Time Reversibility of INAR(1) Processes and Testing for Poisson Innovations

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

Sebastian Schweer
Department of Mathematics,
University of Heidelberg

Christian H. Weiß
Department of Mathematics & Statistics,
Helmut Schmidt University, Hamburg

MATH
STAT

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH STAT

Introduction

Motivation & Outline

Popular count-data counterpart to conventional AR(1) model:
INAR(1) model by McKenzie (1985).

Let $(\epsilon_t)_{\mathbb{Z}}$ be i.i.d. with range $\mathbb{N}_0 = \{0, 1, \ldots\}$,
denote $E[\epsilon_t] = \mu_\epsilon$, $V[\epsilon_t] = \sigma_\epsilon^2$.
Let "$\alpha\circ$" be **binomial thinning** operator with $\alpha \in (0;1)$
(Steutel & van Harn, 1979).
$(X_t)_{\mathbb{Z}}$ referred to as **INAR(1) process** if

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

together with appropriate independence assumptions.

INAR(p) models by Alzaid & Al-Osh (1990), Du & Li (1991).

Properties of **INAR(1) models**:

Homogeneous Markov chain with

$$P(X_t = k \mid X_{t-1} = l) \;=\; \sum_{j=0}^{\min\{k,l\}} \binom{l}{j} \alpha^j (1-\alpha)^{l-j} \cdot P(\epsilon_t = k-j).$$

If INAR(1) process stationary (Heathcote, 1966), then

$$\mathrm{pgf}_X(z) \;=\; \mathrm{pgf}_X(1-\alpha+\alpha z) \cdot \mathrm{pgf}_\epsilon(z).$$

Autocorrelation function: $\rho_X(k) = \alpha^k$, i. e., AR(1)-type.

For further properties and references, see Weiß (2008).

Most popular instance of INAR(1) family: **Poisson** INAR(1), where innovations $(\epsilon_t)_{\mathbb{Z}}$ i.i.d. Poi(λ), such that $\mu_\epsilon = \sigma_\epsilon^2 = \lambda$.

But also many other choices for distribution of ϵ_t, e. g.,

- with $\sigma_\epsilon^2 > \mu_\epsilon$ (**overdispersion**) (Schweer & Weiß, 2014),
- with $\sigma_\epsilon^2 < \mu_\epsilon$ (**underdispersion**) (Weiß, 2013).

Note that the dispersion behaviour carries over to observations:

$$\mu_X = \frac{\mu_\epsilon}{1-\alpha}, \qquad I_{\text{disp}} := \frac{\sigma_X^2}{\mu_X} = \frac{\frac{\sigma_\epsilon^2}{\mu_\epsilon} + \alpha}{1+\alpha}.$$

How discriminate between Poisson INAR(1) (=null hypothesis) and alternative type of INAR(1) model?

Diagnostic testing of null of *Poisson* INAR(1) model:

- Meintanis & Karlis (2014): test based on empirical bivariate probability generating function; comprehensive ("omnibus test"), but computationally very demanding (complex bootstrap implementation).
- Schweer & Weiß (2014): test based on empirical dispersion index $\hat{I}_{\mathrm{disp}} := S_X^2/\bar{X}$; easy to implement (simple closed-form formulae), powerful only if dispersion behaviour affected.

Diagnostic testing of null of *Poisson* INAR(1) model:

Test procedure that is easy to implement,
but can deal with alternatives that
do not affect (much) dispersion behaviour?

Is there *unique feature* of Poisson INAR(1) model
that could be used to discriminate Poisson INAR(1)
from alternative type of INAR(1) model?

Idea: time-reversibility.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Time Reversibility of INAR(1) Processes

Properties & Results

Stationary $(X_t)_{\mathbb{Z}}$ **time-reversible** iff for all $t_1, \ldots, t_k, h \in \mathbb{Z}$, $(X_{h+t_1}, \ldots, X_{h+t_n})$ and $(X_{h-t_1}, \ldots, X_{h-t_n})$ have same joint distribution ($\approx$ "symmetry in time").

Let $(X_t)_{\mathbb{Z}}$ be **stationary INAR(1) process**.

- Theorem 19.1 in Schweer (2015):
  Also let $P(\epsilon = 0) \in (0; 1)$. Then
  $(X_t)_{\mathbb{Z}}$ **time-reversible** iff ϵ Poisson-distributed.

- Theorem 2.2 in Schweer & Weiß (2015):
  Also let $(X_t)_{\mathbb{Z}}$ be irreducible. Then
  $(X_t)_{\mathbb{Z}}$ **time-reversible** iff ϵ Poisson-distributed.

Essentially:

Exactly *Poisson* INAR(1) processes are time-reversible.

Analogous result for continuous-valued ARMA processes:

Weiss (1975), Theorem 2:

Real-valued ARMA process is time-reversible iff

innovations' distribution is Gaussian.

So in integer-valued case,

Poisson distribution takes place of Gaussian distribution.

Does this extend to INAR(p) models?

Two types of INAR(p) models:

- **INAR(p) model** by Alzaid & Al-Osh (1990):
  - ✔ Theorem 19.2 in Schweer (2015): Let $(X_t)_{\mathbb{Z}}$ be stationary AA-INAR(p) process with $P(\epsilon = 0) \in (0; 1)$. Then $(X_t)_{\mathbb{Z}}$ **time-reversible** iff ϵ Poisson-distributed.
  - ✔ Alzaid & Al-Osh (1990): AA-INAR(p) process with Poisson innovations has Poisson observations.
  - ✘ Alzaid & Al-Osh (1990): AA-INAR(p) process has ARMA(p, p–1) autocorrelation structure.

(. . .)

Two types of INAR(p) models:

- **INAR(p) model** by Alzaid & Al-Osh (1990): (. . .)
- **INAR(p) model** by Du & Li (1991):
  - ✘ Theorem 19.3 in Schweer (2015): DL-INAR(p) process time-reversible only if parameters take degenerate values.
  - ✘ Du & Li (1991): DL-INAR(p) process with Poisson innovations does not have Poisson observations.
  - ✔ Du & Li (1991): DL-INAR(p) process has AR(p) autocorrelation structure.

Let us look back to INAR(1) case.

Time-reversibility is *unique feature* of *Poisson* INAR(1).

How can we utilize time-reversibility

to discriminate Poisson INAR(1) model

from alternative type of INAR(1) model?

Idea: Look at moments $\mu_{i,j}(k) := E[X_t^i X_{t-k}^j]$;

if time-reversible, then $\mu_{i,j}(k) = \mu_{j,i}(k)$ has to hold.

1st approach: "generalized autocovariances"

$$\beta(k) \;:=\; \mu_{2,1}(k)-\mu_{1,2}(k) \;=\; E[X_{t+k}^2 X_t]-E[X_{t+k}X_t^2].$$

Using moment formulae from Schweer & Weiß (2014),

$$\beta(k) \;=\; \alpha^k(1-\alpha^k)\,(\bar{\mu}_{X,3}-\sigma_X^2),$$

which equals 0 for Poisson INAR(1) model.

2nd approach: "skewness index"

$$I_{\text{skew}} \;:=\; \frac{\bar{\mu}_{X,3}}{\sigma_X^2},$$

which equals 1 for Poisson INAR(1) model.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Testing Time Reversibility of INAR(1) Processes

Statistics & Properties

Considered test statistics:

1st approach: "generalized autocovariances"

$$\widehat{\beta}(k) \;:=\; \tfrac{1}{T-k}\textstyle\sum_{t=1}^{T-k}(X_{t+k}^2 X_t - X_{t+k}X_t^2),$$

which have mean *exactly* 0 if time-reversible.

2nd approach: "skewness index"

$$\widehat{I}_{\text{skew}} \;:=\; \frac{m_{X,3}}{S_X^2},$$

where $S_X^2 = \frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})^2$ and $m_{X,3} = \frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})^3$
with $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$.

Theorem: Let $(X_t)_{\mathbb{Z}}$ be *Poisson* INAR(1) process. Then

$$\sqrt{T-1}\cdot\widehat{\beta}(1) \quad \stackrel{\mathcal{D}}{\to} \quad \mathrm{N}\left(0, \sigma_{\mathrm{a}}^2\right) \qquad \text{for } T\to\infty,$$

where

$$\sigma_{\mathrm{a}}^2 \;=\; 4\mu_X^2(1-\alpha)^2\left(2\,\frac{\alpha}{1+\alpha} \;+\; \mu_X\frac{(1-\alpha)(1+\alpha)^2}{1+\alpha+\alpha^2}\right).$$

Exact variance $V[\sqrt{T-1}\widehat{\beta}(1)]$ given by

$$\sigma_{\mathrm{a}}^2 + \;\tfrac{2\mu_X^2}{T-1}\Big(\mu_X(\mu_X+2\alpha)^2(1-\alpha^{T-1})$$
$$+\;2\big(\mu_X(1+\alpha)+2\alpha\big)^2\,\tfrac{1-\alpha^{2T-2}}{(1+\alpha)^2} \;+\; 6\alpha^2\mu_X\,\tfrac{1-\alpha^{3T-3}}{(1+\alpha+\alpha^2)^2}\Big).$$

Theorem: Let $(X_t)_{\mathbb{Z}}$ be *Poisson* INAR(1) process. Then

$$\sqrt{T}(\hat{I}_{\mathsf{skew}}-1) \quad \stackrel{\mathcal{D}}{\rightarrow} \mathrm{N}\left(0,\ 8\,\frac{1+\alpha^2}{1-\alpha^2}+6\mu_X\,\frac{1+\alpha^3}{1-\alpha^3}\right) \qquad \text{for } T\to\infty.$$

Improved approximation for mean:

$$E[\hat{I}_{\mathsf{skew}}] \;\approx\; 1-\frac{2}{T}\,\frac{3+2\alpha+3\alpha^2}{1-\alpha^2}.$$

Both asymptotic results can be used to construct tests based on $\widehat{\beta}(1)$ or $\hat{I}_{\mathsf{skew}}$, respectively, by plugging-in moment estimates for μ_X and α.

Proofs: (sketch)

- Poisson INAR(1) α-mixing with expon. decreasing weights (Schweer & Weiß, 2014), existing moments of any order.
- Carries over to vector-valued process

$$\boldsymbol{Y}_t := (X_t - \mu_X, X_t^2 - \mu_X - \mu_X^2, X_t^3 - \mu_X - 3\mu_X^2 - \mu_X^3)^\top.$$

- CLT of Ibragimov (1962) applicable.
- Use result about cumulants for Poisson INAR(1) by Pickands & Stine (1997), Schweer & Weiß (2015):

$$\mathrm{cum}(X_{i_1}, \ldots, X_{i_r}) \;=\; \mu_X\, \alpha^{i_r - i_1}, \qquad i_1 \leq \cdots \leq i_r.$$

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH STAT

Testing Time Reversibility of INAR(1) Processes

Empirical Investigations

Finite-sample performance of asymptotic approximation for mean and std. dev. of $\hat{I}_{\text{skew}}$ (Poi. INAR(1) with $\epsilon_t \sim$ Poi(λ)):

Mean	λ	0.5		1		2		4	
α	T	sim.	asymp.	sim.	asymp.	sim.	asymp.	sim.	asymp.
0.5	100	0.887	0.873	0.887	0.873	0.886	0.873	0.879	0.873
	250	0.950	0.949	0.948	0.949	0.951	0.949	0.953	0.949
	500	0.977	0.975	0.974	0.975	0.980	0.975	0.978	0.975
	1000	0.988	0.987	0.989	0.987	0.987	0.987	0.991	0.987

S.D.	λ	0.5		1		2		4	
α	T	sim.	asymp.	sim.	asymp.	sim.	asymp.	sim.	asymp.
0.5	100	0.359	0.459	0.456	0.536	0.565	0.665	0.762	0.866
	250	0.265	0.290	0.312	0.339	0.394	0.420	0.523	0.548
	500	0.194	0.205	0.228	0.240	0.291	0.297	0.379	0.387
	1000	0.142	0.145	0.167	0.170	0.205	0.210	0.272	0.274

Simulated size for Poisson INAR(1) with $\epsilon_t \sim \text{Poi}(\lambda)$:

$\widehat{\beta}(1)$-test:

		Size (specified parameters)				Size (estimated parameters)			
α	$T \setminus \lambda$	0.5	1	2	4	0.5	1	2	4
0.5	100	0.062	0.055	0.054	0.059	0.045	0.045	0.044	0.055
	250	0.055	0.048	0.054	0.052	0.051	0.044	0.050	0.050
	500	0.052	0.050	0.049	0.053	0.053	0.050	0.050	0.053
	1000	0.056	0.049	0.051	0.052	0.054	0.049	0.050	0.050

$\widehat{I}_{\text{skew}}$-test:

		Size (specified parameters)				Size (estimated parameters)			
α	$T \setminus \lambda$	0.5	1	2	4	0.5	1	2	4
0.5	100	0.023	0.029	0.026	0.029	0.019	0.028	0.025	0.029
	250	0.032	0.035	0.036	0.040	0.033	0.035	0.036	0.040
	500	0.037	0.037	0.047	0.044	0.038	0.037	0.047	0.044
	1000	0.043	0.044	0.045	0.050	0.044	0.044	0.044	0.049

Simulated power for NB-INAR(1) with $I_\epsilon = 1.5$:

	n	1	$(\mu_\epsilon = 0.5)$		4	$(\mu_\epsilon = 2)$	
α	T	$\widehat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\widehat{I}_{\text{skew}}$	$\widehat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\widehat{I}_{\text{skew}}$
0.3	100	0.556	0.370	0.394	0.600	0.283	0.289
	250	0.869	0.542	0.723	0.915	0.399	0.557
	500	0.987	0.738	0.938	0.997	0.540	0.827
	1000	1.000	0.890	0.999	1.000	0.719	0.979
0.7	100	0.226	0.328	0.126	0.234	0.129	0.085
	250	0.435	0.646	0.276	0.440	0.244	0.164
	500	0.677	0.884	0.460	0.697	0.523	0.268
	1000	0.922	0.985	0.736	0.926	0.841	0.450

$\widehat{\beta}(1)$ superior for large α and low μ_ϵ.

An **equidispersed alternative** to Poisson distribution:

Good distribution (Weiß, 2013).

(q', ν)	Good(q', ν) Mean	Var.	Skew.	Exc.
$(-2.5119, -2.6470)$	0.500	0.500	1.455	2.335
$(-1.9560, -2.9135)$	1.000	1.000	1.103	1.571
$(-1.5139, -3.5533)$	2.000	2.000	0.926	1.289
$(-1.2498, -5.2489)$	4.000	4.000	0.800	0.961

λ	Poi(λ) Mean	Var.	Skew.	Exc.
0.5	0.500	0.500	1.414	2.000
1	1.000	1.000	1.000	1.000
2	2.000	2.000	0.707	0.500
4	4.000	4.000	0.500	0.250

Simulated power for Good-INAR(1):

	(q', ν)	(−1.5139, −3.5533)			(−1.2498, −5.2489)		
α	T	$\hat{I}_{\text{disp}}$	$\hat{\beta}(1)$	$\hat{I}_{\text{skew}}$	$\hat{I}_{\text{disp}}$	$\hat{\beta}(1)$	$\hat{I}_{\text{skew}}$
0.3	100	0.052	0.078	0.078	0.058	0.076	0.108
	250	0.066	0.084	0.126	0.060	0.091	0.174
	500	0.064	0.104	0.180	0.061	0.107	0.290
	1000	0.066	0.119	0.279	0.061	0.166	0.486
0.5	100	0.047	0.069	0.053	0.045	0.070	0.061
	250	0.054	0.086	0.078	0.053	0.081	0.104
	500	0.054	0.102	0.108	0.055	0.118	0.148
	1000	0.056	0.148	0.145	0.058	0.192	0.247

$\hat{I}_{\text{skew}}$ superior, but generally low power.

Comparison with $W_{T,2}$-test from Meintanis & Karlis (2014):

(Italic values for $W_{T,2}$ taken from Table 1 in Meintanis & Karlis (2014))

	$T = 100$				$T = 500$			
Alternative model	$W_{T,2}$	$\widehat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\widehat{I}_{\text{skew}}$	$W_{T,2}$	$\widehat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\widehat{I}_{\text{skew}}$
NB INAR(1)								
$\mu_\epsilon = 1,\ \alpha = 0.5,\ I_\epsilon = 1.1$	*0.120*	0.074	0.078	0.046	*0.149*	0.141	0.122	0.100
$\mu_\epsilon = 1,\ \alpha = 0.5,\ I_\epsilon = 1.5$	*0.327*	0.403	0.323	0.229	*0.897*	0.941	0.782	0.718
$\mu_\epsilon = 1,\ \alpha = 0.5,\ I_\epsilon = 2.0$	*0.686*	0.798	0.651	0.538	*1.000*	1.000	0.987	0.993
Poisson INAR(2)								
$\alpha_1 = 0.5,\ \alpha_2 = 0.3,\ \lambda = 1$	*0.191*	0.353	0.116	0.100	*0.487*	0.908	0.224	0.452
$\alpha_1 = 0.5,\ \alpha_2 = 0.1,\ \lambda = 1$	*0.089*	0.082	0.060	0.039	*0.127*	0.167	0.082	0.108
$\alpha_1 = 0.3,\ \alpha_2 = 0.1,\ \lambda = 1$	*0.088*	0.055	0.047	0.035	*0.035*	0.072	0.053	0.055

$\widehat{I}_{\text{disp}}$ most powerful.

Comparison with $W_{T,2}$-test from Meintanis & Karlis (2014):

(Italic values for $W_{T,2}$ taken from Table 1 in Meintanis & Karlis (2014))

	$T = 100$				$T = 500$			
Alternative model	$W_{T,2}$	$\hat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\hat{I}_{\text{skew}}$	$W_{T,2}$	$\hat{I}_{\text{disp}}$	$\widehat{\beta}(1)$	$\hat{I}_{\text{skew}}$
INGARCH(1,1)								
$\gamma_1 = 0.4,\ \gamma_2 = 0.1,\ \delta = 0.2$	*0.012*	0.052	0.065	0.035	*0.050*	0.059	0.061	0.054
$\gamma_1 = 0.4,\ \gamma_2 = 0.2,\ \delta = 0.2$	*0.042*	0.082	0.078	0.059	*0.108*	0.172	0.101	0.145
$\gamma_1 = 0.1,\ \gamma_2 = 0.5,\ \delta = 1.0$	*0.152*	0.449	0.178	0.208	*0.512*	0.975	0.390	0.745
$\gamma_1 = 0.1,\ \gamma_2 = 0.5,\ \delta = 0.5$	*0.248*	0.433	0.205	0.259	*0.724*	0.962	0.461	0.835
$\gamma_1 = 0.4,\ \gamma_2 = 0.4,\ \delta = 0.1$	*0.296*	0.385	0.218	0.264	*0.902*	0.938	0.466	0.868
$\gamma_1 = 0.1,\ \gamma_2 = 0.6,\ \delta = 0.6$	*0.376*	0.686	0.273	0.374	*0.988*	1.000	0.608	0.967
$\gamma_1 = 0.2,\ \gamma_2 = 0.6,\ \delta = 0.1$	*0.534*	0.627	0.367	0.479	*0.998*	0.997	0.748	0.991
$\gamma_1 = 0.5,\ \gamma_2 = 0.45,\ \delta = 0.1$	*0.669*	0.792	0.348	0.477	*0.986*	1.000	0.704	0.998

$\hat{I}_{\text{disp}}$ most powerful.

Conclusions

- Time reversibility as unique feature of Poisson INAR(1).
- Test statistic $\widehat{\beta}(1)$ based on generalized autocovariance as well as skewness index $\widehat{I}_{\text{skew}}$.
- Closed-form formulae for asymptotic distributions, hence easy to implement (no bootstrap required).
- $\widehat{\beta}(1)$-test powerful for large autocorrelation, hence complements $\widehat{I}_{\text{disp}}$-test.
- $\widehat{I}_{\text{skew}}$-test useful if equidispersed alternative like Good INAR(1) model.

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`

Schweer & Weiß (2015):
Testing for Poisson Arrivals in INAR(1) Processes. TEST, appears.

Alzaid & Al-Osh (1990): *An integer-valued pth-order autoregressive structure (INAR(p)) process*. J. Appl. Probab. 27, 314–324.

Du & Li (1991): *The integer-valued autoregressive (INAR(p)) model*. J. Time Ser. Anal. 12, 129–142.

Heathcote (1966): *A branching process allowing immigration*. J. Roy. Stat. Soc. B 28, 213–217.

Ibragimov (1962): *Some limit theorems for stationary processes*. Theor. Probab. Appl. 7, 349–382.

McKenzie (1985): *Some simple models for discrete variate time series*. Water Res. Bull. 21, 645–650.

Meintanis & Karlis (2014): *Validation tests for the innovation distribution in INAR time series models*. Comput. Stat. 29, 1221–1241.

Pickands & Stine (1997): *Estimation for an $M/G/\infty$ queue with incomplete information*. Biometrika 84, 295–308.

(. . .)

Literature

Schweer & Weiß (2015):
Testing for Poisson Arrivals in INAR(1) Processes. TEST, appears.

(. . .)

Schweer (2015): *On the time-reversibility of integer-valued autoregressive processes of general order*. In Steland (eds) Stochastic models, statistics and their applications, Springer Proc. Math. Stat. 122, 169–177.

Schweer & Weiß (2014): *Compound P. INAR(1) processes: stoch. properties and testing for overdispersion*. Comput. Stat. Data Anal. 77, 267–284.

Steutel & van Harn (1979): *Discrete analogues of self-decomposability and stability* Ann. Probab. 7, 893–899.

Weiss (1975): *Time-reversibility of linear stochastic processes*. J. Appl. Probab. 12, 831–836

Weiß (2008): *Thinning operations for modelling time series of counts – a survey*. Adv. Stat. Anal. 92, 319–341.

Weiß (2013): *Integer-valued autoregressive models for counts showing underdispersion*. J. Appl. Stat. 40, 1931–1948.