Testing the Compounding Structure of the CP-INARCH Model

Christian H. Weiß

Department of Mathematics & Statistics,
Helmut Schmidt University, Hamburg

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

Esmeralda Gonçalves,
Nazaré Mendes Lopes

Departamento de Matemática,
Universidade de Coimbra

MATH STAT

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Compound Poisson INARCH Models

Introduction

INARCH(p) models assume

linear conditional mean $M_t := E[X_t \mid X_{t-1}, \ldots]$:

$$M_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i} \quad \text{with } \alpha_0 > 0,\ \alpha_1, \ldots, \alpha_p \geq 0.$$

Poisson INARCH(p) model (Heinen, 2003; Ferland et al., 2006):

$$X_t \,\Big|\, X_{t-1}, X_{t-2}, \ldots \quad \sim \ \mathrm{Poi}(M_t),$$

i. e., unconditional overdispersion $\quad V[X_t] > E[X_t]$,

but conditional equidispersion:

$$V[X_t \mid X_{t-1}, \ldots] = E[X_t \mid X_{t-1}, \ldots] = M_t.$$

Dispersed INARCH(p) model (Xu et al., 2012):

$$V[X_t \mid X_{t-1}, \ldots] \;=\; \theta\, M_t \quad \text{with some } \theta > 1.$$

Here, **compound Poisson (CP)** INARCH(p) model

by Gonçalves et al. (2015):

$$\text{pgf}_{X_t|X_{t-1},\ldots}(z) \;=\; \exp\left(\frac{M_t}{H'(1)}\,\big(H(z)-1\big)\right),$$

where $H(z)$ is pgf of compounding distribution

(normalized to $H(0) = 0$ for uniqueness). So

$$V[X_t \mid X_{t-1}, \ldots] \;=\; M_t\, \underbrace{\big(1 + H''(1)/H'(1)\big)}_{=\theta}.$$

Example: NB-INARCH(p) model

with conditional NB(n, π)-distribution (Xu et al., 2012):

$$X_t \mid X_{t-1}, X_{t-2}, \ldots \quad \sim \; \mathrm{NB}\Big(\frac{M_t}{\theta-1},\, \frac{1}{\theta}\Big) \quad \text{with some } \theta > 1.$$

Belongs to CP-INARCH family with

$$H(z) \;=\; 1 - \frac{\ln\big(\theta + (1-\theta)z\big)}{\ln\theta}.$$

Using that $\quad H'(1) \;=\; -\frac{1-\theta}{\ln\theta}, \quad H''(1) \;=\; \frac{(1-\theta)^2}{\ln\theta},$

it follows that

$$V[X_t \mid X_{t-1}, \ldots] \;=\; \theta\, M_t \quad \text{with some } \theta > 1.$$

Aim in the sequel: Test to distinguish between

$$H_0: \ (X_t)_{\mathbb{Z}} \text{ is a Poisson INARCH process;}$$
$$H_1: \ (X_t)_{\mathbb{Z}} \text{ is a true CP-INARCH process.}$$

Equivalently expressed as

$$H_0: \ H(z) = z;$$
$$H_1: \ H(z) \neq z.$$

Equivalently expressed as

$$H_0: \ H'(1) = 1;$$
$$H_1: \ H'(1) > 1.$$

Note: $H'(1)$ is mean of compounding distribution.

Compounding Structure of CP-INARCH Models

Properties

CP-INARCH model (Gonçalves et al., 2015)

$$\mathrm{pgf}_{X_t|X_{t-1},\ldots}(z) \;=\; \exp\left(\frac{M_t}{H'(1)}\big(H(z)-1\big)\right)$$

implies following data-generating mechanism:

Given past observations $X_{t-1},\ldots$,

- generate stopping count $N_t \;\sim\; \mathrm{Poi}\big(M_t/H'(1)\big)$,
- generate (independently) N_t i. i. d. counts $Y_{t,1},\ldots,Y_{t,N_t}$ according to compounding model $H(z)$,
- compute next observation as $X_t = Y_1 + \ldots + Y_{N_t}$.

To distinguish between

$$H_0: \quad H'(1) = 1 \qquad \text{and} \qquad H_1: \quad H'(1) > 1,$$

mean statistic

$$\frac{1}{T}\sum_{t=1}^{T} \frac{Y_{t,1} + \ldots + Y_{t,N_t}}{N_t} \;=\; \frac{1}{T}\sum_{t=1}^{T} \frac{X_t}{N_t}$$

would be reasonable to infer $H'(1)$.

Problem: we do not observe N_t in practice,

we only know that it has mean $M_t/H'(1)$.

2nd trial: we may consider a slightly modified version,

$$\frac{1}{T}\sum_{t=1}^{T}\frac{X_t}{M_t},$$

should give values close to 1.

Note that X_t/M_t are residuals ε_t in Zhu & Wang (2010).

So possible test statistic:

$$\widehat{C}_p \;:=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{X_t}{M_t},$$

but does it allow to distinguish between H_0 and H_1?

Properties of $\widehat{C}_p$: Since

$$E\left[\frac{X_t}{M_t}\right] = 1, \qquad Cov\left[\frac{X_t}{M_t}, \frac{X_{t-k}}{M_{t-k}}\right] = 0 \quad \text{for } k \geq 1,$$

and

$$V\left[\frac{X_t}{M_t}\right] = \left(1 + \frac{H''(1)}{H'(1)}\right) \underbrace{E\left[M_0^{-1}\right]}_{\text{inverse moment!}},$$

we obtain

$$E\left[\widehat{C}_p\right] = 1, \qquad V\left[\widehat{C}_p\right] = \frac{1}{T-p}\left(1 + \frac{H''(1)}{H'(1)}\right) E\left[\frac{1}{M_0}\right].$$

So variance of $\widehat{C}_p$ inflated by $1 + H''(1)/H'(1)$,

but mean of $\widehat{C}_p$ is always 1, independent of CP-INARCH model.

3rd trial: higher-order extension.

Conditional Poisson distribution of H_0 implies

$$E\big[(X_t)_{(r)}\mid X_{t-1},\ldots\big]\ =\ M_t^r.$$

r^{th}**-order statistic** (we consider $r\geq 2$):

$$\widehat{C}_{p;\,r}\ :=\ \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{(X_t)_{(r)}}{M_t^r}$$

$$=\ \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{X_t(X_t-1)\cdots(X_t-r+1)}{\big(\alpha_0+\sum_{i=1}^{p}\alpha_i X_{t-i}\big)^r}.$$

r^{th}-order statistic

$$\widehat{C}_{p;\,r} \;=\; \frac{1}{T-p}\sum_{t=p+1}^{T}\frac{(X_t)_{(r)}}{M_t^r}$$

satisfies under H_0:

$$E\big[\widehat{C}_{p;\,r}\big] \;=\; 1, \qquad V\big[\widehat{C}_{p;\,r}\big] \;=\; \frac{1}{T-p}\sum_{k=1}^{r}\binom{r}{k}^2 k!\,E\big[M_t^{-k}\big].$$

Example: $r=2$ and CP-INARCH(p) model:

$$E\big[\widehat{C}_{p;\,2}\big] \;=\; 1+(\theta-1)\,E\big[M_0^{-1}\big].$$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Compounding Structure of CP-INARCH Models

Significance Test

In the sequel,

to get feasible analytic expressions,

we concentrate of $p = 1$, i. e., **CP-INARCH(1) model**.

For $p > 1$, bootstrap implementation would be required.

Under H_0 (Poi-INARCH(1) model), $\widehat{C}_{1;\,r}$ is asympt. normal (Neumann (2011): mixing properties; Ibragimov (1962): CLT), and we have (see before)

$$E\big[\widehat{C}_{1;\,r}\big] \;=\; 1, \qquad V\big[\widehat{C}_{1;\,r}\big] \;=\; \frac{1}{T-1}\sum_{k=1}^{r} \binom{r}{k}^2 k!\, q_{0,k}.$$

On previous slide, abbreviation

$$q_{k,l} := q_{k,l}(\alpha_0, \alpha_1) \;:=\; E\Big[\frac{X^k}{(\underbrace{\alpha_0 + \alpha_1 X}_{=M})^l}\Big] \qquad \text{for } k, l \geq 0.$$

Computation:

$$q_{k,l} \;=\; \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} \frac{\alpha_0^{k-j}}{\alpha_1^k} \cdot E\big[(\alpha_0 + \alpha_1 X)^{j-l}\big],$$

where $$E\big[(\alpha_0 + \alpha_1 X)^{j-l}\big] \;=\; \begin{cases} q_{0,l-j} & \text{if } j < l, \\ \sum_{i=0}^{j-l} \binom{j-l}{i} \alpha_0^{j-l-i} \alpha_1^i \mu_i & \text{if } j \geq l. \end{cases}$$

So $q_{k,l}$ traced back to usual moments μ_k

or purely inverse moments $q_{0,l}$.

Computation of $\quad q_{0,l} = E\big[(\alpha_0 + \alpha_1\, X)^{-l}\big]$:

First note bounds $\quad 0 < q_{0,l} < \alpha_0^{-l}$.

Numerical computation via **Markov chain approach**:

$$\mathbf{P}_M := (p_{r|s})_{r,s=0,\ldots,M} \quad \text{with } p_{r|s} \;=\; e^{-\alpha_0 - \alpha_1\, s}\, \frac{(\alpha_0 + \alpha_1\, s)^r}{r!},$$

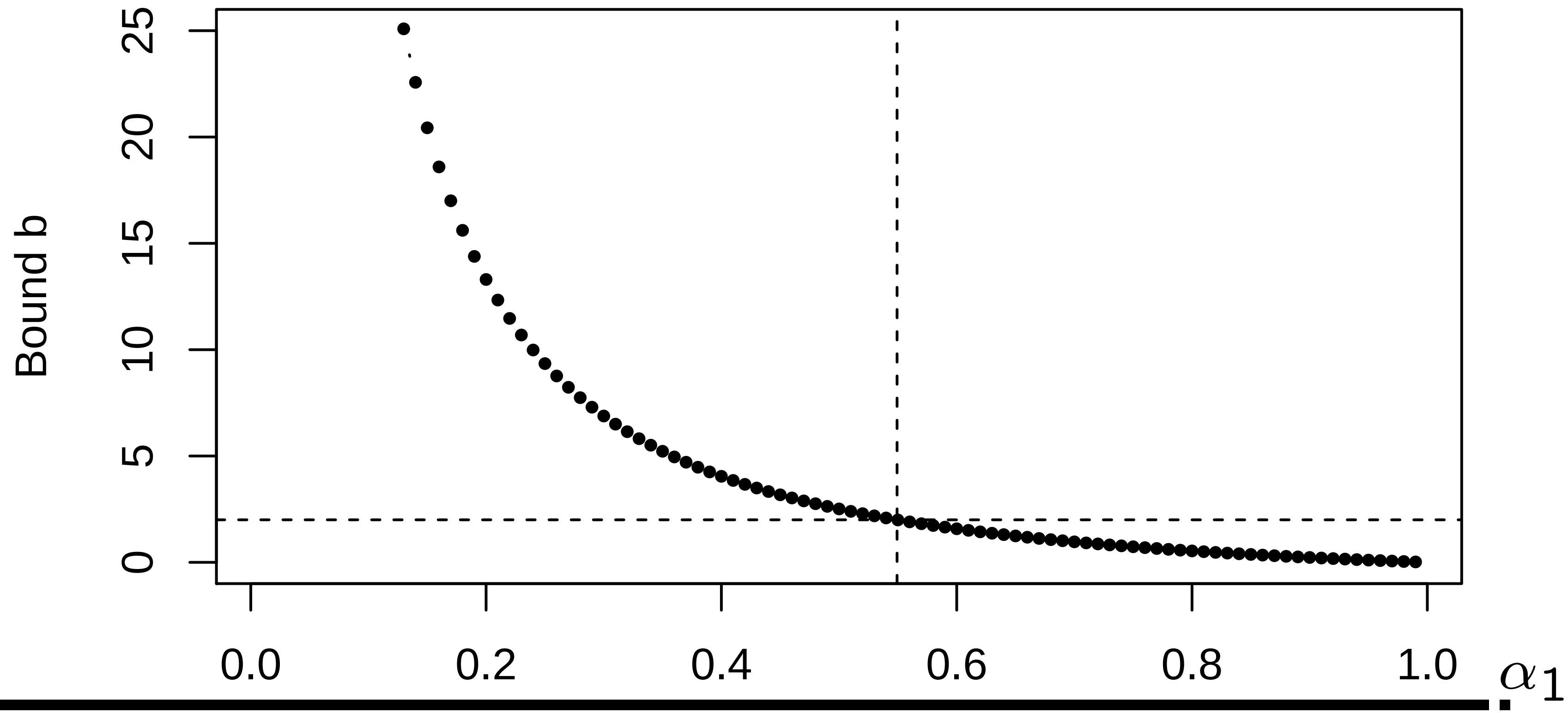
then solve $\mathbf{P}_M\, \boldsymbol{p} = \boldsymbol{p}$ (invariance equation) in $\boldsymbol{p}$. Then,

$$q_{0,l} \;\approx\; \sum_{r=0}^{M} \frac{1}{(\alpha_0 + \alpha_1\, r)^l} \cdot p_r, \qquad M \text{ sufficiently large.}$$

Analytic solution to inverse moments $q_{0,l}$:

Necessary condition for existence of mgf:

If mgf's radius of convergence equals $b > 0$, then $\alpha_1 < \frac{\ln(b+1)}{b}$.



Analytic solution to $q_{0,l}$ using Adell et al. (1996):

If mgf's radius of convergence satisfies $b > 2$, then

$$q_{0,l} \;=\; \frac{1}{\alpha_1^l} \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}\, E[M^n]\, \sum_{j=0}^{n} \binom{n}{j} \frac{(-1)^j}{\left(\frac{\alpha_0}{\alpha_1} + j\right)^l}.$$

Note that above bound $\alpha_1 < \ln(3)/2 \approx 0.549$ for $b = 2$.

α_0	α_1	l	$q_{0,l}$ by MC	n. s.	$q_{0,l}$ by sum	α_0	α_1	l	$q_{0,l}$ by MC	n. s.	$q_{0,l}$ by sum
2	0.2	1	0.4064081	11	0.4064081	1	0.6	1	0.4973967	48	0.4973967
		2	0.1676993	11	0.1676993			2	0.3046319	52	0.3046320
		3	0.0702093	12	0.0702093			3	0.2212899	55	0.2212899
		4	0.0298009	12	0.0298009			4	0.1815225	57	0.1815225
1.5	0.4	1	0.4299554	18	0.4299554	0.5	0.8	1	0.8060558	100	$2.247 \cdot 10^{26}$
		2	0.1980567	19	0.1980567			2	1.1167550	100	$1.706 \cdot 10^{27}$
		3	0.0972296	20	0.0972296			3	1.9693380	100	$7.069 \cdot 10^{27}$
		4	0.0505194	20	0.0505194			4	3.7735840	100	$2.155 \cdot 10^{28}$

Previous asymptotic normal distribution for $\widehat{C}_{1;\,r}$ allows to apply test if **parameters** α_0, α_1 **specified**.

In practice, we want to test H_0 with $\widehat{C}_{1;\,r}$ **in the presence of estimated parameters**, $\hat{\alpha}_0$ and $\hat{\alpha}_1$.

Idea: first-order Taylor approximation of

$$\widehat{C}_{1;\,r}(\alpha_0, \alpha_1) \;=\; \frac{1}{T-1}\,\sum_{t=2}^{T} \frac{(X_t)_{(r)}}{(\alpha_0 + \alpha_1\, X_{t-1})^r},$$

which has partial derivatives

$$\tfrac{\partial}{\partial \alpha_0}\,\widehat{C}_{1;\,r} \;=\; \tfrac{1}{T-1}\,\sum_{t=2}^{T} \tfrac{-r\,(X_t)_{(r)}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}, \qquad \tfrac{\partial}{\partial \alpha_1}\,\widehat{C}_{1;\,r} \;=\; \tfrac{1}{T-1}\,\sum_{t=2}^{T} \tfrac{-r\,(X_t)_{(r)}\,X_{t-1}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}.$$

Testing CP-INARCH Models

By conditioning, it follows that

$$E\Big[\frac{(X_t)_{(r)}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}\Big] \;=\; q_{0,1}, \qquad E\Big[\frac{(X_t)_{(r)}\,X_{t-1}}{(\alpha_0+\alpha_1\,X_{t-1})^{r+1}}\Big] \;=\; q_{1,1}.$$

So we approximate $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$ by

$$\widetilde{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1) \;:=\; \widehat{C}_{1;\,r}(\alpha_0,\alpha_1) \;-\; r\,q_{0,1}\,\big(\hat{\alpha}_0-\alpha_0\big) \;-\; r\,q_{1,1}\,\big(\hat{\alpha}_1-\alpha_1\big).$$

Approximation of distribution of $\widehat{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$
by distribution of $\widetilde{C}_{1;\,r}(\hat{\alpha}_0,\hat{\alpha}_1)$.

Example: use moment estimators $\widehat{\alpha}_0 := \bar{X}\big(1-\widehat{\rho}(1)\big)$ and $\widehat{\alpha}_1 := \widehat{\rho}(1)$, see Weiß & Schweer (2016) for asymptotics.

Then

$$E\big[\widehat{C}_{1;\,r}(\widehat{\alpha}_0,\widehat{\alpha}_1)\big] \;\approx\; 1 \;-\; r\,\tfrac{q_{0,1}}{T-1}\,\left(\tfrac{1+3\alpha_1}{1-\alpha_1}\,\alpha_0+\tfrac{2\alpha_1^2(1+2\alpha_1^2)}{1-\alpha_1^3}\right)$$
$$+\;r\,\tfrac{q_{1,1}}{T-1}\,\left(1+3\alpha_1+\tfrac{\alpha_1}{\alpha_0}\left(1+\tfrac{2\alpha_1(1+2\alpha_1^2)}{1+\alpha_1+\alpha_1^2}\right)\right),$$

and approximate variance $\sigma_{1;\,r}^2/(T-1)$ with

$$\sigma_{1;\,r}^2 \;=\; \textstyle\sum_{k=1}^r\,\binom{r}{k}^2 k!\,q_{0,k} \;-\; 2r^2\,q_{0,1} \;+\; r^2\,q_{0,1}^2\,\tfrac{\alpha_0}{1-\alpha_1}\big(\alpha_0(1+\alpha_1)+\tfrac{1+2\alpha_1^4}{1+\alpha_1+\alpha_1^2}\big)$$
$$+\;r^2\,q_{1,1}^2\,(1-\alpha_1^2)\big(1+\tfrac{\alpha_1(1+2\alpha_1^2)}{\alpha_0(1+\alpha_1+\alpha_1^2)}\big) \;-\; 2r^2\,q_{0,1}q_{1,1}\,\big(\alpha_0(1+\alpha_1)+\tfrac{(1+2\alpha_1)\alpha_1^3}{1+\alpha_1+\alpha_1^2}\big).$$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Finite-Sample Performance of $\widehat{C}_{1;\,r}$-Test with Estimated Parameters

Simulation Study

Simulations with 10 000 replications.

Properties under H_0 (Poi-INARCH(1) model with $\mu = 2.5$):

			$E[\widehat{C}_{1;2}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;2}(\cdot)]$		$E[\widehat{C}_{1;3}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;3}(\cdot)]$		$E[\widehat{C}_{1;4}(\cdot)]$		$\sqrt{V}[\widehat{C}_{1;4}(\cdot)]^{1/2}$	
α_0	α_1	T	appr	simul	appr	simul	appr	simul	appr	simul	appr	simul	appr	simul
2	0.2	100	0.999	0.992	0.058	0.060	0.999	0.976	0.186	0.184	0.998	0.953	0.444	0.426
		250	1.000	0.997	0.037	0.037	1.000	0.993	0.118	0.115	0.999	0.985	0.280	0.273
		500	1.000	0.999	0.026	0.026	1.000	0.995	0.083	0.082	1.000	0.991	0.198	0.195
		1000	1.000	0.999	0.018	0.018	1.000	0.997	0.059	0.058	1.000	0.995	0.140	0.139
1.5	0.4	100	0.996	0.990	0.064	0.066	0.994	0.981	0.206	0.207	0.992	0.964	0.501	0.504
		250	0.998	0.997	0.041	0.041	0.997	0.992	0.130	0.130	0.997	0.985	0.316	0.313
		500	0.999	0.998	0.029	0.029	0.999	0.995	0.092	0.093	0.998	0.990	0.223	0.225
		1000	1.000	0.999	0.020	0.020	0.999	0.998	0.065	0.066	0.999	0.998	0.158	0.161
1	0.6	100	0.982	0.987	0.088	0.091	0.973	0.972	0.269	0.277	0.964	0.946	0.698	0.685
		250	0.993	0.996	0.056	0.057	0.989	0.994	0.170	0.174	0.986	0.994	0.440	0.464
		500	0.996	0.998	0.039	0.040	0.995	0.994	0.120	0.120	0.993	0.989	0.311	0.310
		1000	0.998	0.999	0.028	0.028	0.997	0.997	0.085	0.086	0.996	0.996	0.220	0.234
0.5	0.8	100	0.876	0.958	0.219	0.180	0.813	0.942	0.616	0.840	0.751	0.926	1.933	3.401
		250	0.951	0.984	0.138	0.136	0.926	0.977	0.388	0.418	0.901	0.967	1.219	1.472
		500	0.975	0.992	0.098	0.098	0.963	0.995	0.274	0.320	0.951	1.010	0.861	1.311
		1000	0.988	0.995	0.069	0.069	0.982	0.994	0.194	0.197	0.975	0.996	0.609	0.643

Simulations with 10 000 replications (5 % level).

Properties under H_1 (NB-INARCH(1) model with $\mu = 2.5$):

			$\widehat{C}_{1;2}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;3}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$				$\widehat{C}_{1;4}(\hat{\alpha}_0, \hat{\alpha}_1)$; $\theta =$			
α_0	α_1	T	1	1.2	1.4	1.6	1	1.2	1.4	1.6	1	1.2	1.4	1.6
2	0.2	100	0.051	0.354	0.720	0.901	0.051	0.328	0.667	0.874	0.051	0.272	0.561	0.786
		250	0.049	0.636	0.966	0.999	0.053	0.581	0.947	0.997	0.056	0.478	0.878	0.985
		500	0.051	0.874	0.999	1.000	0.052	0.829	0.998	1.000	0.057	0.717	0.988	1.000
		1000	0.049	0.989	1.000	1.000	0.055	0.975	1.000	1.000	0.056	0.927	1.000	1.000
1.5	0.4	100	0.053	0.337	0.691	0.893	0.054	0.305	0.644	0.855	0.049	0.244	0.532	0.755
		250	0.053	0.608	0.956	0.999	0.055	0.561	0.929	0.996	0.053	0.448	0.848	0.979
		500	0.051	0.848	0.999	1.000	0.058	0.805	0.997	1.000	0.060	0.680	0.983	1.000
		1000	0.052	0.984	1.000	1.000	0.060	0.969	1.000	1.000	0.061	0.905	1.000	1.000
1	0.6	100	0.063	0.305	0.617	0.830	0.050	0.266	0.579	0.805	0.036	0.193	0.448	0.668
		250	0.061	0.522	0.910	0.991	0.060	0.487	0.888	0.987	0.047	0.353	0.760	0.942
		500	0.059	0.748	0.993	1.000	0.057	0.722	0.989	1.000	0.048	0.555	0.944	0.997
		1000	0.056	0.944	1.000	1.000	0.060	0.932	1.000	1.000	0.058	0.803	0.999	1.000
0.5	0.8	100	0.054	0.206	0.443	0.642	0.040	0.189	0.412	0.617	0.019	0.111	0.277	0.456
		250	0.056	0.336	0.695	0.891	0.052	0.325	0.696	0.894	0.025	0.186	0.493	0.745
		500	0.057	0.522	0.891	0.985	0.062	0.503	0.896	0.989	0.032	0.294	0.696	0.923
		1000	0.056	0.729	0.990	1.000	0.061	0.735	0.993	1.000	0.042	0.457	0.917	0.996

Conclusions

- For distinction between Poi-INARCH and true CP-INARCH, we proposed test based on pgf of compounding distribution.
- Asymptotic normality of test statistics either in case, where model parameters specified, or where such parameters estimated.
- Asymptotics involve inverse moments, existence and two methods for calculation.
- Recommendation for practice: use second-order test $\widehat{C}_{1;\,2}(\hat{\alpha}_0, \hat{\alpha}_1)$.

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

Adell et al. (1996) On the Cheney and Sharma... *JMAA* 200, 663–679.

Ferland et al. (2006) Integer-valued GARCH processes. *JTSA* 27, 923–942.

Gonçalves et al. (2015) Infinitely divisible distributions in integer-valued GARCH models. *JTSA* 36, 503–527.

Heinen (2003) Modelling time series count data... *CORE Disc.P.* 2003-63.

Ibragimov (1962) Some limit theorems for stationary processes. *Theory Probab. Appl.* 7(4), 349–382.

Neumann (2011) Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* 17(4), 1268–1284.

Weiß & Schweer (2016) Bias corrections for moment estimators in Poisson INAR(1) and INARCH(1) processes. *SPL* 112, 124–130.

Xu et al. (2012) A model for integer-valued time series with conditional overdispersion. *CSDA* 56(12), 4229–4242.

Zhu & Wang (2010) Diagnostic checking integer-valued ARCH(p) models using conditional residual autocorrelations. *CSDA* 54(2), 496–508.