Process Capability Analysis for Serially Dependent Processes of Poisson Counts

Christian H. Weiß

Department of Mathematics,

Darmstadt University of Technology

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Literature

This talk is based on the paper

Weiß (2010): Process Capability Analysis for Serially Dependent Processes of Poisson Counts. Submitted.

All references mentioned in this talk correspond to the references in this article.

Process Capability Analysis for Poisson Counts

Brief review

Process is **in control** (stable)

if stationary, following particular process model.

Even if production process in control:

Is quality of produced items sufficient for practice?

$\Rightarrow$ Evaluate to what extent the given (external!) target values and specification limits are met

$\Rightarrow$ **process capability analysis** (PC analysis).

Process capability indices (PC indices):

Evaluate actual process capability in a single number.

Huge amount of work concerning
continuous variates (variables data processes),

but only few articles consider
PC analysis of discrete variates (attributes data),
like count data processes.

Examples of **count data processes**:

- Manufacturing industry:
  number of defects or nonconformities.
- Service industry:
  number of complaints of customers per time unit.

Essentially, only one approach in literature for

PC analysis of attributes data:

Target value: Acceptable level of probability

for producing non-conforming/defective item.

In the following, w.l.o.g.: 0.0027.

We measure the **true level** of probability

for producing non-conforming/defective item.

PC analysis compares acceptable level with true level.

Particular case of **Poisson counts:**

– upper specification limit USL (e. g., max. number of non-conformities per produced item, otherwise: item defective),

– acceptable probability level for defective item: 0.0027.

Two proposals in literature:

Perakis & Xekalaki (2005):

$$C_{\mathsf{PX}} \;:=\; \frac{0.0027}{P(X>USL)} \;\in [0.0027;\infty).$$

Borges & Ho (2001):

$$C_{\mathsf{BH}} \;:=\; \frac{1}{3}\cdot\Phi^{-1}\Big(1-\frac{1}{2}\cdot P(X>USL)\Big) \;\in [0;\infty).$$

Perakis & Xekalaki (2005): $C_{\mathsf{PX}} := \dfrac{0.0027}{P(X > USL)}.$

Borges & Ho (2001): $C_{\mathsf{BH}} := \dfrac{1}{3}\cdot\Phi^{-1}\Big(1-\dfrac{1}{2}\cdot P(X > USL)\Big).$

We have:

$P(X > USL) = 0.0027 \Rightarrow C_{\mathsf{PX}} = C_{\mathsf{BH}} = 1.$

$P(X > USL) < 0.0027$ (desirable!) $\Rightarrow C_{\mathsf{PX}}, C_{\mathsf{BH}} > 1.$

C_{BH} scaled like usual C_p index,

C_{PX} easy to interpret as a fraction of two probabilities.

In practice: indices $C_{\mathsf{PX}}, C_{\mathsf{BH}}$ not known, but have to be estimated based on available time series data.

Perakis & Xekalaki (2005):
$(X_t)_{\mathbb{N}}$ i.i.d. with Poisson marginal distribution $Po(\mu)$.

Recommendation:

1. Estimate μ by sample mean $\hat{\mu} := \bar{X}$, then
2. estimate $P(X > USL)$ by
   Poisson probability $P(X > USL \mid \mu = \hat{\mu})$.

Estimates $\hat{C}_{\mathsf{PX}}, \hat{C}_{\mathsf{BH}}$ obtained by inserting this probability into formulae for $C_{\mathsf{PX}}, C_{\mathsf{BH}}$.

Questions:

How do estimators $\widehat{C}_{\mathsf{PX}}, \widehat{C}_{\mathsf{BH}}$ perform?

$\rightarrow$ Perakis & Xekalaki (2005): $\widehat{C}_{\mathsf{PX}}$ for **i.i.d.** counts.

But: i.i.d.-assumption often violated in practice!

$\Rightarrow$ Performance in presence of serial dependence?

How to define and estimate such indices at all for serially dependent Poisson counts?

Poisson INAR(1) Model

Definition & Properties

Definition of Poisson INAR(1) process:

Let $\lambda > 0$ and $\alpha \in (0;1)$.

Innovations process (usually unobservable):

$(\epsilon_t)_{\mathbb{N}}$ i.i.d. with marginal distribution $Po(\lambda)$.

Observations process $(X_t)_{\mathbb{N}_0}$: $\quad X_0 \sim Po(\frac{\lambda}{1-\alpha})$,

$$X_t \;=\; \alpha \circ X_{t-1} \;+\; \epsilon_t, \qquad t \geq 1,$$

plus sufficient independence conditions.

McKenzie (1985), Al-Osh & Alzaid (1987, 1988)

Poisson INAR(1) Processes

Basic properties of Poisson INAR(1) processes:

- Stationary Markov chain with transition probabilities
  $p_{k|l} \;:=\; P(X_t = k \mid X_{t-1} = l) \;=\; \sum_{j=0}^{\min(k,l)} \binom{l}{j}\alpha^j(1-\alpha)^{l-j} \cdot e^{-\lambda}\frac{\lambda^{k-j}}{(k-j)!}$,

- Poisson marginal distribution $Po(\mu)$ with mean $\mu = \frac{\lambda}{1-\alpha}$,

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

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

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

Binomial thinning

$$\alpha \circ X \; := \; \sum_{i=1}^{X} Y_i,$$

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

Guarantees that right-hand side always integer-valued:

$$X_t \; = \; \alpha \circ X_{t-1} \; + \; \epsilon_t.$$

Interpretation: $\alpha \circ X$ is number of survivors.

Interpretation of INAR(1) process:

$$\underbrace{X_t}_{\text{Population at time } t} = \underbrace{\alpha \circ X_{t-1}}_{\text{Survivors of time } t-1} + \underbrace{\epsilon_t.}_{\text{Immigration}}$$

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

- X_t: number of faults, ϵ_t: number of new faults, $\alpha \circ X_{t-1}$: number of previous faults not rectified yet.
- X_t: number of unanswered complaints of customers, ϵ_t new complaints, $\alpha \circ X_{t-1}$ past complaints.

Observations $(X_t)_{\mathbb{N}_0}$: marg. dist. $Po(\mu)$ with $\mu = \frac{\lambda}{1-\alpha}$,

Innovations $(\epsilon_t)_{\mathbb{N}}$: marg. dist. $Po(\lambda)$.

We observe $X_1, \ldots, X_T$ from a Poisson INAR(1) process.

If **observations** X_t themselves are quantity of interest:

Estimate observations' mean μ,

then process capability in terms of $C_{\mathsf{PX},X}, C_{\mathsf{BH},X}$.

If **innovations** ϵ_t are quantity of interest:

Estimate innovations' mean λ, (from $X_1, \ldots, X_T$!)

then process capability in terms of $C_{\mathsf{PX},\epsilon}, C_{\mathsf{BH},\epsilon}$.

Example,

where **innovations** ϵ_t may be quantity of interest:

We observe the total **number of faults** in system: X_t.

ϵ_t: number of new faults,

$\alpha \circ X_{t-1}$: number of previous faults not rectified yet.

Then, perhaps,

not total number of unremedied faults most important,

but number of new faults may better describe capability of production process.

Poisson INAR(1) Processes

Observations $(X_t)_{\mathbb{N}_0}$: marg. dist. $Po(\mu)$ with $\mu = \frac{\lambda}{1-\alpha}$,

Innovations $(\epsilon_t)_{\mathbb{N}}$: marg. dist. $Po(\lambda)$.

We observe $X_1, \ldots, X_T$ from a Poisson INAR(1) process.

We consider **two questions** in the following:

How to estimate observations' PC in terms of $C_{\mathsf{PX},X}, C_{\mathsf{BH},X}$, and how do estimators perform?

How to estimate innovations' PC in terms of $C_{\mathsf{PX},\epsilon}, C_{\mathsf{BH},\epsilon}$, and how do estimators perform?

Process Capability for Observations of Poisson INAR(1)

Estimation

Observations $(X_t)_{\mathbb{N}_0}$: marg. dist. $Po(\mu)$ with $\mu = \frac{\lambda}{1-\alpha}$.

PC Indices:

$$C_{\mathrm{PX},X} \;:=\; \frac{0.0027}{P(X > USL)},$$

$$C_{\mathrm{BH},X} \;:=\; \frac{1}{3} \cdot \Phi^{-1}\big(1 - \frac{1}{2} \cdot P(X > USL)\big).$$

Tasks:

1. Estimate observations' mean μ: $\widehat{\mu}$.
2. Estimate $P(X > USL)$ by

   Poisson probability $P(X > USL \mid \mu = \widehat{\mu})$.

PC Analysis for Observations

Task 1: estimation of observations' mean μ.

Moment estimator: $\hat{\mu}_{\mathsf{MM}} = \bar{X}_T = \frac{1}{T} \cdot \sum_{t=1}^{T} X_t.$

Same estimator as in i.i.d. case,

but different asymptotic distribution:

(Freeland & McCabe, 2005)

$$\sqrt{T}(\hat{\mu}_{\mathsf{MM}} - \mu) \underset{a}{\sim} N\Big(0,\ \mu \cdot \tfrac{1+\alpha}{1-\alpha}\Big).$$

Interval estimation for μ:

one-sided confidence intervals $(0; u)$ particularly important,

since in practice, **worst-case scenarios** most relevant.

Task 1: (continued)

Point estimator: $\hat{\mu}_{\mathsf{MM}} = \bar{X}_T = \frac{1}{T} \cdot \sum_{t=1}^{T} X_t.$

Approximate **confidence interval** for μ on level γ:

$(0; \hat{u}_{\mu,\gamma})$ with

$$\hat{u}_{\mu,\gamma} := \hat{\mu}_{\mathsf{MM}} + \frac{z_\gamma^2}{2T} \cdot \frac{1+\hat{\alpha}_{\mathsf{MM}}}{1-\hat{\alpha}_{\mathsf{MM}}}$$

$$+ \frac{z_\gamma}{\sqrt{T}} \cdot \sqrt{\hat{\mu}_{\mathsf{MM}} \cdot \frac{1+\hat{\alpha}_{\mathsf{MM}}}{1-\hat{\alpha}_{\mathsf{MM}}} + \frac{z_\gamma^2}{4T} \cdot \left(\frac{1+\hat{\alpha}_{\mathsf{MM}}}{1-\hat{\alpha}_{\mathsf{MM}}}\right)^2},$$

where $z_\gamma = \Phi^{-1}(\gamma)$ is γ-quantile of $N(0,1)$-distribution.

Task 1: (continued)

Finite-sample performance of $\widehat{\mu}_{\mathsf{MM}}$ and $(0;\widehat{u}_{\mu,\gamma})$:

- Asymptotic variance $\frac{\mu}{T}\cdot\frac{1+\alpha}{1-\alpha}$ for $\widehat{\mu}_{\mathsf{MM}}$
  very good approximation already for $T\geq 25$;

- asymptotic distribution $N\big(\mu,\ \frac{\mu}{T}\cdot\frac{1+\alpha}{1-\alpha}\big)$ for $\widehat{\mu}_{\mathsf{MM}}$
  very good approximation for $T\geq 100$;

- true coverage probability of $(0;\widehat{u}_{\mu,\gamma})$ close to nominal level γ for $T\geq 100$.

Task 2: estimation of observations' capability.

Given: estimators $\widehat{\mu}_{\mathsf{MM}}$ and $(0; \widehat{u}_{\mu,\gamma})$.

Point estimates of $C_{\mathsf{PX},X}$ and $C_{\mathsf{BH},X}$:

$$\widehat{C}_{\mathsf{PX},X} \;:=\; 0.0027 \big/ P(X > USL \mid \mu = \widehat{\mu}),$$

$$\widehat{C}_{\mathsf{BH},X} \;:=\; \tfrac{1}{3} \cdot \Phi^{-1}\Big(1 - \tfrac{1}{2} \cdot P(X > USL \mid \mu = \widehat{\mu})\Big).$$

Interval estimates of $C_{\mathsf{PX},X}$ and $C_{\mathsf{BH},X}$:

$$\widehat{I}_{C_{\mathsf{PX},X}} \;:=\; \Big(0.0027 \big/ P(X > USL \mid \mu = \widehat{u}_{\mu,\gamma});\; \infty\Big),$$

$$\widehat{I}_{C_{\mathsf{BH},X}} \;:=\; \Big(\tfrac{1}{3} \cdot \Phi^{-1}\Big(1 - \tfrac{1}{2} \cdot P(X > USL \mid \mu = \widehat{u}_{\mu,\gamma})\Big);\; \infty\Big).$$

Task 2: (continued)

Finite-sample performance: Good coverage for $T \geq 100$,

but $(USL = 11)$

		$C_{\mathsf{PX},X}$	mean of $\hat{C}_{\mathsf{PX},X}$			$C_{\mathsf{BH},X}$	mean of $\hat{C}_{\mathsf{BH},X}$		
μ	$\alpha \setminus T$		25	100	400		25	100	400
3	0	37.82	81.92	44.93	39.44	1.324	1.328	1.325	1.324
	0.25		143.79	50.57	40.47		1.330	1.326	1.324
	0.5		628.01	64.28	42.91		1.334	1.327	1.325
4	0	2.950	4.731	3.285	3.027	1.105	1.108	1.106	1.105
	0.25		6.540	3.519	3.074		1.110	1.106	1.105
	0.5		14.935	4.088	3.170		1.114	1.107	1.105
4.57	0	1.000	1.435	1.087	1.021	1.000	1.003	1.001	1.000
	0.25		1.832	1.146	1.031		1.005	1.001	1.000
	0.5		3.268	1.279	1.057		1.007	1.001	1.000

Task 2: (continued)

Huge bias of $\widehat{C}_{\mathsf{PX},X}$

also clear from large-sample approximation:

Applying Delta theorem with Taylor expansion up to order 2,

we obtain bias approximations ($USL = 11$, $\mu = 3$)

$\frac{618.7}{T} \cdot \frac{1+\alpha}{1-\alpha}$ for $\widehat{C}_{\mathsf{PX},X}$, $\quad$ $\frac{0.096}{T} \cdot \frac{1+\alpha}{1-\alpha}$ for $\widehat{C}_{\mathsf{BH},X}$.

$\Rightarrow C_{\mathsf{PX},X}$ easy to interpret (quotient between nominal and actual defect probability), but

scaled in uncommon way and difficult to estimate

$\Rightarrow C_{\mathsf{BH},X}$ better choice for PC analysis of attributes data.

Process Capability for Innovations of Poisson INAR(1)

Estimation

Innovations $(\epsilon_t)_{\mathbb{N}}$: marg. dist. $Po(\lambda)$.

PC Indices:

$$C_{\mathsf{PX},\epsilon} \;:=\; \frac{0.0027}{P(\epsilon > USL)},$$

$$C_{\mathsf{BH},\epsilon} \;:=\; \frac{1}{3}\cdot \Phi^{-1}\Big(1-\frac{1}{2}\cdot P(\epsilon > USL)\Big).$$

Tasks:

1. Estimate innovations' mean λ: $\hat{\lambda}$.
2. Estimate $P(\epsilon > USL)$ by

   Poisson probability $P(\epsilon > USL \mid \lambda = \hat{\lambda})$.

Task 1: estimation of innovations' mean λ.

Moment estimator: $\hat{\lambda}_{\mathsf{MM}} = \bar{X}_T \cdot (1 - \hat{\alpha}_{\mathsf{MM}})$

where $\hat{\alpha}_{\mathsf{MM}} := \hat{\rho}(1)$, with asymptotic distribution

(Freeland & McCabe, 2005)

$$\sqrt{T}(\hat{\lambda}_{\mathsf{MM}} - \lambda) \underset{a}{\sim} N\Big(0,\ \lambda(1 + \lambda\tfrac{1+\alpha}{1-\alpha})\Big).$$

Disadvantages:

generally biased (also see simulation study below!),

no exact properties known.

Task 1: (continued)

New **jumps estimator:** $\hat{\lambda}_{\mathsf{J}} := \frac{1}{2(T-1)} \cdot \sum_{t=2}^{T} (X_t - X_{t-1})^2$

with asymptotic distribution

$$\sqrt{T-1}(\hat{\lambda}_{\mathsf{J}} - \lambda) \underset{a}{\sim} N\Big(0,\ \lambda(1+\lambda\tfrac{3+\alpha}{1+\alpha})\Big).$$

Advantages:

exactly unbiased, expression for exact variance available.

More details and proofs in Weiß (2010).

Task 1: (continued)

Point estimator: $\hat{\lambda}_{\mathsf{MM}} = \bar{X}_T \cdot (1 - \hat{\alpha}_{\mathsf{MM}})$.

Approximate **confidence interval** for λ on level γ:

$(0; \hat{u}_{\lambda,\gamma})$ with

$$\hat{u}_{\lambda,\gamma} := \hat{\lambda}_{\mathsf{MM}} + \frac{z_\gamma}{\sqrt{T}} \cdot \sqrt{\hat{\lambda}_{\mathsf{MM}}\left(1 + \hat{\lambda}_{\mathsf{MM}} \frac{1+\hat{\alpha}_{\mathsf{MM}}}{1-\hat{\alpha}_{\mathsf{MM}}}\right)},$$

where $z_\gamma = \Phi^{-1}(\gamma)$ is γ-quantile of $N(0,1)$-distribution.

Task 1: (continued)

Point estimator: $\hat{\lambda}_{\mathsf{J}} := \frac{1}{2(T-1)} \cdot \sum_{t=2}^{T} (X_t - X_{t-1})^2.$

Approximate **confidence interval** for λ on level γ:

$(0; \hat{u}_{\lambda,\gamma})$ with

$$\hat{u}_{\lambda,\gamma} := (1 - \frac{z_\gamma^2}{T-1} \cdot \frac{3+\hat{\alpha}_{\mathsf{MM}}}{1+\hat{\alpha}_{\mathsf{MM}}})^{-1} \cdot \Big(\hat{\lambda}_{\mathsf{J}} + \frac{z_\gamma^2}{2(T-1)}$$

$$+ \frac{z_\gamma}{\sqrt{T-1}} \cdot \sqrt{\hat{\lambda}_{\mathsf{J}}(1 + \hat{\lambda}_{\mathsf{J}} \frac{3+\hat{\alpha}_{\mathsf{MM}}}{1+\hat{\alpha}_{\mathsf{MM}}}) + \frac{z_\gamma^2}{4(T-1)}}\Big),$$

where $z_\gamma = \Phi^{-1}(\gamma)$ is γ-quantile of $N(0,1)$-distribution.

Task 1: (continued)

Finite-sample performance of $\hat{\lambda}_{\mathsf{MM}}, \hat{\lambda}_{\mathsf{J}}$ and resulting $(0; \hat{u}_{\lambda,\gamma})$:

- $\hat{\lambda}_{\mathsf{MM}}$ shows a large bias, variance and skewness for small T and especially for large α;
- also asymptotic variance of $\hat{\lambda}_{\mathsf{MM}}$ is much smaller than empirically observed variance in these cases;
- for small α, $\hat{\lambda}_{\mathsf{MM}}$ has less variance and skewness than $\hat{\lambda}_{\mathsf{J}}$, also clear from asymptotic variances:

$$\tfrac{3+\alpha}{1+\alpha} < \tfrac{1+\alpha}{1-\alpha} \quad \text{iff} \quad \alpha > \sqrt{2} - 1 \approx 0.4142;$$

Task 1: (continued)

Finite-sample performance of $\hat{\lambda}_{\mathsf{MM}}, \hat{\lambda}_{\mathsf{J}}$ and resulting $(0; \hat{u}_{\lambda,\gamma})$:

- for $T \leq 50$, all intervals $(0; \hat{u}_{\lambda,\gamma})$ very conservative;
- $\hat{\lambda}_{\mathsf{MM}}$-interval conservative even for $T = 200$ if α large;
- $\hat{\lambda}_{\mathsf{J}}$-interval only moderately conservative for $T \geq 100$, independent of α.

Summary:

$\hat{\lambda}_{\mathsf{J}}$ (and corresponding interval) in general preferable,
$\hat{\lambda}_{\mathsf{MM}}$ only good for small α.

Task 2: estimation of innovations' capability.

Given: estimators $\widehat{\lambda}_{\mathsf{MM}}, \widehat{\lambda}_{\mathsf{J}}$ and resulting $(0; \widehat{u}_{\lambda,\gamma})$.

Point estimates of $C_{\mathsf{PX},\epsilon}$ and $C_{\mathsf{BH},\epsilon}$:

$$\widehat{C}_{\mathsf{PX},\epsilon} := 0.0027 \big/ P(\epsilon > USL \mid \lambda = \widehat{\lambda}),$$

$$\widehat{C}_{\mathsf{BH},\epsilon} := \tfrac{1}{3} \cdot \Phi^{-1}\big(1 - \tfrac{1}{2} \cdot P(\epsilon > USL \mid \lambda = \widehat{\lambda})\big).$$

Interval estimates of $C_{\mathsf{PX},\epsilon}$ and $C_{\mathsf{BH},\epsilon}$:

$$\widehat{I}_{C_{\mathsf{PX},\epsilon}} := \big(0.0027 \big/ P(\epsilon > USL \mid \lambda = \widehat{u}_{\lambda,\gamma});\ \infty\big),$$

$$\widehat{I}_{C_{\mathsf{BH},\epsilon}} := \big(\tfrac{1}{3} \cdot \Phi^{-1}\big(1 - \tfrac{1}{2} \cdot P(\epsilon > USL \mid \lambda = \widehat{u}_{\lambda,\gamma})\big);\ \infty\big).$$

Task 2: (continued)

Finite-sample performance:

- coverage of $\widehat{I}_{C_{\mathsf{PX},\epsilon}}, \widehat{I}_{C_{\mathsf{BH},\epsilon}}$ determined by that of $(0; \widehat{u}_{\lambda,\gamma})$, see above;
- again very large bias for $\widehat{C}_{\mathsf{PX},\epsilon}$, especially if based on $\widehat{\lambda}_{\mathsf{J}}$ (!) for small T;
- $\widehat{C}_{\mathsf{BH},\epsilon}$ shows a small bias for $T \geq 50$ if based on $\widehat{\lambda}_{\mathsf{J}}$ (or if based on $\widehat{\lambda}_{\mathsf{MM}}$ for $\alpha \leq 0.25$).

PC Analysis for Serially Dependent Poisson Counts

Conclusions

Conclusions

- PC indices $C_{\mathsf{PX}}, C_{\mathsf{BH}}$ for processes with Poisson marg.;
- Poisson INAR(1) model: capability expressed for observation or innovation process;
- interval estimation possible for both situations and for both indices in satisfactory way (based on new estimator $\hat{\lambda}_{\mathsf{J}}$ if innovation process), but:
- point estimation of C_{PX} problematic due to huge bias (with dangerous tendency to overestimation!)
  $\Rightarrow C_{\mathsf{BH}}$ seems to be preferable for practice!

Thank You

for Your Interest!

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Christian H. Weiß

Department of Mathematics

Darmstadt University of Technology

`weiss@mathematik.tu-darmstadt.de`