Detecting Mean Increases in Poisson INAR(1) Processes with EWMA Control Charts

Christian H. Weiß
Department of Mathematics
Darmstadt University of Technology
For references in this talk, see


**Weiß (2010)**: Detecting mean increases in Poisson INAR(1) processes with EWMA control charts. Appears in *Journal of Applied Statistics*.

Poisson INAR(1) Processes

Definition & Properties
Definition of Poisson INAR(1) process:

Let \((\epsilon_t)_N\) be i.i.d. process with marginal distribution \(Po(\mu(1-\alpha))\), where \(\mu > 0\) and \(\alpha \in (0; 1)\). Let \(N_0 \sim Po(\mu)\).

If the process \((N_t)_{N_0}\) satisfies

\[
N_t = \alpha \circ N_{t-1} + \epsilon_t, \quad t \geq 1,
\]

plus sufficient independence conditions, then it follows a stationary \textit{Poisson INAR(1) model} with marginal distribution \(Po(\mu)\).

Binomial thinning, due to Steutel & van Harn (1979):

$N$ discrete random variable with range $\{0, \ldots, n\}$ or $\mathbb{N}_0$.

**Binomial thinning**

$$\alpha \circ N := \sum_{i=1}^{N} X_i,$$

where $X_i$ are independent Bernoulli trials $\sim B(1, \alpha)$.

Guarantees that right-hand side always integer-valued:

$$N_t = \alpha \circ N_{t-1} + \epsilon_t.$$

**Interpretation:** $\alpha \circ N$ is number of survivors.
Poisson INAR(1) Processes

Basic properties of Poisson INAR(1) processes:

- Stationary Markov chain with $Po(\mu)$-marginals and

$$p_{k|l} := P(N_t = k \mid N_{t-1} = l) = \frac{\sum_{j=0}^{\min(k,l)} \binom{l}{j} \alpha^j (1 - \alpha)^{l-j} \cdot e^{-\mu(1-\alpha)} \frac{(\mu(1-\alpha))^{k-j}}{(k-j)!}}{e^{-\mu(1-\alpha)}} ,$$

- autocorrelation $\rho(k) := Corr[N_t, N_{t-k}] = \alpha^k$.

Estimation from time series $N_1, \ldots, N_T$:

$$\hat{\mu} := \frac{1}{T} \cdot \sum_{t=1}^{T} N_t, \quad \hat{\alpha} = \frac{\sum_{t=2}^{T} (N_t - \bar{N}_T)(N_{t-1} - \bar{N}_T)}{\sum_{t=1}^{T} (N_t - \bar{N}_T)^2}.$$
Interpretation of INAR(1) process:

\[ N_t = \alpha \circ N_{t-1} + \epsilon_t \]

Population at time \( t \) = Survivors of time \( t-1 \) + Immigration

Interpretation applies well to many real-world problems, e.g.:

- \( N_t \): number of users accessing web server, \( \epsilon_t \): number of new users, \( \alpha \circ N_{t-1} \): number of previous users still active.
- \( N_t \): number of faults, \( \epsilon_t \): number of new faults, \( \alpha \circ N_{t-1} \): number of previous faults not rectified yet.

Christian H. Weiß — Darmstadt University of Technology
Poisson INAR(1) Processes

The Poisson INAR(1) model . . .

- is of simple structure,
- essential properties known explicitly,
- is easy to fit to data,
- is easy to interpret,
- applies well to real-world problems, . . .

**In a nutshell:** A simple model for autocorrelated counts, which is well-suited for SPC!

Christian H. Weiß — Darmstadt University of Technology
Controlling Poisson INAR(1) Processes

Control Concepts
Poisson INAR(1) model:

\((N_t)_{N_0}\) is stationary Poisson INAR(1) process with innovations \((\epsilon_t)_N \sim Po(\mu(1 - \alpha))\). So \(N_t \sim Po(\mu)\).

State of statistical control: \(\mu = \mu_0\) and \(\alpha = \alpha_0\).
Weiß (2007) analyzed \textit{c-Chart} for Poisson INAR(1):

+ exact \textit{ARLs} via Markov chain approach,

+ easily designed and interpreted,

+ effective for very large shifts,

− but very insensitive otherwise!
Weiß (2009) analyzed **Combined EWMA Chart**:

+ exact *ARLs* via Markov chain approach,

+ applicable for very different types of out-of-control situation,

− difficult to design, six design parameters!

Christian H. Weiß — Darmstadt University of Technology
In practice:
Often only interested in detecting increases in process mean compared to in-control mean, e.g.,

- counts of defects in manufacturing industry,
- counts of complaints in service industry,
- number of certain infections in epidemiology.

In such situations, combined EWMA chart appears to be overparametrized.

Christian H. Weiß — Darmstadt University of Technology
Controlling INAR(1) Processes – Concepts

→ One-sided CUSUM chart of Weiß & Testik (2009) attractive alternative:

+ exact ARLs via Markov chain approach,
+ easily designed (three design parameters),
+ very sensitive already to small mean shifts, etc.

⇒ benchmark chart in the following!

Simplified EWMA possible with similar performance?

Christian H. Weiß — Darmstadt University of Technology
One-Sided Poisson INAR(1) EWMA Chart

Definition & Properties
One-sided **EWMA chart** for detecting positive shifts in $\mu$:

$$Q_0 = q_0,$$

$$Q_t = \text{round}(\lambda \cdot N_t + (1 - \lambda) \cdot Q_{t-1}), \quad t = 1, 2, \ldots$$

$q_0 \geq 0$: starting value, typically $q_0 = 0$.

Fast Initial Response (FIR) feature if $q_0 > 0$.

$\lambda \in (0; 1]$: smoothing parameter.

$u > 0$: upper control limit.

$(N_t)_N$ considered in control unless alarm $Q_t \geq u$ triggered.
Properties: One-sided EWMA chart coincides with combined EWMA chart of Weiß (2009) if design parameters $l_c = l_e = 0$ and $u_c = \infty$.

⇒ All results of Weiß (2009) directly apply to one-sided EWMA chart.

In particular, $ARL$s can be computed with MC approach outlined there.
**Real-data example** of Weiß & Testik (2009):
Time series of counts of IP addresses, conjectured to stem from following in-control model:
Poisson INAR(1) with $\mu_0 = 1.28$ and $\alpha_0 = 0.29$.

Based on this model, Weiß & Testik (2009) designed a $c$ chart with UCL $u = 6$ and $ARL_0 \approx 504.949$, and four CUSUM charts with $ARL_0$s between 502.586 and 507.447.

**Aim:** Find EWMA chart design with $ARL_0$ around 500.
Considered control charts:

- \((u, \lambda, q_0) = (2, 0.11, 1)\) with \(ARL_0 \approx 504.949\);
- \((u, \lambda, q_0) = (3, 0.16, 2)\) with \(ARL_0 \approx 504.949\);
- \((u, \lambda, q_0) = (4, 0.37, 3)\) with \(ARL_0 \approx 592.584\);
- \((u, \lambda, q_0) = (5, 0.63, 1)\) with \(ARL_0 \approx 464.239\).

We expect design 4 to show too much false alarms, while design 3 too robust.

Designs 1 and 2 exactly same \(ARL_0\) as \(c\) chart with UCL 6!
In fact, it follows from definition of one-sided EWMA for designs 1 and 2 that statistic $Q_t$ reaches its respective limit $u$ for first time iff $N_t = 6$ for first time.

⇒ Both charts lead to equivalent decision rule as $c$ chart.

⇒ Both show same (bad) performance as $c$ chart.
Due to small value of $\lambda$ and rounding operation of $Q_t$, data clearly oversmoothed.

Deficiencies of one-sided EWMA chart:
On one hand, sensitivity becomes better if $\lambda$ decreases.
On other hand, effect of smoothing increases for decreasing $\lambda$.

⇒ Not possible to choose $\lambda$ as small as perhaps required to reach certain sensitivity.
The $s$-EWMA Control Chart

Definition & Properties
Basic idea: New rounding operation!

Let $s \in \mathbb{N}$, define $\mathbb{Q}_s := \left\{ \frac{r}{s} \mid r \in \mathbb{Z} \right\}$ as set of all rationals with denominator $s$. (Note: $\mathbb{Q}_1 = \mathbb{Z}$)

Define function $s$-round: $\mathbb{R} \rightarrow \mathbb{Q}_s$ by

$$s\text{-round}(x) = z \text{ iff } x \in \left[ z - \frac{1}{2s}; z + \frac{1}{2s} \right].$$

So $s$-round maps $x$ onto nearest fraction with denominator $s$.

Note that $1$-round coincides with round.

Christian H. Weiß — Darmstadt University of Technology
One-sided \textbf{s-EWMA chart} for detecting positive shifts in $\mu$:

\begin{align*}
Q_0 &= q_0, \\
Q_t &= s\text{-round}\left(\lambda \cdot N_t + (1 - \lambda) \cdot Q_{t-1}\right), \quad t = 1, 2, \ldots
\end{align*}

(N_t)_N considered in control unless alarm $Q_t \geq u$ triggered.

$u \in \mathbb{Q}_s^+$: upper control limit.

$q_0 \in \{0, \ldots, u - \frac{1}{s}\}$: starting value, typically $q_0 = 0$.

Fast Initial Response (FIR) feature if $q_0 > 0$.

$\lambda \in (0; 1]$: smoothing parameter.
Weiß (2010) shows that still \((N_t, Q_t)_N\) is homogeneous bivariate Markov chain.

\[\Rightarrow\] Again MC approach for exact ARL computation.

However: Dimension of involved matrices increases with \(s\).

\[\Rightarrow\] Values like \(s \in \{1, 2, 4\}\) reasonable for practice.
Design of One-Sided $s$-EWMA Chart:

1. Find best possible 1-EWMA design $(u_1, \lambda_1)$ according to Weiß (2009),

2. increase $s$ and choose $u \leq u_1$ such that in-control $ARL$ below desired $ARL_0$, 

3. decrease $\lambda \in (0; \lambda_1]$ to adjust in-control $ARL$ close to desired $ARL_0$, 

4. use $q_0 \in \{0,\ldots,u - \frac{1}{s}\}$ for fine-tuning of in-control $ARL$. 

Christian H. Weiß — Darmstadt University of Technology
Above real-data example: We find

- 2-EWMA chart with \((u, \lambda, q_0) = (\frac{7}{2}, 0.295, 3)\) and \(ARL_0 \approx 518.459\),

- 4-EWMA chart with \((u, \lambda, q_0) = (\frac{14}{4}, 0.323, 3)\) and \(ARL_0 \approx 505.301\).

Both charts have similar in-control performance like considered \(c\) chart and CUSUM charts.
Out-of-control performance of two $s$-EWMA charts $c$ chart and two most sensitive CUSUM charts

\[ ARL \]

\[ c: (l,u)=(0,6) \]

\[ (h,k,c_0): \]
- $(26/4, 9/4, 0)$
- $(27/4, 9/4, 21/4)$

\[ (s,u,\lambda,q_0): \]
- $(2, 7/2, 0.295, 3)$
- $(4, 14/4, 0.323, 3)$

Christian H. Weiß — Darmstadt University of Technology
Both EWMA charts noticeably better than $c$ chart, but outperformed by CUSUM chart with FIR.

4-EWMA chart applied to IP data detects the out-of-control situation already at time $t = 10$, like FIR-CUSUM:

![Graph showing EWMA control chart example]

Christian H. Weiß — Darmstadt University of Technology
Extensive performance study in Weiß (2010):

- $s$-EWMA chart clearly outperforms $c$ chart,

- best sensitivity for small $\lambda$ although very large shifts often better detected for large $\lambda$,

- choice of $q_0 > \mu_0$ (FIR) slightly improves sensitivity concerning very large shifts ($\geq 100\%$),

- but CUSUM shows better performance for small shifts (5\% to 20\%); performance of $s$-EWMA relative to CUSUM better for increasing $\alpha_0$.  

Christian H. Weiß — Darmstadt University of Technology
In Weiß (2010), also robustness against model misspecification analyzed:

$s$-EWMA chart based on assumption of INAR(1) process with Poisson marginals.

Moderate departures may not be detected.

Most common misspecification: overdispersion.

Most popular approaches for modeling overdispersion: negative binomial and generalized Poisson distribution.

Both are possible marginal distributions of INAR(1) model.
In Weiß (2010), generalized Poisson considered.

$s$-EWMA chart affected by overdispersion: in-control performance influenced most severely, ooc performance for large shifts in $\mu$ nearly constant.

Robustness becomes better if $\lambda$ decreased, while additional FIR feature does not affect robustness.

For very small $\lambda$ ($\lambda \approx 0.1$), $ARL$ performance reasonably close to expected one even if up to 20% overdispersion.

Christian H. Weiß — Darmstadt University of Technology
Future Research

Work in progress together with Murat Testik:  
**Robustness of one-sided CUSUM.**

It seems that

- one-sided CUSUM quite sensitive to overdispersion, better robustness for large $h$, small $k$, without FIR;

- but a new **Winsorized CUSUM**
  - allows approximate $ARL$ comp. with MC approach,
  - has nearly unaffected ooc performance
    (Wins. CUSUM *with FIR* even improved ooc perf.),
  - shows a clearly improved robustness.

Christian H. Weiß — Darmstadt University of Technology
Thank You
for Your Interest!

Christian H. Weiß
Department of Mathematics
Darmstadt University of Technology