) is a column vector (resp. row vector) of $x + 1$ ones (resp. zeros).

The associated expected value, survival function and failure (or alarm) rate function are given by

$$E[HT^i(\lambda_0, \alpha_0)] = \underline{\mathbf{e}}_i^\top \times [\mathbf{I} - \mathbf{Q}(\lambda, \alpha)]^{-1} \times \underline{\mathbf{1}} , \tag{5.5}$$

$$\overline{F}_{HT^i(\lambda,\alpha)}(m) = \underline{\mathbf{e}}_i^\top \times [\mathbf{Q}(\lambda, \alpha)]^m \times \underline{\mathbf{1}} , \qquad m \in \mathbb{N} , \tag{5.6}$$

$$\begin{aligned} h_{HT^i(\lambda,\alpha)}(m) &= \frac{P[HT^i(\lambda, \alpha) = m]}{P[HT^i(\lambda, \alpha) \geq m]} \\ &= \frac{\overline{F}_{HT^i(\lambda,\alpha)}(m-1) - \overline{F}_{HT^i(\lambda,\alpha)}(m)}{\overline{F}_{HT^i(\lambda,\alpha)}(m-1)} , \qquad m \in \mathbb{N} \end{aligned} \tag{5.7}$$

(respectively), where $\underline{\mathbf{e}}_i$ represents the $(i + 1)^{\text{th}}$ vector of the orthonormal basis of $\mathbb{R}^{x+1}$.

The failure (or alarm) rate function was proposed by Margavio *et al.* (1995) and represents the conditional probability that the critical level x has been exceeded at time m, given that this threshold has not been crossed before. Even though the alarm rate function is defined in terms of HT probabilities, it will bring forth insights into the chart detection capability, as we progress with the sampling procedure, insights that cannot be provided by the ARL $E[HT^i(\lambda_0,\alpha_0)]$, as illustrated by Margavio *et al.* (1995) and Morais and Pacheco (2012).

The parameters $\lambda_0 = 1$, $\alpha_0 = 0.4$ and $k = 3$ yield an upper one-sided c-chart for the mean of the Poisson INAR(1) process with $x = 5$ and in-control ARL equal to $E[HT^0(\lambda_0,\alpha_0)] = 157.457$ and $E[HT^3(\lambda_0,\alpha_0)] = 153.971$.

$E[HT^i(\lambda_0,\alpha_0)]$ should be calculated for a wide range of changes in the parameters λ and α in order to assess the chart detection ability to several out-of-control conditions. For instance, an increase of 10% in λ leads to out-of-control ARL of $E[HT^0(1.1\,\lambda_0,\alpha_0)] = 104.554$ and $E[HT^3(1.1\,\lambda_0,\alpha_0)] = 101.548$, whereas an increase of the same magnitude in α yields $E[HT^0(\lambda_0,1.1\,\alpha_0)] = 120.560$ and $E[HT^3(1.1\,\lambda_0,1.1\,\alpha_0)] = 117.018$.

The values and graphs of the alarm rate function in Table 1 and Figure 1 give additional insights to the performance of the chart as we proceed with the sampling, and to the impact of the adoption of a head start.

Table 1: Values of: the alarm rate function $h_{HT^i(\lambda,\alpha)}(m)$, for $\lambda_0 = 1$, $\alpha_0 = 0.4$, $x = 5$, $i = 0,2$ and several values of m; the associated ARL values.

m	$h_{HT}(m)$					
	$HT^0(\lambda_0,\alpha_0)$	$HT^3(\lambda_0,\alpha_0)$	$HT^0(1.1\,\lambda_0,\alpha_0)$	$HT^3(1.1\,\lambda_0,\alpha_0)$	$HT^0(\lambda_0,1.1\,\alpha_0)$	$HT^3(\lambda_0,1.1\,\alpha_0)$
1	0.000594	0.012317	0.000968	0.016344	0.000594	0.014636
2	0.003143	0.009673	0.004939	0.013533	0.003591	0.012668
3	0.005033	0.007672	0.007755	0.011176	0.006166	0.010237
4	0.005884	0.006891	0.008980	0.010259	0.007468	0.009171
5	0.006220	0.006598	0.009449	0.009919	0.008035	0.008733
10	0.006422	0.006424	0.009722	0.009725	0.008427	0.008435
20	0.006423	0.006423	0.009724	0.009724	0.008432	0.008432
30	0.006423	0.006423	0.009724	0.009724	0.008432	0.008432
40	0.006423	0.006423	0.009724	0.009724	0.008432	0.008432
50	0.006423	0.006423	0.009724	0.009724	0.008432	0.008432
$E(HT)$	157.457	153.971	104.554	101.548	120.560	117.018

When no head start has been adopted, the IFR character of HT means that signaling, given that no observation has previously exceeded the upper control limit, becomes more likely as we proceed with the collection of samples, as previously noted by Morais and Pacheco (2012) for other control charts for i.i.d. output, contributing to a considerable decrease of the inconvenient initial inertia of this chart in the out-of-control situation.

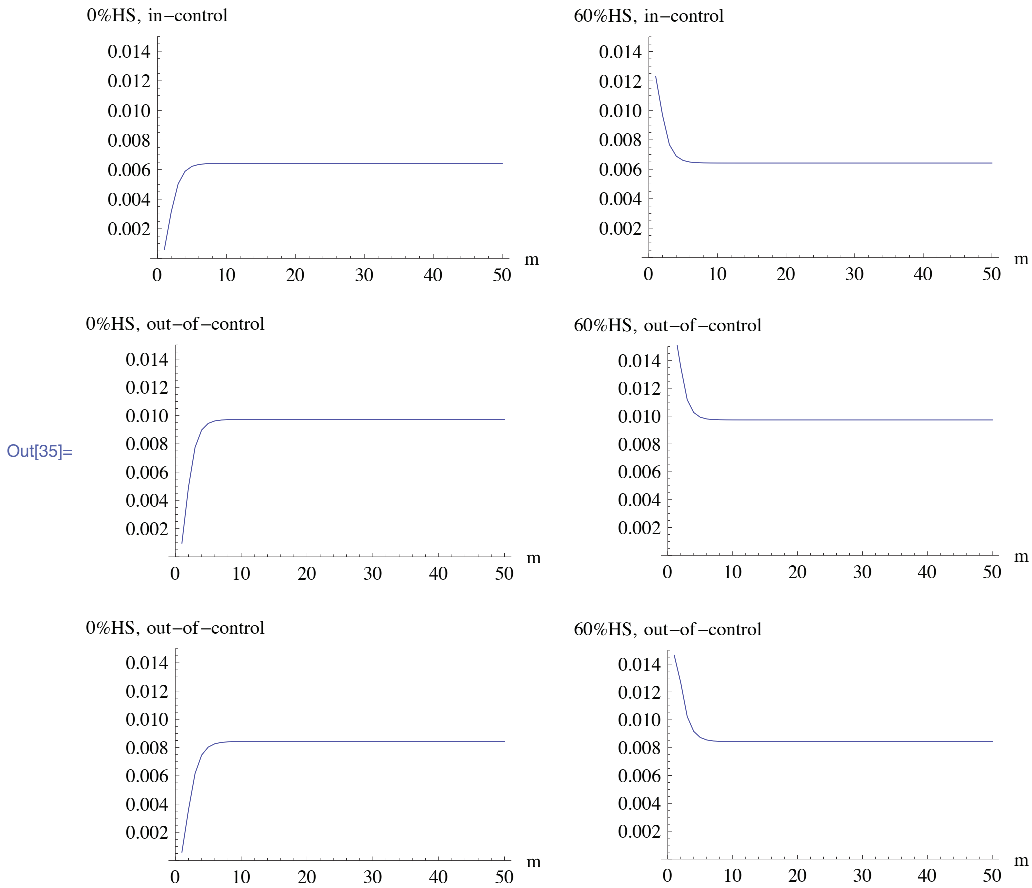

Figure 1: Alarm rates of $HT^i(\lambda, \alpha)$, for $i = 0$ (on the left) and $i = 3$ (on the right), and $(\lambda, \alpha) = (\lambda_0, \alpha_0)$, $(1.1\,\lambda_0, \alpha_0), (\lambda_0, 1.1\,\alpha_0)$ (top, center, bottom).

We ought to also note that, although the adoption of a 60% head start is responsible for mild reductions in the in-control and out-of-control ARL, adding this head-start radically changes the monotone behaviour of the alarm rate function, as shown by Figure 1: $HT^3(\lambda_0, \alpha_0)$, $HT^3(1.1\,\lambda_0, \alpha_0)$, $HT^3(\lambda_0, 1.1\,\alpha_0) \notin IFR$. Figure 1 also suggests a practical meaning of the impact of the adoption of a head start in the absence and in the presence of assignable causes: the false alarm (resp. valid signal) rate conveniently (resp. inconveniently) increases at the first samples.

We strongly believe that the results in this example show that the alarm rate function provides a more insightful portrait of the performance of the control chart than the one based on the ARL.

Even though $HT^i(\lambda,\alpha)$ may not be IFR for $i \neq 0$, it has a weaker ageing property:

$$(5.8) \qquad HT^i(\lambda,\alpha) \in \text{NBU}\,, \qquad i = 0,1,...,x\,.$$

Brown and Chaganty (1983) devised a sufficient condition to deal with a HT with such property. However, stating this condition requires the definition of another stochastic order, a related class of stochastically monotone matrices/processes and a ordering between stochastic matrices.

Definition 5.2. Let X and Y be two nonnegative integer r.v. Then X is said to be stochastically smaller than Y in the usual sense — $X \leq_{st} Y$ — if

$$(5.9) \qquad P(X > m) \leq P(Y > m)\,, \qquad m \in \mathbb{N}_0$$

(Shaked and Shanthikumar, 1994, p. 3).

If the Markov chain $\{X_t : t \in \mathbb{Z}\}$ with TPM $\mathbf{P}$ satisfies

$$(5.10) \qquad (X_t \mid X_{t-1} = i) \leq_{st} (X_t \mid X_{t-1} = m)\,, \qquad i \leq m\,,$$

for any $t \in \mathbb{Z}$, then it is said to be stochastically monotone in the usual sense (Kijima, 1995, p. 129). In this case we write $\{X_t : t \in \mathbb{Z}\} \in \mathcal{M}_{st}$ or $\mathbf{P} \in \mathcal{M}_{st}$, where $\mathcal{M}_{st}$ denotes the class of stochastic processes that are stochastically monotone in the usual sense.

Let $\mathbf{P}$ and $\mathbf{P}'$ two stochastic matrices governing two Markov chains $\{X_t : t \in \mathbb{Z}\}$ and $\{X'_t : t \in \mathbb{Z}\}$ defined in the same state space. Then $\mathbf{P}$ is said to be smaller than $\mathbf{P}'$ in the usual sense (or in the Kalmykov sense) — $\mathbf{P} \leq_{st} \mathbf{P}'$ — if

$$(5.11) \qquad (X_t \mid X_{t-1} = i) \leq_{st} \left(X'_t \mid X'_{t-1} = m\right)\,, \qquad i \leq m\,.$$

Since the stochastic orders $\leq_{st}$ and $\leq_{lr}$ can be related — after all Theorem 1.C.1 of Shaked and Shanthikumar (2007, p. 42) leads to $X \leq_{lr} Y \Longrightarrow X \leq_{st} Y$ —, we naturally have $\mathbf{P} \in \mathcal{M}_{lr} \Longrightarrow \mathbf{P} \in \mathcal{M}_{st}$. Furthermore, Brown and Chaganty (1983) proved that $\mathbf{P} \in \mathcal{M}_{st}$ is sufficient to be dealing with NBU HT. Consequently, the Poisson INAR(1) process satisfies what Shaked and Li (1997, p. 13) called the *NBU Theorem*:

$$(5.12) \qquad HT^i(\lambda,\alpha) \in \text{NBU}\,, \qquad i = 0,1,...,x\,.$$

Consequently,

$$(5.13) \qquad HT^i(\lambda,\alpha) \in \text{NBUE}\,, \qquad i = 0,1,...,x\,.$$

By invoking Corollary 2.1 from Morais and Pacheco (2012),[3] we can add an implication of (5.13):

$$(5.14) \qquad V[HT^i(\lambda,\alpha)] \leq V(Y)\,,$$

[3]This result reads as follows: if X is a discrete NBUE r.v. and Y is a geometric r.v. such that $E(X) \leq E(Y)$, then $V(X) \leq V(Y)$.

where is Y has a geometric distribution with parameter $p \leq \frac{1}{E[HT^i(\lambda,\alpha)]}$. In other words, if we hypothetically replace the upper one-sided c-chart by a chart with a geometrically distributed RL and this results in an aggravation of the ARL, then an increase in the standard deviation of the RL will also follow.

As put by Morais and Pacheco (2012), quality control practitioners should be reminded of Chebyshev's inequality and that considerable benefit it is to be gained by adopting a chart with a smaller standard deviation of the RL, thus diminishing the possibility of having observations beyond the UCL much sooner or much later than expected.

Now, we turn our attention to the HT for a Poisson INAR(1) process whose initial value is a r.v. $X_0(\lambda,\alpha) \sim \text{Poisson}\left(\frac{\lambda}{1-\alpha}\right)$. Following Weiss (2009d, p. 422), this could be called *overall RL* of the upper one-sided c-chart for the mean of such a process. This HT is a mixture of $(x+1)$ r.v. $HT^i(\lambda,\alpha)$, $i = 0,1,...,x$, and a zero-valued r.v. because any value of $X_0(\lambda,\alpha)$ beyond the UCL would lead to a null RL. The associated weights are $P[X_0(\lambda,\alpha) = i]$, $i = 0,1,...,x$, and $P[X_0(\lambda,\alpha) > x]$.

The next proposition provides a thorough characterization of this HT, represented from now on by $HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)$.

Proposition 5.1. *Let* $\{X_t(\lambda,\alpha) : t \in \mathbb{N}_0\}$ *be a Poisson INAR(1) process, where* $X_0(\lambda,\alpha) \sim$ *Poisson* $\left(\frac{\lambda}{1-\alpha}\right)$. *Then the HT*

$$HT^{X_0(\lambda,\alpha)}(\lambda,\alpha) = \min\{t \in \mathbb{N}_0 : X_t(\lambda,\alpha) > x\} \tag{5.15}$$

has expected value, survival function and failure rate function given by

$$E[HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)] = \sum_{i=0}^{x} E[HT^i(\lambda,\alpha)] \times P\left[X_0(\lambda,\alpha) = i\right], \tag{5.16}$$

$$\overline{F}_{HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)}(m) = \begin{cases} 1 - P[X_0(\lambda,\alpha) > x]\,, & m = 0\,, \\ 1 - \sum_{u=0}^{x} F_{HT^i(\lambda,\alpha)}(m) \times P\left[X_0(\lambda,\alpha) = i\right] & \\ \quad - P[X_0(\lambda,\alpha) > x]\,, & m \in \mathbb{N}\,, \end{cases} \tag{5.17}$$

$$\begin{aligned} h_{HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)}(m) &= \frac{P[HT^{X_0(\lambda,\alpha)} = m]}{P[HT^{X_0(\lambda,\alpha)} \geq m]} \\ &= \begin{cases} P[X_0(\lambda,\alpha) > x]\,, & m = 0\,, \\ 1 - \frac{\overline{F}_{HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)}(m)}{\overline{F}_{HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)}(m-1)}\,, & m \in \mathbb{N}_0 \end{cases} \end{aligned} \tag{5.18}$$

(respectively).

Not much can be said about the ageing properties of $HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)$ because the classes of NBU and NBUE r.v. are not closed under mixtures even though they are closed under convolutions (Barlow and Proschan, 1975/1981, pp. 104 and 187).

Table 2: Values of: the alarm rate function $h_{HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)}(m)$, for $\lambda_0 = 1$, $\alpha_0 = 0.4$, $x = 5$ and several values of m; the associated ARL values.

m	$h_{HT^{X_0}}(m)$		
	$HT^{X_0(\lambda_0,\alpha_0)}(\lambda_0,\alpha_0)$	$HT^{X_0(1.1\,\lambda_0,\alpha_0)}(1.1\,\lambda_0,\alpha_0)$	$HT^{X_0(\lambda_0,1.1\,\alpha_0)}(\lambda_0,1.1\,\alpha_0)$
0	0.007302	0.011272	0.010011
1	0.006551	0.009957	0.008677
2	0.006462	0.009795	0.008512
3	0.006437	0.009748	0.008462
4	0.006428	0.009732	0.008444
5	0.006425	0.009727	0.008437
10	0.006423	0.009724	0.008432
20	0.006423	0.009724	0.008432
30	0.006423	0.009724	0.008432
40	0.006423	0.009724	0.008432
50	0.006423	0.009724	0.008432
$E(HT^{X_0})$	154.525	101.648	70.147

Nonetheless, extensive numerical results, illustrated here by the values in Table 2 and the graphs in Figure 2, suggest that we are dealing with a HT with a decreasing failure rate.

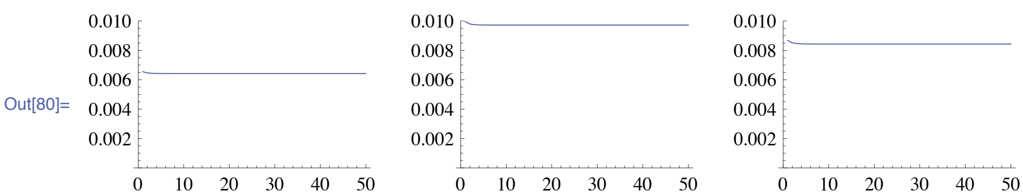

Figure 2: Alarm rates of $HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)$, for $(\lambda,\alpha) = (\lambda_0,\alpha_0), (1.1\,\lambda_0,\alpha_0), (\lambda_0, 1.1\,\alpha_0)$ (left, center, right).

Finally, we qualitatively assess the impact of an increase in λ or α on $HT^{X_0(\lambda,\alpha)}(\lambda,\alpha)$ in the next proposition, stated without a proof since it follows from the fact that $\leq_{lr} \Longrightarrow \leq_{st}$ and an adaptation of Corollary 3.13 from Morais (2002, p. 46).

Proposition 5.2. *Let $\{X_t(\lambda_j,\alpha_j) : t \in \mathbb{N}_0\}$ be a Poisson INAR(1) process with initial state $X_0(\lambda_j,\alpha_j)$, for $j = 1,2$. If $\lambda_1 \leq \lambda_2$ and $\alpha_1 \leq \alpha_2$ then*

$X_0(\lambda_1,\alpha_1) \le_{st} X_0(\lambda_2,\alpha_2)$ and more importantly:

$$\mathbf{P}(\lambda_1,\alpha_1) \ \le_{st} \ \mathbf{P}(\lambda_2,\alpha_2) \ , \tag{5.19}$$

$$HT^{X_0(\lambda_1,\alpha_1)}(\lambda_1,\alpha_1) \ \ge_{st} \ HT^{X_0(\lambda_2,\alpha_2)}(\lambda_2,\alpha_2) \ , \tag{5.20}$$

that is, $HT^{X_0(\lambda,\alpha)}(\lambda,\alpha) \downarrow_{st}$ with λ,α.

The stochastic ordering result (5.20) from Proposition 5.2 can be interpreted as follows: the upper one-sided c-chart for the mean of a Poisson INAR(1) process stochastically increases its detection speed as the increase in λ or α becomes more severe. This result parallels with the notion of a sequentially repeated test possessing what Ramachandran (1958) called the *monotonicity property*.

Results such as (5.20) also remind us of the notion of the level crossing ordering introduced by Irle and Gani (2001). For instance, a Markov chain $\{Y_t : t\in\mathbb{N}_0\}$ is slower in level crossing than a Markov chain $\{Z_t : t\in\mathbb{N}_0\}$ if it takes $\{Y_t : t\in\mathbb{N}_0\}$ stochastically longer than $\{Z_t : t\in\mathbb{N}_0\}$ to exceed any given level. Thus, instead of comparing two stochastic processes through all their finite dimensional distributions as for st-ordering, the lc-ordering compare two stochastic processes through their hitting times (Ferreira and Pacheco, 2007).

In light of this definition we can add that result (5.20) translates as follows: for $\lambda_1 \le \lambda_2$ and $\alpha_1 \le \alpha_2$, the Poisson INAR(1) process $\{X_t(\lambda_1,\alpha_1) : t\in\mathbb{N}_0\}$ is said to be slower in level-crossing in the st-sense than $\{X_t(\lambda_2,\alpha_2) : t\in\mathbb{N}_0\}$.

6. ON GOING AND FURTHER WORK

More than 50 years after Samuel Karlin's first and astounding contributions on total positivity, we illustrate how this concept and its implications provide insights on the performance of quality control charts for the mean of the Poisson INAR(1) process.

Directions for future work include trying to prove the conjecture $HT^0(\lambda,\alpha) \downarrow_{lr}$ with λ.

So far we can add that extensive numerical results, such as the ones shown in Figure 3, suggest this conjecture is valid. In this figure, we can find the likelihood ratio functions $\frac{P[HT^0((j+0.1)\,\lambda_0,\alpha_0)=m]}{P[HT^0(j\,\lambda_0,\alpha_0)=m]}$, for $j=1, 1.1, 1.2, 1.3$, when $\lambda_0=1$, $\alpha_0=0.4$ and $k=3$, as in Example 5.1. All these likelihood ratios are nonincreasing functions suggesting that

$$HT^0((j+0.1)\,\lambda_0,\alpha_0)=m] \ \le_{lr} HT^0(j\,\lambda_0,\alpha_0) \ , \quad j=1, 1.1, 1.2, 1.3 \ . \tag{6.1}$$

Interestingly, additional numerical results led to the conclusion that $HT^0(\lambda,\alpha) \not\downarrow_{lr}$ with α.

Out[162]=

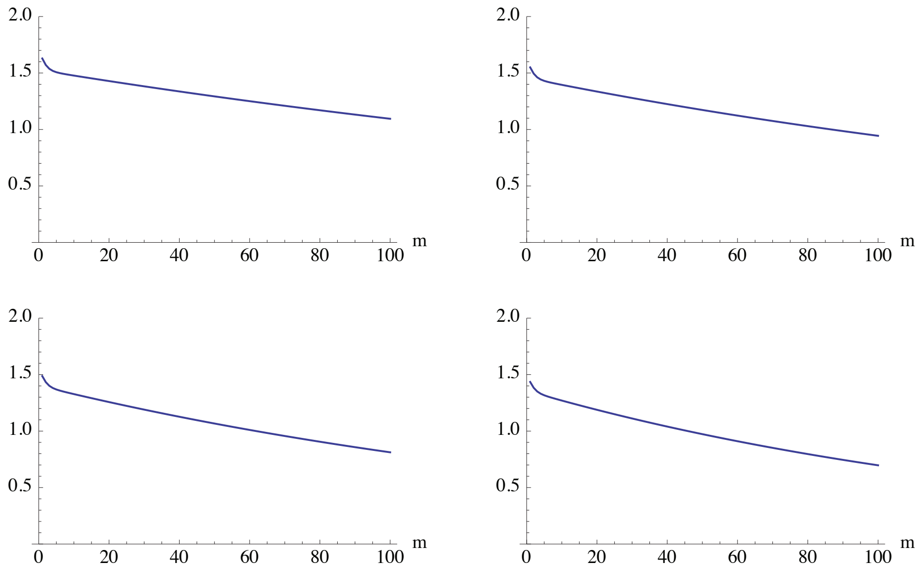

Figure 3: Likelihood ratios functions: $\frac{P[HT^0((j+0.1)\,\lambda_0,\alpha_0)=m]}{P[HT^0(j\,\lambda_0,\alpha_0)=m]}$, for $\lambda_0 = 1$, $\alpha_0 = 0.4$, $k = 3$, and $j = 1$ (top left), 1.1 (top right), 1.2 (bottom left), 1.3 (bottom right).

The simplicity of the Shewhart control charts, such as the upper one-sided c-chart we used, was responsible for their widespread popularity among quality control practitioners. The fact that Shewhart-type charts only use the last observed value of their control statistics to trigger (or not) a signal is responsible for a serious limitation: they are not effective in the detection of small and moderate shifts in the parameter being monitored. As put by Ramos (2013, p. 5), this limitation led to the cumulative sum control chart (CUSUM) proposed by Page (1954) and the exponentially weighted moving average control chart (EWMA) introduced by Roberts (1959), originally designed to monitor the process mean. CUSUM and EWMA control charts make use of recursive control statistics that account for the information contained in every collected sample of the process and prove to be more sensitive to small and moderate shifts in the process mean.

As a consequence, we also plan to conduct a similar analysis on the HT of these more sophisticated quality control charts to monitor $\frac{\lambda}{1-\alpha}$. A few difficulties may arise in the derivation of a result such as (5.20), namely because the CUSUM and EWMA control statistics constitute Markov chains with two-dimensional state spaces, as noted by Weiss and Testik (2009) and Weiss (2009c).

ACKNOWLEDGMENTS

This work was partially supported by FCT (Fundação para a Ciência e a Tecnologia) through projects PEst-OE/MAT/UI0822/2014 and PEst-OE/MAT/UI4080/2014.

REFERENCES

[1] Al-Osh, M.A. and Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process, *Journal of Time Series Analysis*, **8**, 261–275.

[2] Assaf, D.; Shaked, M. and Shanthikumar, J.G. (1985). First-passage times with PF_r densities, *Journal of Applied Probability*, **22**, 185–196.

[3] ASQ (n.d.) *The Early 20th Century*, American Society for Quality, accessed from `http://asq.org/learn-about-quality/history-of-quality/overview/20th-century.html` on 2012-07-02.

[4] Barlow, R.E. and Proschan, F. (1975/1981). *Statistical Theory of Reliability and Life Testing: Probability Models* (2nd printing), Silver Spring: Holt, Rinehart and Winston, Inc.

[5] Boucher, J.-P.; Denuit, M. and Guillén, M. (2008). Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions, *Variance*, **2**, 135–162.

[6] Brook, D. and Evans, D.A. (1972). An approach to the probability distribution of CUSUM run length, *Biometrika*, **59**, 539–549.

[7] Brown, M. and Chaganty, N.R. (1983). On the first passage time distribution for a class of Markov chains, *The Annals of Probability*, **11**, 1000–1008.

[8] Casella, G. and Berger, R.L. (2002). *Statistical Inference* (second edition), Duxbury.

[9] De Santis, E. and Spizzichino, F. (2014). Stochastic Comparisons between hitting times for Markov Chains and words' occurrences, `http://arxiv.org/abs/1210.1116v2`

[10] Ferreira, F. and Pacheco, A. (2007). Comparison of level-crossing times for Markov and semi-Markov processes, *Statistics & Probability Letters*, **77**, 151–157.

[11] Gan, F.F. (1990). Monitoring Poisson observations using modified exponentially weighted moving average control charts, *Communications in Statistics: Simulation and Computation*, **19**, 103–124.

[12] Gan, F.F. (1993). An optimal design of CUSUM control charts for binomial counts, *Journal of Applied Statistics*, **20**, 445–460.

[13] Greenberg, I. (1997). Markov chain approximation methods in a class of level-crossing problems, *Operations Research Letters*, **21**, 153–158.

[14] Irle, A. and Gani, J. (2001). The detection of words and an ordering for Markov chains, *Journal of Applied Probability*, **38A**, 66–77.

[15] Karasu, I. and Özekici, S. (1989). NBUE and NWUE properties of increasing Markov processes, *Journal of Applied Probability*, **27**, 827–834.

[16] Karlin, S. (1964). Total positivity, absorption probabilities and applications, *Transactions of the American Mathematical Society*, **11**, 33–107.

[17] Karlin, S. (1968). *Total Positivity – Vol. I*, California: Stanford University Press.

[18] Karlin, S. and Proschan, F. (1960). Pólya type distributions of convolutions, *Annals of Mathematical Statistics*, **31**, 721–736.

[19] Kijima, M. (1997). *Markov Processes for Stochastic Modelling*, London: Chapman and Hall.

[20] Kim, J.S. and Proschan, F. (1982). *Total positivity*. In "Encyclopedia of Statistical Sciences" (S. Kotz, N.L. Johnson, N. Lloyd and C.B. Read, Eds.), Volume 7, 289–297, New York: John Wiley & Sons.

[21] Li, H. and Shaked, M. (1995). On the first passage times for Markov processes with monotone convex transition kernels, *Stochastic Processes and their Applications*, **58**, 205–216.

[22] Li, H. and Shaked, M. (1997). Ageing first-passage times of Markov processes: a matrix approach, *Journal of Applied Probability*, **34**, 1–13.

[23] Margavio, T.M.; Conerly, M.D.; Woodall, W.H. and Drake, L.G. (1995). Alarm rates for quality control charts, *Statistics & Probability Letters*, **24**, 219–224.

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

[25] Montgomery, D.C. (2009). *Introduction to Statistical Quality Control* (6th. edition), New York: John Wiley & Sons.

[26] Morais, M.J.C. (2002). *Stochastic Ordering in the Performance Analysis of Quality Control Schemes*, PhD Thesis, Technical University of Lisbon, Portugal.

[27] Morais, M.C. and Pacheco, A. (2012). A note on the aging properties of the run length of Markov-type control charts, *Sequential Analysis*, **31**, 88–98.

[28] Nelson, L.S. (1982). *Control charts*. In "Encyclopedia of Statistical Sciences" (S. Kotz and N.L. Johnson, Eds.), Volume 2, pp. 176–183, John Wiley & Sons.

[29] Page, E.S. (1954). Continuous inspection schemes, *Biometrika*, **41**, 100–115.

[30] Ramachandran, K.V. (1958). A test of variances, *Journal of the American Statistical Society*, **53**, 741–747.

[31] Ramos, P.A.A.C.F. (2013). Performance analysis of simultaneous control schemes for the process mean (vector) and (co)variance (matrix), PhD Thesis, Technical University of Lisbon, Portugal.

[32] Roberts, S.W. (1959). Control charts tests based on geometric moving averages, *Technometrics*, **1**, 239–250.

[33] Ross, S.M. (1983). *Stochastic Processes*, New York: John Wiley & Sons.

[34] SHAKED, M. and LI, H. (1997). Aging first-passage times. In "Encyclopedia of Statistical Sciences: Update Volume 1" (S. Kotz, C.B. Read and D.L. Banks, Eds.), pp. 11–20, John Wiley & Sons.

[35] SHAKED, M. and SHANTHIKUMAR, J.G. (1994). *Stochastic Orders and Their Applications*, London: Academic Press.

[36] SHAKED, M. and SHANTHIKUMAR, J.G. (2007). *Stochastic Orders*, New York: Springer-Verlag.

[37] SILVA, I. (2005). *Contributions to the Analysis of Discrete-valued Time Series*, PhD Thesis, Faculdade de Ciências da Universidade do Porto, Portugal.

[38] SILVA, N.; PEREIRA, I. and SILVA, M.E. (2009). Forecasting in INAR(1) model, *REVSTAT – Statistical Journal*, **7**, 119–134.

[39] STEUTEL, F.W. and VAN HARN, K. (1979). Discrete analogues of self-decomposability and stability, *The Annals of Probability*, **7**, 893–899.

[40] WEISS, C.H. (2007). Controlling correlated processes of Poisson counts, *Quality Reliability Engineering International*, **23**, 741–754.

[41] WEISS, C.H. (2009a). Monitoring correlated processes with binomial marginals, *Journal of Applied Statistics*, **36**, 399–414.

[42] WEISS, C.H. (2009b). Controlling jumps in correlated processes of Poisson counts, *Applied Stochastic Models in Business and Industry*, **25**, 551–564.

[43] WEISS, C.H. (2009c). EWMA Monitoring of correlated processes of Poisson counts, *Quality Technology and Quantitative Management*, **6**, 137–153.

[44] WEISS, C.H. (2009d). *Categorical Time Series Analysis and Applications in Statistical Quality Control*, PhD Thesis, Fakultät für Mathematik und Informatik der Universität Würzburg, dissertation.de — Verlag im Internet GmbH.

[45] WEISS, C.H. and TESTIK, M.C. (2009). CUSUM monitoring of first-order integer-valued autoregressive processes of Poisson counts, *Journal of Quality Technology*, **41**, 389–400.

[46] ZEGER, S.L. (1988). A regression model for time series of counts, *Biometrika*, **75**, 621–629.