Mixed thinning INAR(1) model

Ana D. Janjić, Miroslav M. Ristić, Aleksandar S. Nastić

Faculty of Science and Mathematics, University of Niš, Serbia

June 9, 2014
Introduction

2 Thinning operators and INAR(1) models based on them
   - Binomial thinning operator
   - Generalized thinning operator

3 Mixed thinning INAR(1) model
   - Motivation
   - The mixed thinning operator
   - The mixed model

4 Application

5 References
Integer-valued time series

Fields of usage of counting processes

- meteorology (earthquakes counting),
- insurance theory (counts of accidents),
- communications (transmitted messages),
- medicine (number of patients),
- law and social sciences (crime victimizations) and so on.

History of development of integer-valued time series

- Model based on Markov chains (Cox and Miller (1965))
- MTD models (Raftery (1985a))
- DARMA models (Jacobs and Lewis (1978a,b,c))
First ordered integer-valued autoregressive model (INAR(1))

Defined by the recursion

\[ X_t = \sum_{i=1}^{X_{t-1}} B_i(t) + \epsilon_t, \quad t \in \mathbb{Z}, \]

with demands:

- \( \{B_i(t)\} \) and \( \{\epsilon_t\} \) are integer-valued,
- \( \{B_i(t)\} \) is i.i.d. sequence independent of \( X_{t-1} \) and \( \epsilon_t \),
- \( \{\epsilon_t\} \) is i.i.d. sequence independent of \( X_{t-i} \), for \( i \geq 1 \).

Note that \( X_{t-1} = 0 \Rightarrow X_t = \epsilon_t. \)
Thinning operators

Probabilistic operations that can be applied to random counts.

- **Purpose**: *shrinking* the observed population
- **Method**: *randomly deletes* some members of the population

Many different types of thinning operators, refer to a survey of Weiss (2008).
Let $X$ be integer-valued r.v. and $\alpha \in [0, 1]$. Define a random variable

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

where $\{Y_i\}$ are i.i.d. Bernoulli indicators with parameter $\alpha$ (called *counting series*), which are independent of $X$.

We say: $\alpha \circ X$ arises from $X$ by *binomial thinning*, and "$\circ$" is the *binomial thinning operator*.

- $\alpha \circ X | (X = x) : \text{Bin}(x, \alpha)$
- $\alpha \circ X \leq X$

$\Rightarrow$ The term is entirely justified.
Interpretation of $\alpha \circ X$

- Observe the population of size $X$ at certain time $t$.

- At next time point $t + 1$ the population may be shrinked, because some of the elements have left between time points $t$ and $t + 1$.

- Assume that elements under the study leave independently of each other with probability $1 - \alpha$.

$\Rightarrow$ Size of the observed population at time point $t + 1$ is $\alpha \circ X$. 
Some properties of binomial thinning

\[
\begin{align*}
\alpha \circ (X + Y) & \overset{d}{=} \alpha \circ X + \alpha \circ Y, \text{ for independent r. v. } X, Y \\
\alpha \circ (\beta \circ X) & \overset{d}{=} (\alpha\beta) \circ X \\
1 \circ X & \overset{wp^1}{=} X \\
0 \circ X & \overset{wp^1}{=} 0 \\
E[\alpha \circ X] & = \alpha E[X] \\
Var[\alpha \circ X] & = \alpha^2 Var[X] + \alpha(1 - \alpha)E[X]
\end{align*}
\]
INAR(1) model based on binomial thinning

Let $\alpha \in (0, 1)$. The model is defined by the recursion

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z},$$

with demands:

- Thinning operations are performed independently of each other and of $\{\epsilon_t\}$
- At each time $t$ thinning operations at that time and $\epsilon_t$ are independent of $\{X_s\}_{s<t}$

Special case: geometric marginals (GINAR(1) introduced by Alzaid and Al-Osh (1988))

Ana D. Janjić, Miroslav M. Ristić, Aleksandar S. Nastić
Interpretations

Basic interpretation

- $X_t$ - size of the population at time $t$
- $\alpha \circ X_{t-1}$ - survivors of time $t - 1$
- $\epsilon_t$ - immigration

An alternative interpretation

- $X_t$ - customers at time $t$
- $\epsilon_t$ - new customers arrived between time points $t - 1$ and $t$
- $X_{t-1} - \alpha \circ X_{t-1}$ - customers that have been lost between time points $t - 1$ and $t$
Some generalizations of binomial thinning

Obtained by relaxing conditions specified in the definition of binomial thinning.

- counting variables have full range $\mathbb{N}_0$ - *generalized thinning*
  - special case: *negative binomial thinning*

- negative integers are included - *signed thinning*

- *random coefficient thinning*

- dependent Bernoulli indicators
Model based on negative binomial thinning

Defined by the recursion

\[ X_t = \alpha \ast X_{t-1} + \epsilon_t, \quad \text{where} \quad \alpha \ast X = \sum_{i=1}^{X} Y_i, \quad \text{for} \quad Y_i : Geom \left( \frac{\alpha}{1 + \alpha} \right). \]

- Operator "\( \ast \)" is not actually a "thinning", because \( \alpha \ast X \leq X \) is not always true.
- Special case: geometric marginals (NGINAR(1) introduced by Ristić et al. (2009))
The main differences are
- \(0 \circ X \overset{w.p.}{=} 0\) and \(0 \ast X \overset{w.p.}{=} 0\), but
- \(1 \circ X \overset{w.p.}{=} X\), while \(1 \ast X \overset{d}{=} \begin{cases} 0, & \text{w.p. } \frac{1}{1+\mu} \\ X, & \text{w.p. } \frac{\mu^2}{(1+\mu)^2} \\ X + Y, & \text{w.p. } \frac{\mu^2}{(1+\mu)^2} \end{cases}\)

where \(Y\) is geometric \(\left(\frac{1+\mu}{2+\mu}\right)\) independent of \(X\).
- \(\beta \circ (\alpha \circ X) = (\beta \alpha) \circ X\), where counting sequences of "\(\alpha \circ \)" and "\(\beta \circ \)" are independent, unfortunately

\[\beta \ast (\alpha \ast X) \overset{d}{=} \begin{cases} 0, & \text{w.p. } \frac{1+\alpha}{1+\alpha+\alpha\mu} \\ (\beta \alpha) \ast X + Y_1, & \text{w.p. } \frac{\alpha^2\mu^2}{(1+\alpha+\alpha\mu)(1+\alpha\mu)} \\ (\beta \alpha) \ast X + Y_2, & \text{w.p. } \frac{\alpha^2\mu^2}{(1+\alpha+\alpha\mu)(1+\alpha\mu)} + \frac{\beta\alpha}{1+\beta\alpha} \end{cases}\]

\(Y_1\) and \(Y_2\) are independent and geometrically distributed with parameters \(\frac{\beta\alpha}{1+\beta\alpha}\) and \(\beta(1+\alpha+\alpha\mu) \cdot \frac{\beta\alpha}{1+\beta(1+\alpha+\alpha\mu)}\), respectively and are independent of \(X\).
- \(E(\alpha \circ X)^2 = \alpha^2E(X^2) + \alpha(1-\alpha)E(X)\), similarly
- \(E(\alpha \ast X)^2 = \alpha^2E(X^2) + \alpha(1+\alpha)E(X)\).
Model based on dependent Bernoulli counting series

1. Generate a sequence of dependent Bernoulli r. v. as

\[ U_i = (1 - V_i)W_i + V_iZ, \]

where

- \{W_i\} is i.i.d. with \( Ber(\alpha) \) distribution,
- \{V_i\} is i.i.d. with \( Ber(\theta) \) distribution,
- \( Z : Ber(\alpha) \).

\[ U_1 + U_2 + \ldots + U_n \overset{d}{=} \begin{cases} 
    Bin(n, \alpha(1 - \theta)), & \text{w.p. } 1 - \alpha \\
    Bin(n, \alpha + \theta - \alpha\theta), & \text{w.p. } \alpha
\end{cases} \]

2. Define thinning operator as \( \alpha \circ \theta X = \sum_{i=1}^{X} U_i \).

3. Obtained model

\[ X_t = \alpha \circ \theta X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}. \]

Introduced by Ristić et al. (2013)
Models based on binomial thinning

- elements can enter/survive/leave (contribution to the overall thinning sum 0 or 1)

Models based on negative binomial thinning

- elements by replicating themselves contribute to the overall thinning sum more than 1

By mixing we could deal with elements which are active in some period and passive in another

Applications:

- the number of patients with certain transmitting disease
- the number of crimes in some police district
- the number of bacteria
Construction of the mixed thinning operator

Let \( \{W_i\}_{i \in N} \) be i.i.d. sequence, defined as

\[
W_i = \begin{cases} 
B_i, & \text{w.p. } p, \\
G_i, & \text{w.p. } 1 - p,
\end{cases}
\quad p \in [0, 1], \quad B_i : Ber(\alpha), \quad G_i : Geom\left(\frac{\alpha}{1 + \alpha}\right).
\]

The new thinning operator is

\[
\alpha \circ_p X = \sum_{i=0}^{X} W_i,
\]

where \( W_0 = 0 \), \( X \) is nonnegative integer-valued r.v. independent of the counting series \( \{W_i\}_{i \in N} \).
Basic properties of the mixed thinning operator

Using p.g.f. of $W_i$, we obtain

$$\alpha \bullet_p X \mid \{X = x\} \overset{d}{=} \begin{cases} 
NB \left( x, \frac{\alpha}{1+\alpha} \right) & \text{w.p. } (1 - p)^x \\
Bin(i, \alpha) + NB \left( x - i, \frac{\alpha}{1+\alpha} \right) & \text{w.p. } \binom{x}{i} p^i (1 - p)^{x-i} \\
Bin(x, \alpha) & \text{w.p. } p^x
\end{cases}$$

for $1 \leq i \leq x - 1$.

Let $J : Bin(x, p)$. Then

$$\alpha \bullet_p X \mid \{X = x\} \overset{d}{=} Bin(J, \alpha) + NB \left( x - J, \frac{\alpha}{1+\alpha} \right) \overset{d}{=} \alpha \circ J + \alpha \ast (x - J).$$

Also,

$$E[\alpha \bullet_p X] = \alpha E[X]$$

$$Var[\alpha \bullet_p X] = \alpha^2 Var[X] + \alpha (1 + \alpha - 2\alpha p) E[X]$$
Construction of the mixed model with geometric marginals (MixGINAR(1))

MixINAR(1) model is defined by recursion

\[ X_t = \alpha_p X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z} \quad (1) \]

with demands:
- \( \{\varepsilon_t\}_{t \in \mathbb{Z}} \) is a sequence of i.i.d. r. v. independent of the counting series \( \{W_i\}_{i \in \mathbb{N}} \),
- r. v. \( X_{t-i} \) and \( \varepsilon_t \) are independent for all \( i \geq 1 \).

The mixed model (1) with geometric marginals (MixGINAR(1)) contains two existing models as special cases:
- for \( p = 0 \) \( \Rightarrow \) MixGINAR(1) \( \equiv \) NGINAR(1),
- for \( p = 1 \) \( \Rightarrow \) MixGINAR(1) \( \equiv \) GINAR(1).
Using the p.g.f. of innovation r.v. we obtain the next result:

Let $X_t : \text{Geom} \left( \frac{\mu}{1 + \mu} \right)$ for $t \in \mathbb{Z}$ and $\mu \geq \alpha(1 - \alpha p)/(1 - \alpha)$. Then

$$
\varepsilon_t \overset{d}{=} \begin{cases} 
0, & \text{with probability } \alpha p, \\
\text{Geom} \left( \frac{\alpha}{1 + \alpha} \right), & \text{with probability } \frac{\alpha \mu (1 - p)}{\mu - \alpha}, \\
\text{Geom} \left( \frac{\mu}{1 + \mu} \right), & \text{with probability } \frac{\mu - \alpha (1 + \mu - \alpha p)}{\mu - \alpha}.
\end{cases}
$$
The unknown parameters $\alpha$, $\mu$ and $p$ need to be estimated. Since the conditional expectation $E(X_t|X_{t-1}) = \alpha X_{t-1} + (1 - \alpha) \mu$ only depends on the first two parameters $\alpha$ and $\mu$, we will use the two-step conditional least squares approach considered by Karlesen and Tjøstheim (1986).

- **Step one**: estimation of the unknown parameters $\alpha$ and $\mu$,
- **Step two**: estimation of the unknown parameter $p$ using the conditional least squares estimates of the parameters $\alpha$ and $\mu$ obtained in the first step.
Obtained conditional least squares estimates

\[ \hat{\alpha}_{\text{cls}} = \frac{(n - 1)^{-1} \sum_{t=2}^{n-1} X_t X_{t-1} - (n - 1)^{-2} \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n - 1)^{-1} \sum_{t=2}^{n} X_t^2 - (n - 1)^{-2} (\sum_{t=2}^{n} X_{t-1})^2} \]

\[ \hat{\mu}_{\text{cls}} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{\text{cls}} \sum_{t=2}^{n} X_{t-1}}{(n - 1)(1 - \hat{\alpha}_{\text{cls}})} \]

\[ \hat{p}_{\text{cls}} = \frac{\sum_{t=2}^{n} \hat{Z}_t (X_{t-1} - \hat{\mu}_{\text{cls}})}{2\hat{\alpha}_{\text{cls}}^2 \sum_{t=2}^{n} (X_{t-1} - \hat{\mu}_{\text{cls}})^2}, \]

where

\[ \hat{Z}_t = -\hat{Y}_t + \hat{\alpha}_{\text{cls}} (1 + \hat{\alpha}_{\text{cls}}) X_{t-1} + \hat{\mu}_{\text{cls}} (1 - \hat{\alpha}_{\text{cls}} - 2\hat{\alpha}_{\text{cls}}^2 + \hat{\mu}_{\text{cls}} - \hat{\alpha}_{\text{cls}}^2 \hat{\mu}_{\text{cls}}) \]

\[ \hat{Y}_t = X_t - \hat{\alpha}_{\text{cls}} X_{t-1} - (1 - \hat{\alpha}_{\text{cls}}) \hat{\mu}_{\text{cls}}. \]
We will compare GINAR(1), NGINAR(1) and MixGINAR(1).

Consider data series representing a monthly counting of committing a light criminal activity, public drunkenness, in a period from January 1990 to December 2001, constituting sequence of 144 observations.

Series PubDrunk-22 is created by the 22nd police car beat of Pittsburgh and can be downloaded from a website Forecasting Principles (http://www.forecastingprinciples.com).

The sample mean, variance and autocorrelation of the PubDrunk-22 are respectively, 1.34, 10.1 and 0.761.
### Mixed thinning INAR(1) model

<table>
<thead>
<tr>
<th>Model</th>
<th>MLE</th>
<th>MLV</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GINAR(1)</td>
<td>$\hat{q} = 0.6030$</td>
<td>$\hat{\alpha} = 0.4800$</td>
<td>188.099</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGINAR(1)</td>
<td>$\hat{\mu} = 1.3729$</td>
<td>$\hat{\alpha} = 0.5720$</td>
<td>178.010</td>
</tr>
<tr>
<td>MTGINAR(1)</td>
<td>$\hat{\mu} = 1.4544$</td>
<td>$\hat{\rho} = 0.2234$</td>
<td>$\hat{\alpha} = 0.6167$</td>
</tr>
</tbody>
</table>

MLE- maximum likelihood parameter estimates  
MLV - maximum log-likelihood values  
RMS - the root mean squares of differences between the observations and predicted values
References


Thank You for Your Attention