Model Diagnostics for Poisson INARMA Processes using Bivariate Dispersion Indexes

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

Christian H. Weiß

Boris Aleksandrov

Department of Mathematics & Statistics,
Helmut Schmidt University, Hamburg

MATH STAT

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Bivariate Dispersion Indexes for Bivariate Poisson Counts

Introduction

Let $\epsilon_0, \epsilon_1, \epsilon_2$ be independent r.v. with $\epsilon_i \sim \text{Poi}(\lambda_i)$.

Bivariate random vector $\boldsymbol{X} := (\epsilon_1 + \epsilon_0, \epsilon_2 + \epsilon_0)^\top$

said to be **bivariately Poisson distributed**

according to $\text{BPoi}(\lambda_0;\ \lambda_1, \lambda_2)$ (Johnson et al., 1997).

Components univariately Poisson distributed: $X_i \sim \text{Poi}(\lambda_i + \lambda_0)$.

Equidispersion: $\mu_i := E[X_i] = \lambda_i + \lambda_0 = V[X_i] =: \sigma_i^2$.

But common summand ϵ_0 causes cross-correlation:

$$\gamma := Cov[X_i, X_j] \ = \ \lambda_0,$$

$$\rho := Corr[X_i, X_j] \ = \ \lambda_0 \cdot \big((\lambda_i + \lambda_0)(\lambda_j + \lambda_0)\big)^{-1/2}.$$

Univariate count distributions: quantify the extent of dispersion in terms of (Poisson) **index of dispersion**,

$$I = \frac{\sigma^2}{\mu}.$$

Departures from 1 indicate non-Poisson distribution.

Bivariate count distributions: several proposals for dispersion index to uncover *deviations from bivariate Poisson distribution*.

(. . .)

Crockett (1979):

$$I_{\mathsf{C}} = \frac{\mu_2^2\,(\sigma_1^2-\mu_1)^2+\mu_1^2\,(\sigma_2^2-\mu_2)^2-2\,(\sigma_1^2-\mu_1)\,(\sigma_2^2-\mu_2)\,\gamma^2}{2(\mu_1^2\,\mu_2^2-\gamma^4)}$$

$$= \frac{(\frac{\sigma_1^2}{\mu_1}-1)^2+(\frac{\sigma_2^2}{\mu_2}-1)^2-2\,(\frac{\sigma_1^2}{\mu_1}-1)\,(\frac{\sigma_2^2}{\mu_2}-1)\,\frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}\,\rho^2}{2(1-\frac{\sigma_1^4}{\mu_1^2}\,\frac{\sigma_2^4}{\mu_2^2}\,\rho^4)}.$$

Simplified type of Crockett's index by Best & Rayner (1997):

$$I_{\mathsf{BR}} = \frac{(\frac{\sigma_1^2}{\mu_1}-1)^2+(\frac{\sigma_2^2}{\mu_2}-1)^2-2\,(\frac{\sigma_1^2}{\mu_1}-1)\,(\frac{\sigma_2^2}{\mu_2}-1)\,\rho^2}{2(1-\rho^4)}.$$

If marginals of $\boldsymbol{X}$ equidispersed: $I_{\mathsf{C}} = I_{\mathsf{BR}} = 0$.

Loukas & Kemp (1986):

$$I_{\mathsf{LK}} = \frac{\mu_2\,\sigma_1^2 + \mu_1\,\sigma_2^2 - 2\,\gamma^2}{\mu_1\,\mu_2 - \gamma^2} = \frac{\frac{\sigma_1^2}{\mu_1} + \frac{\sigma_2^2}{\mu_2} - 2\,\frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}\,\rho^2}{1 - \frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}\,\rho^2}.$$

Simplified version of I_{LK} by Rayner & Best (1995):

$$I_{\mathsf{RB}} = \frac{\frac{\sigma_1^2}{\mu_1} + \frac{\sigma_2^2}{\mu_2} - 2\sqrt{\frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}}\,\rho^2}{1 - \rho^2}.$$

If marginals of $\boldsymbol{X}$ equidispersed: $I_{\mathsf{LK}} = I_{\mathsf{RB}} = 2$.

Sample counterparts from i. i. d. $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$:

$$n\,\hat{I}_{\mathsf{C}},\ n\,\hat{I}_{\mathsf{BR}} \underset{\mathsf{a}}{\sim} \chi^2_2, \qquad n\,\hat{I}_{\mathsf{LK}},\ n\,\hat{I}_{\mathsf{RB}} \underset{\mathsf{a}}{\sim} \chi^2_{2n-3}.$$

Above indexes take bivariate Poisson as benchmark.

Kokonendji & Puig (2018) consider null of

X_1, X_2 being *independent* Poisson variates:

$$I_{\mathrm{KP},1} \;=\; \frac{\mu_1\,\sigma_1^2 + \mu_2\,\sigma_2^2 + 2\,\sqrt{\mu_1\,\mu_2}\,\gamma}{\mu_1^2 + \mu_2^2} \;=\; \frac{\mu_1^2\,\frac{\sigma_1^2}{\mu_1} + \mu_2^2\,\frac{\sigma_2^2}{\mu_2} + 2\,\mu_1\,\mu_2\,\sqrt{\frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}}\,\rho}{\mu_1^2 + \mu_2^2},$$

$$I_{\mathrm{KP},2} \;=\; \frac{\sigma_1^2\,\sigma_2^2 - \gamma^2}{\mu_1\,\mu_2} \;=\; \frac{\sigma_1^2}{\mu_1}\,\frac{\sigma_2^2}{\mu_2}\,(1-\rho^2).$$

If marginals equidispersed:

$$I_{\mathrm{KP},1} = (\mu_1^2 + \mu_2^2 + 2\,\mu_1\,\mu_2\,\rho)/(\mu_1^2 + \mu_2^2), \quad I_{\mathrm{KP},2} = 1-\rho^2.$$

To become 1, indexes additionally require $\rho = 0$.

Aim: Adapt these **bivariate** indexes

for model diagnostics of **univariate** count processes $(X_t)_{\mathbb{Z}}$.

Focus on Poisson INARMA family.

There, many models not only Poisson marginal distribution,

but pairs (X_t, X_{t-k}) being bivariate Poisson.

Asymptotic distribution of sample versions for indexes,

where H_0 either Poisson INAR(1) or Poisson INMA(1) DGP.

Finite-sample performance of resulting dispersion tests

investigated with simulations.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Intermezzo: Poisson INARMA Processes

Definition & Properties

Idea: Substitute multiplication in ARMA recursion by **binomial thinning** $\alpha \circ X | X \sim \text{B}(X, \alpha)$.

INAR(1) model by McKenzie (1985), Alzaid & Al-Osh (1988):

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

where $\alpha \in [0;1)$, i. i. d. count innovations $(\epsilon_t)_{\mathbb{Z}}$ with $E[\epsilon_t] = \mu_\epsilon$, $V[\epsilon_t] = \sigma_\epsilon^2$, and appropriate independence assumptions.

ACF $\rho(k) = \alpha^k$, observations' dispersion index: $I = \dfrac{I_\epsilon + \alpha}{1 + \alpha}$.

Poisson INAR(1): $\epsilon_t \sim \text{Poi}(\lambda)$, $X_t \sim \text{Poi}(\mu)$ with $\mu = \frac{\lambda}{1-\alpha}$.

Pairs $(X_t, X_{t-k}) \sim \text{BPoi}\big(\alpha^k \mu;\ (1-\alpha^k)\,\mu, (1-\alpha^k)\,\mu\big)$.

INAR(p) model by Alzaid & Al-Osh (1990):

$X_t \;=\; \alpha_1 \circ X_{t-1} + \ldots + \alpha_\mathrm{p} \circ X_{t-\mathrm{p}} + \epsilon_t \qquad$ with $\alpha_\bullet := \sum_{j=1}^{\mathrm{p}} \alpha_j < 1.$

where $(\alpha_1 \circ X_t, \ldots, \alpha_\mathrm{p} \circ X_t)$ given X_t multinomially distributed.

Poisson INAR(p): $\epsilon_t \sim$ Poi(λ), $X_t \sim$ Poi(μ) with $\mu = \frac{\lambda}{1-\alpha_\bullet}$.

Example of **Poisson INAR(2) model**:

$$\rho(1) \;=\; \alpha_1, \qquad \rho(k) \;=\; \alpha_1\, \rho(k-1) + \alpha_2\, \rho(k-2) \quad \text{for } k \geq 2,$$

$$(X_t, X_{t-k}) \;\sim\; \text{BPoi}\Big(\rho(k)\, \mu;\; \big(1-\rho(k)\big)\, \mu,\; \big(1-\rho(k)\big)\, \mu\Big).$$

INMA(q) models defined by

$$X_t \;=\; \beta_0 \circ \epsilon_t \;+\; \beta_1 \circ \epsilon_{t-1} \;+ \ldots +\; \beta_{\mathsf{q}} \circ \epsilon_{t-\mathsf{q}}, \qquad \mathsf{q} \geq 1.$$

Different choices possible for distribution of

$(\beta_0 \circ \epsilon_t, \beta_1 \circ \epsilon_t, \ldots, \beta_{\mathsf{q}} \circ \epsilon_t)$ given ϵ_t, see Weiß (2008),

but it *always* holds:

Poisson INMA(q): $\epsilon_t \sim \mathrm{Poi}(\lambda)$, $X_t \sim \mathrm{Poi}(\mu)$ with $\mu = \lambda\, \beta_\bullet$.

$$(X_t, X_{t-k}) \;\sim\; \mathrm{BPoi}\Big(\rho(k)\, \mu;\; \big(1-\rho(k)\big)\, \mu,\; \big(1-\rho(k)\big)\, \mu\Big),$$

where ACF determined from specific INMA(q) model.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Bivariate Dispersion Indexes for Univariate Count Time Series

Definition & Relations

Idea: use bivariate dispersion indexes as diagnostic tools to check if count time series from Poisson INARMA DGP.

If applied univariate stationary count process $(X_t)_{\mathbb{Z}}$ with mean μ, variance σ^2 and ACF $\rho(k)$, above indexes simplify as follows:

$$I_{\mathsf{C}}(k) \;=\; \frac{(\frac{\sigma^2}{\mu}-1)^2}{1+\frac{\sigma^4}{\mu^2}\,\rho(k)^2}, \qquad I_{\mathsf{BR}}(k) \;=\; \frac{(\frac{\sigma^2}{\mu}-1)^2}{1+\rho(k)^2},$$

$$I_{\mathsf{LK}}(k) \;=\; \frac{2\,\frac{\sigma^2}{\mu}\left(1-\frac{\sigma^2}{\mu}\,\rho(k)^2\right)}{1-\frac{\sigma^4}{\mu^2}\,\rho(k)^2}, \qquad I_{\mathsf{RB}}(k) \;=\; 2\,\frac{\sigma^2}{\mu}.$$

Here, $I_{\mathsf{LK}}(k) < 0$ iff $\frac{\sigma^2}{\mu} \in \left(\frac{1}{|\rho(k)|};\, \frac{1}{\rho(k)^2}\right]$.

Indexes by Kokonendji & Puig (2018) become

$$I_{\mathsf{KP},1}(k) \;=\; \tfrac{\sigma^2}{\mu}\big(1+\rho(k)\big), \qquad I_{\mathsf{KP},2}(k) \;=\; \tfrac{\sigma^2}{\mu}\big(1-\rho(k)\big)\, I_{\mathsf{KP},1}(k).$$

Indexes designed to uncover deviations from situation of independent Poisson variates.

So applicability to such count processes, where components of (X_t, X_{t-k}) independent for some lag k.

Within INARMA family, $I_{\mathsf{KP},1}(k), I_{\mathsf{KP},2}(k)$ relevant for Poisson INMA(q) processes for lags $k > \mathsf{q}$ (i. i. d. case for $\mathsf{q} = 0$).

Motivated by $I_{\mathsf{C}}(k), I_{\mathsf{BR}}(k), I_{\mathsf{KP},1}(k), I_{\mathsf{KP},2}(k)$, we define:

$$f_1\big(\mu,\sigma^2,\rho(k)\big) = \frac{\frac{\sigma^2}{\mu}-1}{\sqrt{1+\frac{\sigma^4}{\mu^2}\,\rho(k)^2}}, \quad f_2\big(\mu,\sigma^2,\rho(k)\big) = \frac{\frac{\sigma^2}{\mu}-1}{\sqrt{1+\rho(k)^2}},$$

$$f_3\big(\mu,\sigma^2,\rho(k)\big) = \frac{\sigma^2}{\mu}\big(1+\rho(k)\big), \quad f_4\big(\mu,\sigma^2,\rho(k)\big) = \frac{\sigma^2}{\mu}\sqrt{1-\rho(k)^2}.$$

These either equal to above indexes, or to square root of it.

$f_1, f_2 = 0$ iff $\sigma^2 = \mu$ (equidispersion).

If, in addition, $\rho(k) = 0$, then $f_3, f_4 = 1$.

While univariate index I only considers marginal dispersion,

bivariate indexes also consider autocorrelation structure.

Motivated by $I_{\mathsf{C}}(k), I_{\mathsf{BR}}(k), I_{\mathsf{KP},1}(k), I_{\mathsf{KP},2}(k)$, we define:

$$f_1\big(\mu,\sigma^2,\rho(k)\big) \;=\; \frac{\frac{\sigma^2}{\mu}-1}{\sqrt{1+\frac{\sigma^4}{\mu^2}\,\rho(k)^2}}, \quad f_2\big(\mu,\sigma^2,\rho(k)\big) \;=\; \frac{\frac{\sigma^2}{\mu}-1}{\sqrt{1+\rho(k)^2}},$$

$$f_3\big(\mu,\sigma^2,\rho(k)\big) \;=\; \frac{\sigma^2}{\mu}\,\big(1+\rho(k)\big), \qquad f_4\big(\mu,\sigma^2,\rho(k)\big) \;=\; \frac{\sigma^2}{\mu}\sqrt{1-\rho(k)^2}.$$

f_1, f_2 further inflated if $\rho(k)$ falls below its hypothetical value.

f_3 increased (decreased) if positive (negative) autocorrelation is present (instead of the hypothetical zero autocorrelation).

f_4 always decreased in the presence of autocorrelation.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Sample Indexes for Poisson INARMA Processes

Stochastic Properties

Let $X_1, \ldots, X_T$ be count time series from Poisson INARMA DGP.

Lag-k statistics $\quad \widehat{f}_i(k) \ := \ f_i\big(\bar{X}, \widehat{\gamma}(0), \widehat{\gamma}(k)/\widehat{\gamma}(0)\big)$.

For hypothesis testing, distribution of $\widehat{f}_i(k)$ under null required.

Poisson INAR(p) processes satisfy strong mixing conditions,
Poisson INMA(q) processes q-dependent.
$\Rightarrow$ asymptotic normality of $\widehat{f}_i(k)$ by CLT like Ibragimov (1962).

Closed-form formulae for asymptotic normal distribution generally hard to find $\Rightarrow$ tests relying on $\widehat{f}_i(k)$ to be implemented by parametric bootstrap approach.

Sample Indexes for Poisson INARMA Processes

For two important special cases of Poisson INARMA models, circumvent computationally costly bootstrap implementation: analytical expressions for asymptotics of $\widehat{f}_i(k)$ for Poisson INAR(1) and INMA(1) model (also i. i. d. Poisson).

Both special cases widely applied in practice:

- Cossette et al. (2011) use both models for number of claims within insurance risk model;
- Zhang et al. (2015) use both models in area of reinsurance.
- Hu et al. (2018): applications where claim numbers overdispersed.

So testing null of Poissonity within relevant type of INAR(1) or INMA(1) model of utmost importance for practice.

Poisson INAR(1) process $X_t = \alpha \circ X_{t-1} + \epsilon_t$:

Since $\rho(k) \neq 0$, indexes $\hat{f}_1(k)$ and $\hat{f}_2(k)$ relevant.

Benchmark: univariate index $\hat{I} = \hat{\sigma}^2/\bar{X}$,

asymptotically approximated by normal distribution with

$$\mu_I = 1 - \frac{1}{T}\frac{1+\alpha}{1-\alpha}, \qquad \sigma_I^2 = \frac{2}{T}\frac{1+\alpha^2}{1-\alpha^2}.$$

For bivariate indexes at lag $k=1$, consider

$$\boldsymbol{Y}_t^{(1)} := \big(X_t-\mu,\ X_t^2-\mu(0),\ X_tX_{t-1}-\mu(1)\big)^\top, \qquad \mu(k) := E[X_tX_{t-k}].$$

Weiß & Schweer (2016): $\frac{1}{\sqrt{T}}\sum_{t=1}^T \boldsymbol{Y}_t^{(1)}$ asymptotically normally distributed with (exact) mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}^{(1)}$.

Sample Indexes for Poisson INARMA Processes

By applying Delta method, we derive closed-form expression for asymptotic normal distribution of $\widehat{f}_i(1)$, plus bias correction from 2nd-order Taylor approximation.

Theorem: For Poisson INAR(1) DGP, $\widehat{f}_i(1)$ at lag 1, $i = 1, 2$, asymptotically normal with

$$\mu_{1,1} = 0 - \frac{1}{T}\frac{1+2\alpha+6\alpha^2+2\alpha^3+\alpha^4}{(1-\alpha^2)(1+\alpha^2)^{3/2}}, \qquad \sigma^2_{1,1} = \frac{1}{T}\frac{2}{1-\alpha^2};$$

$$\mu_{2,1} = 0 - \frac{1}{T}\frac{1+\alpha+3\alpha^2-\alpha^3}{(1-\alpha)(1+\alpha^2)^{3/2}}, \qquad \sigma^2_{2,1} = \frac{1}{T}\frac{2}{1-\alpha^2}.$$

So $\widehat{f}_1(1)$ stronger bias than $\widehat{f}_2(1)$.

Note: If use $\widehat{f}_i(1)$ as test statistics, then two-sided decision rule (level a): reject null if $\left|\widehat{f}_i\big(\bar{X}, \widehat{\sigma}^2, \widehat{\rho}(1)\big) - \mu_{i,1}\right| > z_{1-a/2}\, \sigma_{i,1}$.

Since $\mu_{i,1}, \sigma_{i,1}$ depend on α (but not on μ) which is not known in practice, we plug-in $\widehat{\rho}(1)$ instead.

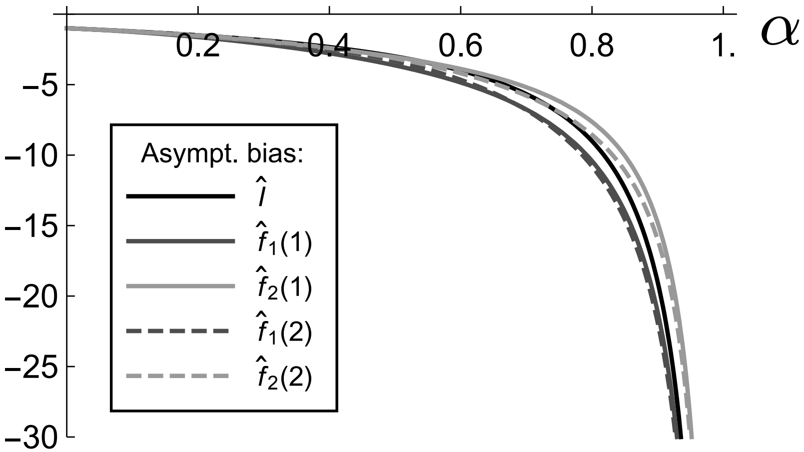
In this case, decision rule for test $\widehat{f}_2(1)$ nearly same as that related to $\widehat{I}$, only mean approximations differ.
So we expect similar performance of these two tests.

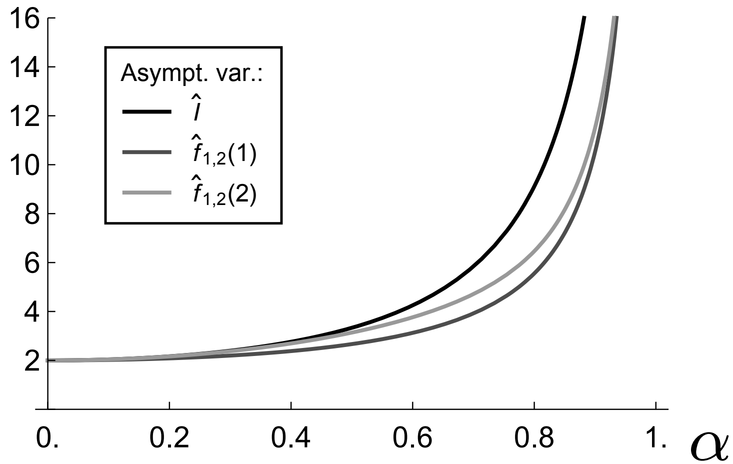
For lag $k = 2$, first derive CLT for $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{Y}_t^{(2)}$. Then ...

Theorem: For Poisson INAR(1) DGP, $\widehat{f}_i(2)$ at lag 2, $i = 1,2$, asymptotically normal with

$$\mu_{1,2} = 0 - \frac{1}{T}\frac{1+2\alpha+\alpha^2+7\alpha^4+2\alpha^5-\alpha^6}{(1-\alpha^2)(1+\alpha^4)^{3/2}}, \qquad \sigma^2_{1,2} = \frac{1}{T}\frac{2}{1+\alpha^4}\frac{1+\alpha^2}{1-\alpha^2};$$

$$\mu_{2,2} = 0 - \frac{1}{T}\frac{1+\alpha+5\alpha^4-3\alpha^5}{(1-\alpha)(1+\alpha^4)^{3/2}}, \qquad \sigma^2_{2,2} = \frac{1}{T}\frac{2}{1+\alpha^4}\frac{1+\alpha^2}{1-\alpha^2}.$$





Poisson INMA(1) process $X_t = \epsilon_t + \beta \circ \epsilon_{t-1}$,

with ACF at lag 1 given by $\rho = \beta/(1+\beta) \in [0; 0.5]$.

Indexes $\widehat{f}_1(1)$ and $\widehat{f}_2(1)$ relevant at lag 1,

indexes $\widehat{f}_1(k), \widehat{f}_2(k)$ plus $\widehat{f}_3(k), \widehat{f}_4(k)$ at lags $k \geq 2$.

Benchmark: univariate index $\hat{I} = \hat{\sigma}^2/\bar{X}$,

asymptotically approximated by normal distribution with

$$\mu_I = 1 - \frac{1}{T}(1+2\rho), \qquad \sigma_I^2 = \frac{2}{T}(1+2\rho^2).$$

Aleksandrov & Weiß (2018): CLT for $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \boldsymbol{Y}_t^{(k)}$. Then ...

Theorem: For Poisson INMA(1) DGP, $\hat{f}_i(1)$ at lag 1, $i = 1, 2$, asymptotically normal with

$$\mu_{1,1} = 0 - \frac{1}{T}\frac{1+2\rho+5\rho^2+2\rho^3}{(1+\rho^2)^{3/2}}, \qquad \sigma^2_{1,1} = \frac{2}{T}\frac{1+2\rho^2}{1+\rho^2};$$

$$\mu_{2,1} = 0 - \frac{1}{T}\frac{1+2\rho+3\rho^2+2\rho^3-4\rho^4}{(1+\rho^2)^{3/2}}, \qquad \sigma^2_{2,1} = \frac{2}{T}\frac{1+2\rho^2}{1+\rho^2}.$$

In analogy to INAR(1) case, test statistics $\hat{I}$ and $\hat{f}_2(1)$ lead to nearly same decision rule, only bias terms differ.

Respective bias and variance always larger in INAR(1) case.

Theorem: For Poisson INMA(1) DGP, $\widehat{f}_i(k)$ at lag $k \geq 2$, $i = 1, \ldots, 4$, asymptotically normal with

$$\mu_{i,k} = 0 - \frac{1}{T}(1+2\rho), \qquad \sigma_{i,k}^2 = \frac{2}{T}(1+2\rho^2) \quad \text{for } i = 1,2;$$

$$\mu_{3,k} = 1 - \frac{2}{T}(1+2\rho), \qquad \sigma_{3,k}^2 = \begin{cases} \frac{1}{T}(3+10\rho^2) & \text{if } k = 2, \\ \frac{3}{T}(1+2\rho^2) & \text{if } k > 2; \end{cases}$$

$$\mu_{4,k} = 1 - \frac{1}{T}(\tfrac{3}{2}+2\rho+\rho^2), \quad \sigma_{4,k}^2 = \frac{2}{T}(1+2\rho^2).$$

Note that $\widehat{f}_1(k), \widehat{f}_2(k)$ at lag $k \geq 2$ have identical asymptotic bias and variance as $\widehat{I}$ although their computations differ.

Case $\rho = 0$: null hypothesis of i. i. d. Poisson counts.

Then all indexes $\hat{f}_1(k), \ldots, \hat{f}_4(k)$ for all lags $k \geq 1$ applicable.

Corollary: For i. i. d. Poisson counts, $\hat{f}_i(k)$ at lag $k \geq 1$, $i = 1, \ldots, 4$, asymptotically normal with

$$\mu_{i,k} = 0 - \frac{1}{T}, \quad \sigma^2_{i,k} = \frac{2}{T} \quad \text{for } i = 1, 2;$$

$$\mu_{3,k} = 1 - \frac{2}{T}, \quad \sigma^2_{3,k} = \frac{3}{T}; \qquad \mu_{4,k} = 1 - \frac{3/2}{T}, \quad \sigma^2_{4,k} = \frac{2}{T}.$$

Asymptotic distribution of $\hat{f}_1(k), \hat{f}_2(k)$ coincides with one of $\hat{I}$.

Asymptotic distributions free of parameters, no plug-in required.

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH
STAT

Sample Indexes for Poisson INARMA Processes

Simulation Study

We analyze finite-sample performance of asymptotic approximations for bivariate dispersion indexes for DGPs Poisson INAR(1), Poisson INMA(1) or i. i. d. Poisson, as well as corresponding hypothesis tests, by simulations (always 10 000 replications per scenario).

Abbreviations:

- SD = standard deviation,
- index "$_{\text{a}}$" = asymptotic,
- index "$_{\text{s}}$" = simulated.

Poisson INAR(1) process with $\mu = 2.5$, $\alpha = 0.5$:

T	$\hat{f}_1(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_2(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_1(2)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_2(2)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$
100	-0.036	-0.034	0.163	0.155	-0.030	-0.029	0.163	0.159	-0.033	-0.031	0.177	0.169	-0.031	-0.029	0.177	0.171
250	-0.015	-0.013	0.103	0.100	-0.012	-0.011	0.103	0.101	-0.013	-0.012	0.112	0.109	-0.013	-0.011	0.112	0.109
500	-0.007	-0.009	0.073	0.072	-0.006	-0.008	0.073	0.073	-0.007	-0.009	0.079	0.078	-0.006	-0.008	0.079	0.079
1000	-0.004	-0.004	0.052	0.051	-0.003	-0.004	0.052	0.051	-0.003	-0.004	0.056	0.055	-0.003	-0.004	0.056	0.055

Poisson INMA(1) process with $\mu = 2.5$, $\rho = 0.4$:

T	$\hat{f}_1(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_2(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_3(2)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_4(2)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$
100	-0.022	-0.019	0.151	0.146	-0.018	-0.016	0.151	0.148	0.964	0.966	0.214	0.207	0.975	0.978	0.162	0.159
250	-0.009	-0.009	0.095	0.094	-0.007	-0.008	0.095	0.095	0.986	0.985	0.136	0.133	0.990	0.989	0.103	0.102
500	-0.004	-0.004	0.067	0.067	-0.004	-0.003	0.067	0.067	0.993	0.993	0.096	0.095	0.995	0.996	0.073	0.072
1000	-0.002	-0.002	0.048	0.048	-0.002	-0.001	0.048	0.048	0.996	0.996	0.068	0.068	0.998	0.998	0.051	0.052

i. i. d. Poisson counts with $\mu = 2.5$:

T	$\hat{f}_1(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_2(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_3(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$	$\hat{f}_4(1)$ Mean$_a$	Mean$_s$	SD$_a$	SD$_s$
100	-0.010	-0.011	0.141	0.140	-0.010	-0.011	0.141	0.140	0.980	0.979	0.173	0.171	0.985	0.984	0.141	0.140
250	-0.004	-0.004	0.089	0.088	-0.004	-0.004	0.089	0.088	0.992	0.992	0.110	0.108	0.994	0.994	0.089	0.088
500	-0.002	-0.002	0.063	0.064	-0.002	-0.002	0.063	0.064	0.996	0.996	0.077	0.078	0.997	0.997	0.063	0.064
1000	-0.001	-0.001	0.045	0.045	-0.001	-0.001	0.045	0.045	0.998	0.998	0.055	0.055	0.999	0.998	0.045	0.045

Simulated size (nominal level 0.05) of two-sided tests . . .

. . . for null hypothesis "DGP is Poisson INAR(1)":

	Under null hypothesis: Poi-INAR(1) with $\mu = 2.5, \alpha = 0.5$					Sensitivity analysis: Poi-INMA(1) with $\mu = 2.5, \rho = 0.4$					Sensitivity analysis: Poi-INAR(2) with $\mu = 2.5, \alpha_1 = 0.5, \alpha_2 = 0.25$				
T	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$
100	0.048	0.045	0.049	0.047	0.049	0.045	0.043	0.045	0.046	0.046	0.120	0.119	0.120	0.102	0.104
250	0.048	0.046	0.049	0.047	0.049	0.045	0.044	0.045	0.047	0.047	0.130	0.129	0.130	0.104	0.104
500	0.049	0.049	0.049	0.049	0.049	0.044	0.044	0.043	0.046	0.046	0.130	0.129	0.131	0.103	0.103
1000	0.050	0.051	0.051	0.051	0.051	0.047	0.046	0.046	0.049	0.049	0.135	0.134	0.135	0.107	0.107

. . . for null hypothesis "DGP is Poisson INMA(1)":

	Under null hypothesis: Poi-INMA(1) with $\mu = 2.5, \rho = 0.4$							Sensitivity analysis: Poi-INAR(1) with $\mu = 2.5, \alpha = 0.4$						
T	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{f}_3(2)$	$\widehat{f}_4(2)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{f}_3(2)$	$\widehat{f}_4(2)$
100	0.046	0.044	0.045	0.044	0.044	0.045	0.045	0.054	0.052	0.053	0.050	0.051	0.130	0.047
250	0.046	0.045	0.047	0.045	0.046	0.045	0.046	0.052	0.051	0.051	0.048	0.049	0.222	0.046
500	0.048	0.046	0.048	0.047	0.048	0.049	0.047	0.052	0.052	0.052	0.049	0.050	0.385	0.047
1000	0.051	0.050	0.051	0.051	0.051	0.051	0.051	0.058	0.057	0.058	0.054	0.055	0.617	0.053

Simulated size (nominal level 0.05) of two-sided tests ...

... for null hypothesis "DGP is i. i. d. Poisson":

	Under null hypothesis: i. i. d. Poisson with $\mu = 2.5$					Sensitivity analysis: Poi-INMA(1) with $\mu = 2.5, \rho = 0.4$					Sensitivity analysis: Poi-INAR(1) with $\mu = 2.5, \alpha = 0.5$				
T	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_3(1)$	$\hat{f}_4(1)$	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_3(1)$	$\hat{f}_4(1)$	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_3(1)$	$\hat{f}_4(1)$
100	0.048	0.047	0.047	0.047	0.047	0.077	0.058	0.060	0.533	0.074	0.109	0.073	0.079	0.612	0.142
250	0.048	0.048	0.048	0.047	0.048	0.087	0.061	0.063	0.857	0.145	0.118	0.079	0.082	0.912	0.333
500	0.053	0.053	0.053	0.050	0.053	0.082	0.061	0.062	0.985	0.259	0.126	0.087	0.088	0.993	0.587
1000	0.049	0.049	0.049	0.048	0.049	0.088	0.067	0.068	1.000	0.477	0.132	0.091	0.091	1.000	0.867

$\hat{f}_3(k) = \hat{I}\left(1 + \hat{\rho}(k)\right)$ most affected by ACF misspecification as values of $1 + \hat{\rho}(k) > 1$ lead to intensified index values.

$\hat{f}_4(k) = \hat{I}\sqrt{1 - \hat{\rho}(k)^2}$ has worse power in detecting model misspecification as $\sqrt{1 - \hat{\rho}(k)^2}$ decreases index value.

Simulated power (nominal level 0.05) of two-sided tests . . .

. . . for null hypothesis "DGP is Poisson INAR(1)":

	NB-INAR(1) with $\mu = 2.5, \alpha = 0.5, I = 1.10$					NB-INAR(1) with $\mu = 2.5, \alpha = 0.5, I = 1.25$					NB-INAR(1) with $\mu = 2.5, \alpha = 0.5, I = 1.50$				
T	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_1(2)$	$\hat{f}_2(2)$	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_1(2)$	$\hat{f}_2(2)$	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_1(2)$	$\hat{f}_2(2)$
100	0.092	0.080	0.094	0.088	0.094	0.283	0.264	0.288	0.279	0.287	0.645	0.630	0.650	0.641	0.647
250	0.152	0.141	0.156	0.150	0.155	0.535	0.520	0.540	0.532	0.538	0.942	0.938	0.943	0.941	0.942
500	0.239	0.229	0.243	0.237	0.241	0.806	0.797	0.810	0.804	0.808	0.998	0.998	0.998	0.998	0.998
1000	0.404	0.396	0.406	0.402	0.406	0.975	0.974	0.976	0.975	0.975	1.000	1.000	1.000	1.000	1.000

. . . for null hypothesis "DGP is Poisson INMA(1)":

	NB-INMA(1) with $\mu = 2.5, \rho = 0.4, I = 1.25$							NB-INMA(1) with $\mu = 2.5, \rho = 0.4, I = 1.50$						
T	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_1(2)$	$\hat{f}_2(2)$	$\hat{f}_3(2)$	$\hat{f}_4(2)$	$\hat{I}$	$\hat{f}_1(1)$	$\hat{f}_2(1)$	$\hat{f}_1(2)$	$\hat{f}_2(2)$	$\hat{f}_3(2)$	$\hat{f}_4(2)$
100	0.332	0.319	0.334	0.328	0.330	0.220	0.330	0.721	0.709	0.724	0.717	0.718	0.534	0.718
250	0.626	0.616	0.628	0.625	0.625	0.442	0.625	0.967	0.966	0.968	0.967	0.967	0.860	0.967
500	0.870	0.866	0.871	0.869	0.870	0.684	0.869	0.999	0.999	0.999	0.999	0.999	0.987	0.999
1000	0.990	0.990	0.990	0.990	0.990	0.920	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Simulation Study

Simulated size (nominal level 0.05) of two-sided tests . . .

. . . for null hypothesis "DGP is i. i. d. Poisson":

	i. i. d. NB with $\mu = 2.5, I = 1.10$					i. i. d. NB with $\mu = 2.5, I = 1.25$					i. i. d. NB with $\mu = 2.5, I = 1.50$				
T	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_3(1)$	$\widehat{f}_4(1)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_3(1)$	$\widehat{f}_4(1)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_3(1)$	$\widehat{f}_4(1)$
100	0.133	0.131	0.131	0.109	0.132	0.406	0.403	0.404	0.311	0.402	0.825	0.823	0.824	0.692	0.824
250	0.215	0.214	0.214	0.170	0.213	0.728	0.725	0.726	0.581	0.727	0.991	0.991	0.991	0.955	0.991
500	0.359	0.358	0.359	0.269	0.359	0.943	0.943	0.943	0.840	0.942	1.000	1.000	1.000	0.999	1.000
1000	0.590	0.589	0.589	0.441	0.589	0.999	0.999	0.999	0.983	0.998	1.000	1.000	1.000	1.000	1.000

. . . for null hypothesis "DGP is Poisson INAR(1)":

	Poi-INARCH(1) with $\mu = 2.5, \alpha = 0.5$					NB-NDARMA(1, 1) with $\mu = 2.5, \phi_1 = 0.5, \varphi_1 = 0.25, I = 1.25$					NB-INMA(1) with $\mu = 2.5, \rho = 0.4, I = 1.25$				
T	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$	$\widehat{I}$	$\widehat{f}_1(1)$	$\widehat{f}_2(1)$	$\widehat{f}_1(2)$	$\widehat{f}_2(2)$
100	0.393	0.372	0.400	0.389	0.397	0.294	0.279	0.297	0.295	0.301	0.341	0.329	0.346	0.346	0.347
250	0.729	0.717	0.732	0.726	0.730	0.479	0.463	0.483	0.479	0.485	0.606	0.596	0.608	0.612	0.612
500	0.945	0.942	0.945	0.945	0.945	0.695	0.684	0.699	0.698	0.702	0.868	0.866	0.869	0.872	0.872
1000	0.999	0.999	0.999	0.999	0.999	0.894	0.890	0.895	0.895	0.897	0.991	0.990	0.991	0.991	0.991

Conclusions

- Univariate time series following Poisson INARMA model have bivariate Poisson for lagged observations
  $\Rightarrow$ adapt indexes to time series context: $\widehat{f}_1(\cdot), \ldots, \widehat{f}_4(\cdot)$.
- $\widehat{f}_1(\cdot), \widehat{f}_2(\cdot)$ used with any Poisson INARMA. $\widehat{f}_2(\cdot)$ often outperforms $\widehat{I}$: $\widehat{f}_2(1)$ general improvement of $\widehat{I}$'s decision rule, outperformed by $\widehat{f}_2(2)$ if also changes in ACF structure.
- For Poisson INMA(q), $\widehat{f}_3(\cdot), \widehat{f}_4(\cdot)$ relevant for lags $\geq$ q.
  Overdispersion within INMA best detected with $\widehat{f}_2(1)$.
  $\widehat{f}_3(\cdot)$ outstanding performance if INMA structure is violated (in addition to possible non-Poissonity) $\Rightarrow$ purposeful use of $\widehat{f}_3(\cdot)$ for such combined alternatives recommended.

Thank You

for Your Interest!

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

Aleksandrov & Weiß (2018): *Parameter estimation . . . for INMA(1) processes*. Submitted.

Alzaid & Al-Osh (1988): *First-order integer-valued autoregr. . . .* Stat. Neerl. 42, 53–61.

Alzaid & Al-Osh (1990): *An integer-valued pth-order . . . .* J. Appl. Prob. 27, 314–324.

Best & Rayner (1997): *Crockett's test of fit for the bivariate Poisson*. Biom. J. 39, 423–430.

Cossette et al. (2011): *Risk models based on time . . . .* Insurance: Math. Econ. 48, 19–28.

Crockett (1979): *A quick test of fit of a bivariate dist*. Interactive Statistics, 185–191.

Hu et al. (2018): *Risk model based on the first-order . . . .* Scand. Act. J. 2018, 412–425.

Ibragimov (1962): *Some limit theorems for stationary . . . .* Theory Prob. Appl. 7, 349–382.

Johnson et al. (1997): *Discrete Multivariate Distributions*. Wiley.

Kokonendji & Puig (2018): *Fisher dispersion index for mult. . . .* J. Mult. Anal. 165, 180–193.

Loukas & Kemp (1986): *The index of dispersion test for . . . .* Biometrics 42, 941–948.

McKenzie (1985): *Some simple models for discrete . . . .* Water Resources Bull. 21, 645–650.

Rayner & Best (1995): *Smooth tests for the bivariate Poisson*. ANZJS 37, 233–245.

Weiß (2008): *Serial dependence . . . of Poisson INARMA models*. JSPI 138, 2975–2990.

Weiß (2018): *An Introduction to Discrete-Valued Time Series*. Wiley.

Weiß & Schweer (2016): *Bias corrections for moment . . . .* SPL 112, 124–130

Zhang et al. (2015): *Optimal reinsurance under adjust. . . .* Scand. Act. J. 2015, 455–467.