Binomial Autoregressive Processes with Density-Dependent Thinning

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

Philip K. Pollett
Department of Mathematics, University of Queensland
Time Series of Counts with a Finite Range

Motivation & Outline
During last 30 years, increasing research interest in count data time series. Initially, research focussed nearly exclusively on counts with unlimited range $\mathbb{N}_0$. Only during last years, considerable interest also in time series with fixed finite range $\{0, 1, \ldots, N\}$.

Examples:
- utilization of computer pools (with $N$ workstations);
- spread of metapopulations (with $N$ patches);
- \ldots

Christian H. Weiß — Helmut Schmidt University, Hamburg
Data example 1: (Weiß & Kim, 2013)

Number of securities companies (among $N = 22$ companies) traded in Korea stock market per 5-min period (Feb. 8, 2011, 09:00–14:50; $T = 70$).
Data example 2: (Robert-Koch-Institut, survstat.rki.de)

Number of districts in Germany (among \( N = 38 \) districts) with new case of hantavirus infection (weekly data, 2011; \( T = 52 \)).
Modeling of time series with fixed finite range \( \{0, 1, \ldots, N\} \)?

Popular **basic approach** (counterpart to AR(1) model): *binomial AR(1) model* (BAR(1)) by McKenzie (1985), which uses **binomial thinning** operation

\[
\alpha \circ X \sim \text{Bin}(X, \alpha)
\]


Parameters \( \pi \in (0; 1) \), \( \rho \in \left( \max \left\{ -\frac{\pi}{1-\pi}, -\frac{1-\pi}{\pi} \right\}; 1 \right) \), define thinning probabilities \( \beta := \pi (1 - \rho) \) and \( \alpha := \beta + \rho \).

BAR(1) recursion

\[
X_{t+1} = \frac{\alpha \circ X_t}{\text{survivors}} + \frac{\beta \circ (N - X_t)}{\text{newly occupied}},
\]

thinnings performed independently, independent of \((X_s)_{s \leq t}\).

Christian H. Weiß — Helmut Schmidt University, Hamburg
A few well-known properties:

Ergodic Markov chain, transition probabilities

\[ P(X_{t+1} = k \mid X_t = l) = \sum_{m=\max\{0, k+l-N\}}^{\min\{k,l\}} \binom{l}{m} \binom{N-l}{k-m} \alpha^m (1-\alpha)^{l-m} \beta^{k-m} (1-\beta)^{N-l+m-k}, \]

uniquely determined stationary distribution: Bin\((N, \pi)\).

Autocorrelation function: \( \rho_X(k) = \rho^k \) for \( k \geq 0 \).

Regression properties:

\[ E[X_{t+1} \mid X_t] = \rho \cdot X_t + N\beta, \]
\[ V[X_{t+1} \mid X_t] = \rho(1-\rho)(1-2\pi) \cdot X_t + N\beta(1-\beta). \]

Christian H. Weiß — Helmut Schmidt University, Hamburg
Limitations:

- Binomial marginal distribution,
- exponentially decaying ACF (AR(1)-like),
- thinning probabilities at time $t$
do not depend on process up to that time, ... 

Approaches for generalization:

- State dependence of parameters $\pi, \rho$ resp. $\alpha, \beta$ at time $t + 1$
  - as function of $X_t/N$: “density dependence”;
    (e.g., proportion of infectives or proportion of occupied patches)
  - acc. to $X_t$ in which regime: “SET approach” (Tong).
Density-Dependent Binomial AR(1) Processes

Definition & Properties
**Definition:** (Weiβ & Pollett, 2014)

Let $\pi : [0; 1] \rightarrow (0; 1)$ and $\rho : [0; 1] \rightarrow (0; 1)$, so functions $\beta(y) := \pi(y) \left(1 - \rho(y)\right)$ and $\alpha(y) := \beta(y) + \rho(y)$ also range $(0; 1)$.

Write $\pi_{t+1} := \pi(X_t/N)$, $\rho_{t+1} := \rho(X_t/N)$,

and $\alpha_{t+1} := \alpha(X_t/N)$, $\beta_{t+1} := \beta(X_t/N)$.

DD-BAR(1) recursion

$$X_{t+1} = \alpha_{t+1} \circ X_t + \beta_{t+1} \circ (N - X_t),$$

thinnings performed independently, independent of $(X_s)_{s \leq t}$. 

Christian H. Weiβ — Helmut Schmidt University, Hamburg
Density-Dependent Binomial AR(1) Processes

General properties: (Weiß & Pollett, 2014)

- time-homogeneous finite-state Markov chain with truly positive transition probabilities

\[
P(k|l) := P(X_t = k \mid X_{t-1} = l) = \sum_{m=\max\{0,k+l-N\}}^{\min\{k,l\}} \binom{l}{m} \binom{N-l}{k-m} (\alpha(l/N))^m (1 - \alpha(l/N))^{l-m} (\beta(l/N))^{k-m} (1 - \beta(l/N))^{N-l+m-k};
\]

- regression properties

\[
E[X_t \mid X_{t-1}] = \rho_t X_{t-1} + N \beta_t,
\]
\[
V[X_t \mid X_{t-1}] = \rho_t (1 - \rho_t) (1 - 2\pi_t) X_{t-1} + N \beta_t (1 - \beta_t);
\]

- ...
General properties: (cont.)

- due to primitive transition matrix $\mathbf{P} = (P(k|l)_{k,l=0,...,N})$, ergodic with unique stationary distribution, obtained numerically from equation
  $$\mathbf{P} \mathbf{p} = \mathbf{p}$$ (eigenvalue problem)

(also allows to evaluate stationary moments);

- second-order moments with time-lag $h$
  from $P(X_t = k, X_{t-h} = l)$ as entries of $\mathbf{P}^h \text{diag}(\mathbf{p})$, etc.;

- ...
Large–$N$ Approximations: Define

\[ f(x) := \alpha(x) x + \beta(x) (1 - x) = \rho(x) x + \beta(x), \]

\[ \nu(x) := \alpha(x) (1 - \alpha(x)) x + \beta(x) (1 - \beta(x)) (1 - x), \]

so

\[ E[X_t/N \mid X_{t-1}/N] = f(X_{t-1}/N), \quad NV[X_t/N \mid X_{t-1}/N] = \nu(X_{t-1}/N). \]

Several large–$N$ results, among others

- law of large numbers \( X_t/N \xrightarrow{P} y_t \), where \( y_{t+1} = f(y_t) \);
- central limit law for scaled fluctuations \( \sqrt{N} (X_t/N - y_t) \);
- if \( y^* \) stable fixed point of \( f \), and \( \kappa = f'(y^*) \), then

\[
\mu_X \approx N y^*, \quad \sigma^2_X \approx N \frac{\nu(y^*)}{1 - \kappa^2}, \quad \rho_X(k) \approx \kappa^k. 
\]

(Buckley & Pollett, 2010; Weiß & Pollett, 2014)

Christian H. Weiß — Helmut Schmidt University, Hamburg
Density-Dependent Binomial AR(1) Processes

Special Cases
Special Case 1: Binomial Over- or Underdispersion

**Binomial index of dispersion** for r.v. with support \(\{0, \ldots, N\}\), and with mean \(\mu\) and variance \(\sigma^2\):

\[
I_d := \frac{N\sigma^2}{\mu(N - \mu)} \in (0; \infty).
\]

For Bin\((N, \pi)\)-distributed r.v., we have \(I_d = 1\) for any \(\pi \in (0; 1)\). If \(I_d > 1\), then **overdispersion** w.r.t. binomial distribution ("extra-binomial variation"), while **underdispersion** refers to \(I_d < 1\).
Special Case 1: Binomial Over- or Underdispersion

**Definition:** Let \( \rho(y) = \rho \in (0; 1) \) be constant, and \( \pi(y) = a + by \) be linear, where \( a, a + b \in (0; 1) \). So

\[
\beta_t = (1 - \rho) \left( a + b \frac{X_{t-1}}{N} \right), \quad \alpha_t = (1 - \rho) \left( a + b \frac{X_{t-1}}{N} \right) + \rho.
\]

also linear, and depending on sign of \( b \), these probabilities increase or decrease with increasing density.

**Properties:** ACF \( \rho_X(k) = \left( \rho + (1 - \rho)b \right)^k \), and

\[
\mu = \frac{N a}{1 - b}, \quad I_d = \frac{1 - \rho^2}{1 - \frac{\rho^2}{N} - (1 - \frac{1}{N})(\rho + (1 - \rho)b)^2}.
\]

Christian H. Weiß — Helmut Schmidt University, Hamburg
Special Case 1: Binomial Over- or Underdispersion

Properties:

(a) Possible Range of \( b \)

(b) Attainable range of \( b \) depending on \( a \) in (a)
dispersion determined by \( b \) and \( \rho \) in (b).

Christian H. Weiß — Helmut Schmidt University, Hamburg
Special Case 1: Binomial Over- or Underdispersion

Data example 1: securities counts, \( N = 22 \).

Analyzing empirical (partial) autocorrelations, first-order autoregressive dependence structure apparent.

Sample mean and variance are \( \bar{x} \approx 9.529 \) and \( s^2 \approx 4.253 \).

Empirical index of dispersion \( \hat{I}_d = s^2/(\bar{x}(1 - \bar{x}/N)) \approx 0.787 \) indicates slight degree of binomial underdispersion.

ML-fitted model:

\[
\hat{a}_{ML} \approx 0.693, \quad \hat{b}_{ML} \approx -0.619, \quad \hat{\rho}_{ML} \approx 0.630.
\]
Binomial INARCH(1) Model

Boundary case $\rho \to 0$ leads to

$$X_t \overset{D}{=} \text{Bin}(N, a + b X_{t-1}/N).$$

Motivated by analogy to (Poisson) INARCH(1) model (Ferland et al., 2006): **binomial INARCH(1) model** (BINARCH(1)).

**Properties:** ACF $\rho_X(k) = b^k$, and

$$\mu = \frac{Na}{1 - b}, \quad I_d = \frac{1}{1 - (1 - \frac{1}{N})b^2} \in [1; N).$$

Consequently, only overdispersion is possible.
Data example 2: hanta counts, $N = 38$.

Sample mean and variance are $\bar{x} \approx 4.173$ and $s^2 \approx 7.793$.

Empirical index of dispersion $\hat{I}_d \approx 2.098$

indicates considerable degree of binomial overdispersion.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\pi}_{ML}$</th>
<th>$\hat{\rho}_{ML}$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAR(1)</td>
<td>0.115</td>
<td>0.535</td>
<td>222.8</td>
<td>226.7</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.071)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD-BAR(1)</td>
<td>$\hat{a}_{ML}$</td>
<td>$\hat{b}_{ML}$</td>
<td>$\hat{\rho}_{ML}$</td>
<td>AIC</td>
</tr>
<tr>
<td></td>
<td>0.030</td>
<td>0.748</td>
<td>0.000</td>
<td>213.4</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.143)</td>
<td>(0.367)</td>
<td></td>
</tr>
<tr>
<td>BINARCH(1)</td>
<td>$\hat{a}_{ML}$</td>
<td>$\hat{b}_{ML}$</td>
<td>AIC</td>
<td>BIC</td>
</tr>
<tr>
<td></td>
<td>0.030</td>
<td>0.748</td>
<td>211.4</td>
<td>215.3</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.108)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Special Case 2: Density-Dependent Colonization

**Definition:** Let $\alpha(y) = \alpha \in (0; 1)$ be constant, and $\beta(y) = \alpha (a + by)$ be linear, where $a, a + b \in (0; 1]$.

**Epidemic context:** If $b > 0$,

prob. for susceptible becoming infected increases if number of infectives already large (infection is spreading), while recovery from infection independent of other infectives.

**Properties:**

$E[X_t \mid X_{t-1}] = (\alpha - \beta_t) X_{t-1} + N \beta_t$ quadratic in $X_{t-1}$,

$V[X_t \mid X_{t-1}]$ cubic polynomial in $X_{t-1}$.

Explicit large-$N$ approximation for marginal moments and ACF.
Self-Exciting Threshold Binomial AR(1) Processes

(jointly with T.A. Möller, M.E. Silva, M.G. Scotto, I. Pereira)

“Work in Progress”
Idea: Like in Monteiro et al. (2012), fix threshold value $0 \leq R < N$, which separates range into two regimes: lower regime \{0, 1, 2, \ldots, R\}, upper regime \{R + 1, R + 2, \ldots, N\}. Depending if previous observation in lower regime ($X_{t-1} \leq R$) or in upper regime ($X_{t-1} > R$), parameters of BAR(1) recursion at time $t$ chosen differently.

Note: Possible generalizations include

- more than two regimes;
- increased delay $X_{t-d} \leq R$ vs. $X_{t-d} > R$ with $d > 1$;
- more complex criteria, e.g., $\max \{X_{t-1}, \ldots, X_{t-d}\} \leq R$.  

Christian H. Weiß — Helmut Schmidt University, Hamburg
**Definition:** Define $\pi_i \in (0; 1)$, $\rho_i \in \left( \max \left\{ -\frac{\pi_i}{1-\pi_i}, -\frac{1-\pi_i}{\pi_i} \right\}; 1 \right)$, and $\beta_i \equiv \pi_i \cdot (1 - \rho_i) \in (0; 1)$ and $\alpha_i \equiv \beta_i + \rho_i \in (0; 1)$.

**SET-BAR(1) process**

\[
X_t = \phi_t \circ X_{t-1} + \eta_t \circ (N - X_{t-1}),
\]

where $\phi_t := \alpha_1 I_{t-1} + \alpha_2 (1 - I_{t-1})$ and $\eta_t := \beta_1 I_{t-1} + \beta_2 (1 - I_{t-1})$ with indicator $I_{t-1} := 1\{X_{t-1} \leq R\}$.

**Note:** If $R = 0$, then $\alpha_1$ no influence (can be chosen arbitrarily).

$\Rightarrow$ $\alpha_1$ unidentifiable during parameter estimation.

Same issue with $\beta_2$ and $R = N - 1$. Therefore, we set $\rho_1 = \rho_2$ for $R = 0, N - 1$ ($\rightarrow$ LSET model from below).
Properties:

SET-BAR(1) model instance of

**DD-BAR(1) models** by Weiß & Pollett (2014)

⇒ adapt results concerning transition probabilities, ergodicity, existence of unique stationary marginal distribution $p$.

Since finite range, always possible to compute $p$ numerically by solving eigenvalue problem $Pp = p$.

We derived closed-form formulae for mean and variance, but very complex.

Christian H. Weiß — Helmut Schmidt University, Hamburg
Data example 3: (Robert-Koch-Institut)

Weekly No. of districts with new measles case (G., 2004/05).

Finite range \( \{0, \ldots, N\} \) with \( N = 38 \) (number of districts).

But ...
Above measles time series: level shift in first half of 2005, but no obvious change in serial dependence structure.

**Idea:** Reduce number of parameters by additional restriction $\rho_1 = \rho_2 =: \rho$.

**Definition:** Let $\rho \in \left( \max \left\{ -\frac{\pi_1}{1-\pi_1}, -\frac{\pi_2}{1-\pi_2}, -\frac{1-\pi_1}{\pi_1}, -\frac{1-\pi_2}{\pi_2} \right\}; 1 \right)$. A SET-BAR(1) process for which $\rho_1 = \rho_2 =: \rho \neq 0$ holds is called an **LSET-BAR(1) process**.
Defining $p := P(X_t \leq R) = E[I_{t-1}]$, $\mu_{IX} := E[I_{t-1}X_{t-1}]$, we can express unconditional mean and variance as

$$
\mu_X = Np \pi_1 + N(1 - p) \pi_2,
$$

$$
\sigma_X^2 = Np \pi_1(1 - \pi_1) + N(1 - p) \pi_2(1 - \pi_2)
$$

$$
+ N^2 p(1 - p) (\pi_2 - \pi_1)^2
$$

$$
+ \frac{2\rho}{1 + \rho} (N - 1) (\pi_2 - \pi_1) (Np \pi_1 - \mu_{IX}).
$$

LSET model able to show over- and underdispersion w.r.t. binomial distribution for appropriate parameter settings.
Forecasting distributions for model $N = 40$, $\rho = 0.3$, $\pi_1 = 0.15$, $\pi_2 = 0.4$ and $R = 10$, conditional on $X_T = 2$: 
Autocorrelation function for model $N = 40$, $\rho = 0.3$, $\pi_1 = 0.15$, $\pi_2 = 0.4$ and $R = 10$ (black points):

Gray triangles show $f(k) := (\rho X(1))^k$ ⇒ longer memory than corresponding AR(1)-like model.

Christian H. Weiß — Helmut Schmidt University, Hamburg
Data example 3: measles counts, $N = 38$.

<table>
<thead>
<tr>
<th>Model</th>
<th>Par. 1</th>
<th>Par. 2</th>
<th>Par. 3</th>
<th>Par. 4</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAR(1)</td>
<td>0.0882</td>
<td>0.4158</td>
<td></td>
<td>-</td>
<td>448.2</td>
<td>453.5</td>
</tr>
<tr>
<td>$(\pi, \rho)$</td>
<td>(0.0070)</td>
<td>(0.0550)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD-BAR(1)</td>
<td>0.0419</td>
<td>0.5270</td>
<td>0</td>
<td>-</td>
<td>436.9</td>
<td>444.9</td>
</tr>
<tr>
<td>$(a, b, \rho)$</td>
<td>(0.0095)</td>
<td>(0.1077)</td>
<td>(0.1765)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BINARCH(1)</td>
<td>0.0419</td>
<td>0.5270</td>
<td></td>
<td>-</td>
<td>434.9</td>
<td>440.2</td>
</tr>
<tr>
<td>$(a, b)$</td>
<td>(0.0060)</td>
<td>(0.0682)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SET-BAR(1)</td>
<td>0.0706</td>
<td>0.1558</td>
<td>0.1916</td>
<td>0.2904</td>
<td>432.5</td>
<td>443.1</td>
</tr>
<tr>
<td>$(\pi_1, \pi_2, \rho_1, \rho_2)$</td>
<td>(0.0056)</td>
<td>(0.0269)</td>
<td>(0.0884)</td>
<td>(0.375)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSET-BAR(1)</td>
<td>0.0707</td>
<td>0.1604</td>
<td>0.1947</td>
<td>0</td>
<td>430.6</td>
<td>438.5</td>
</tr>
<tr>
<td>$(\pi_1, \pi_2, \rho)$</td>
<td>(0.0057)</td>
<td>(0.0169)</td>
<td>(0.0876)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All threshold models include threshold value $R = 5$. 

Christian H. Weiß — Helmut Schmidt University, Hamburg
Conclusions

- Time series of counts with finite range \(\{0, 1, \ldots, N\}\) is an important topic in practice.

- McKenzie’s basic binomial AR(1) model easily generalized in several ways.

- Density-dependent binomial AR(1) model for binomial over- or underdispersion, boundary case of binomial INARCH(1) model.

- Model with density-dependent colonization.

- SET models for counts exhibiting piecewise-type patterns.

Christian H. Weiß — Helmut Schmidt University, Hamburg
Thank You

for Your Interest!

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de