More on Bayesian Infererence for Correlated Negative-Binomial Processes

RLW

Draft 75: 12:23 April 15, 2013

1 Introduction

For any fixed $\alpha > 0$, $\lambda > 0$, and $p \in (0,1)$ there are several different stochastic processes with negative binomial distributions $Y_t \sim \mathsf{NB}(\alpha, p)$ at every time t and with the same autocorrelation $\mathsf{Corr}[Y_s, Y_t] = e^{-\lambda|s-t|}$; I haven't seen much about any of them in print, and I don't know of any work on inference about them. Steutel, Vervaat and Wolfe (1983) talk about the NP Branching Process of Section (2), and cite Phatarfod and Mardia (1973) for its pgf; apparently it was introduced by Edwards and Gurland (1961). The idea of thinning for general ID distributions (and negative binomial in particular) arises in McKenzie (1985, 1986). Here's a sketch of these stationary NB Markov processes, and an idea of how to find posterior distributions of α, λ, p from data.

2 The Branching NB Model

The negative binomial distribution $\mathsf{NB}(\alpha, p)$ is usually parametrized with pmf

$$\mathsf{P}[Y = k] = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)\, k!} p^\alpha q^k = \binom{-\alpha}{k} p^\alpha (-q)^k, \qquad k \in \mathbb{N}_0 = \{0, 1, 2, \dots\} \tag{1}$$

where $q \equiv 1-p$, but it is often more convenient to use $\beta = p/q \in (0, \infty)$ (replacing "$p^\alpha q^k$" with $\beta^\alpha(1+\beta)^{-\alpha-k}$) or even $\log \beta \in \mathbb{R}$ instead of p. For any $\lambda > 0$ there is precisely one process Y_t with $\mathsf{NB}(\alpha, p)$ marginal distributions that is stationary, Markov, time-reversible, has correlation $\mathsf{Corr}(Y_0, Y_1) = e^{-\lambda}$, and has infinitely-divisible marginal distributions of all order (that's the main theorem in (Wolpert and Brown 2011)). That paper has the generating function for its bivariate distributions, and for its Markov transitions, but only recently did I find a closed-form expression for the conditional pmf for Y_t given $\mathcal{F}_s$, $s \le t$, and so only now do we have a good way to evaluate the likelihood function for an observed data-set $\mathbf{y} = \{y_0, \dots, y_J\}$ of values of some random variables $\mathbf{Y} = \{Y_0, \dots, Y_J\}$ at times $T = \{t_0 \le t_1 \le \cdots \le t_J\} \subset \mathbb{R}$ which we model with this $\mathsf{AR}(1)$-like joint distribution. This is better than the indirect way I had discussed before that used data augmentation.

2.1 Conditional and Joint pmf

Wolpert and Brown (2011, §3.5) offer a recursive update scheme from Y_s to Y_t for the $\mathsf{bNB}(\alpha, p, \lambda)$ process of the form

$$Y_t = \xi + \zeta$$

where ξ and ζ are generated from Y_s by

$$\xi \sim \mathsf{Bi}\Big(Y_s,\ \frac{\rho\ p}{1-\rho+\rho p}\Big) \tag{2a}$$

$$\zeta \sim \mathsf{NB}\Big(\alpha+\xi,\ \frac{p}{1-\rho+\rho p}\Big) \tag{2b}$$

with $\rho \equiv \exp(-\lambda|t-s|)$. This leads to a closed form expression for the joint and conditional pmfs (which should perhaps be added to Wolpert and Brown (2011)):

$$\begin{aligned}
p_{t-s}(j \mid i) &= \mathsf{P}[Y_t = j \mid Y_s = i] \\
&= \sum_{\xi=0}^{i\wedge j} \binom{i}{\xi} \left(\frac{\rho\ p}{1-\rho}\right)^{\xi}\ (1-\rho+\rho p)^{-i}(1-\rho)^{i} \\
&\qquad \times \frac{\Gamma(\alpha+j)}{\Gamma(\alpha+\xi)\ (j-\xi)!} p^{\alpha+\xi}[(1-\rho)(1-p)]^{j-\xi}\ (1-\rho+\rho p)^{-\alpha-j} \\
&= i!\ \Gamma(\alpha+j)\ \frac{p^{\alpha}\ (1-\rho)^{i+j}(1-p)^{j}}{(1-\rho+\rho p)^{\alpha+i+j}} \\
&\qquad \times \sum_{\xi=0}^{i\wedge j} \frac{1}{\xi!\ (i-\xi)!\ \Gamma(\alpha+\xi)\ (j-\xi)!} \left(\frac{\rho \qquad p^2}{(1-\rho)^2(1-p)}\right)^{\xi} \\
&= \frac{\Gamma(\alpha+j)}{\Gamma(\alpha)\ j!}\ \frac{p^{\alpha}\ (1-\rho)^{i+j}\ (1-p)^{j}}{(1-\rho+\rho p)^{\alpha+i+j}}\ {}_2F_1(-i,-j;\alpha;z)
\end{aligned} \tag{3a}$$

where $z = \rho p^2(1-\rho)^{-2}(1-p)^{-1}$ and where ${}_2F_1(a,b;c;z)$ is Gauss' hypergeometric function (Abramowitz and Stegun 1964, §15.1.1, available in `R` as `hyperg_2F1()` in package `gsl`). From this and the one-dimensional marginal $Y_0 \sim \mathsf{NB}(\alpha, p)$ the joint pmf (or likelihood) for Y_T at arbitrary finite sets $T \subset \mathbb{R}$ can be found. For example, the bivariate pmf for $T = (s,t)$ is:

$$\begin{aligned}
p(i,j) &= \mathsf{P}[Y_s = i,\ Y_t = j] \\
&= \frac{\Gamma(\alpha+i)\ \Gamma(\alpha+j)}{\Gamma(\alpha)^2 \quad i! \quad j!}\ \frac{p^{2\alpha}\ (1-\rho)^{i+j}\ (1-p)^{i+j}}{(1-\rho+\rho p)^{\alpha+i+j}}\ {}_2F_1(-i,-j;\alpha;z),
\end{aligned} \tag{3b}$$

exactly $(1-\rho+\rho p)^{\alpha}\ {}_2F_1(-i,-j;\alpha;z)$ times the joint pmf for two independent random variables $Y_s, Y_t \overset{\text{iid}}{\sim} \mathsf{NB}\big(\alpha, p/(1-\rho+\rho p)\big)$.

2.1.1 Likelihood

Fix $\alpha > 0$, $0 < p < 1$, and $\lambda > 0$. Let $T = \{t_0 < t_1 < \cdots < t_J\}$ be an increasing set of times and $\mathbf{y} = \{y_0, y_1, \ldots, y_J\} \subset \mathbb{N}_0$ an arbitrary set of nonnegative integers. Then the joint pmf for $\mathbf{Y} \sim \mathsf{bNB}(\alpha, p, \lambda)$ is the product of the marginal and the conditionals,

$$\mathsf{P}[\mathbf{Y} = \mathbf{y} \mid \alpha, p, \lambda] = p(y_0) \prod_{0<j\leq J} p_j\big(y_j \mid y_{j-1}\big) \tag{4a}$$

where, for $x, y \in \mathbb{N}_0$ and $\rho_j = \exp\big(-\lambda|t_j - t_{j-1}|\big)$, the marginal and transition pmfs are:

$$\begin{aligned} p(x) &= \frac{\Gamma(\alpha + x)}{\Gamma(\alpha)\ x!}\ p^\alpha (1-p)^x \\ p_j(y \mid x) &= \frac{\Gamma(\alpha + y)}{\Gamma(\alpha)\ y!}\ \frac{p^\alpha\ (1-\rho_j)^{x+y}\ (1-p)^y}{(1-\rho_j+\rho_j p)^{\alpha+x+y}}\ {}_2F_1\left(-x, -y; \alpha; \frac{\rho_j}{(1-\rho_j)^2}\ \frac{p^2}{(1-p)}\right). \end{aligned}$$

For equally-spaced times $t_i \equiv t_0 + j\Delta$, each $\rho_j = e^{-\lambda\Delta}$ and this simplifies a bit to

$$\begin{aligned} \mathsf{P}[\mathbf{Y} = \mathbf{y} \mid \alpha, p, \lambda] = \prod_{j=0}^{J} &\left\{\frac{p^\alpha \Gamma(\alpha + y_j)}{\Gamma(\alpha)\ y_j!}\right\} (1-p)^{y_+} (1-\rho+\rho p)^{y_0+y_J-2y_+-J\alpha} \\ &\times \prod_{j=1}^{J} {}_2F_1(-y_{j-1}, -y_j; \alpha; z)\ (1-\rho)^{2y_+-(y_0+y_J)}, \end{aligned} \tag{4b}$$

where $y_+ \equiv \sum_0^J y_j$ and $z = \rho p^2 (1-\rho)^{-2} (1-p)^{-1}$.

For long time series (*i.e.*, those large values of J) with equally-spaced times and comparatively few values (so the set R of distinct values among $\{y_j\}$ is small enough that $|R|^2 \ll J$) it is more efficient to evaluate (using (1) and (3b)) the vector $p(i) = \mathsf{P}\{Y_0 = i\}$ and symmetric matrix $P(i,j) = \mathsf{P}[Y_0 = i, Y_\Delta = j]$ of uni- and bi-variate pmfs for $i, j \in R$, and compute from these the transition matrix $Q(i,j) = \mathsf{P}[Y_{t+\Delta} = j \mid Y_t = i] = P(i,j)/p(i)$. The likelihood function can then be evaluated quickly as

$$\mathsf{P}[\mathbf{Y} = \mathbf{y} \mid \alpha, p, \lambda] = p(y_0) \prod_{j=1}^{J} Q(y_{j-1}, y_j) = \frac{\prod_{0\leq j\leq J} P(y_{j-1}, y_j)}{\prod_{0<j<J} p(y_j)}. \tag{4c}$$

2.1.2 MCMC for Inference

No conjugate distributions exist for this family, but convenient conventional choices are independent Beta distributions for p and ρ and a Gamma distribution for $\log\alpha$. A routine Metropolis-Hastings approach should let us generate an MCMC sample from the joint posterior for α, β, λ using steps:

$$\begin{array}{llcccl}
\alpha: & \text{Gaussian SRW on log scale,} & \alpha_t & \rightsquigarrow & \alpha^* & = \alpha_t e^{Z\delta_\alpha}; \\
p: & \text{Gaussian SRW on logit scale,} & \frac{p_t}{1-p_t} & \rightsquigarrow & \frac{p^*}{1-p^*} & = \frac{p_t}{1-p_t} e^{Z\delta_p}; \\
\lambda: & \text{Gaussian SRW on log scale,} & \lambda_t & \rightsquigarrow & \lambda^* & = \lambda_t e^{Z\delta_\lambda}
\end{array}$$

where the Z's all denote iid $\mathsf{No}(0,1)$ variates and the δ's are step sizes, probably about 0.10. Note that the uncorrelated IID $\mathsf{NB}(\alpha,p)$ model is the limiting case as $\lambda \to \infty$, so high posterior values of λ are evidence against correlation in the model.

2.2 The bNB Generator

Let $Y_t \sim \mathsf{bNB}(\alpha,p,\lambda)$ and let $\epsilon > 0$ be small. Set $q = (1-p)$ and $\rho = 1-\lambda\epsilon = \exp(-\lambda\epsilon)+o(\epsilon)$. Then

$$\begin{aligned}
\mathsf{P}[Y_{t+\epsilon} = j \mid Y_t = i] &= \sum_{k=0}^{i\wedge j} \binom{i}{k} \left(\frac{p-\lambda\epsilon p}{p+\lambda\epsilon q}\right)^k \left(\frac{\lambda\epsilon}{p+\lambda\epsilon q}\right)^{i-k} \\
&\times \qquad \frac{\Gamma(\alpha+j)}{\Gamma(\alpha+k)(j-k)!} \left(\frac{p}{p+\lambda\epsilon q}\right)^{\alpha+k} \left(\frac{\lambda\epsilon q}{p+\lambda\epsilon q}\right)^{j-k} \\
&= o(\epsilon) \qquad \text{if } i+j > 2k+1. \text{ The remaining cases are:} \\
\mathsf{P}[Y_{t+\epsilon} = i+1 \mid Y_t = i] &= \left(\frac{p-\lambda\epsilon p}{p+\lambda\epsilon q}\right)^i (\alpha+i) \left(\frac{p}{p+\lambda\epsilon q}\right)^{\alpha+i} \left(\frac{\lambda\epsilon q}{p+\lambda\epsilon q}\right) \\
&= \lambda(q/p)(\alpha+i)\epsilon + o(\epsilon); \\
\mathsf{P}[Y_{t+\epsilon} = i-1 \mid Y_t = i] &= i \left(\frac{p-\lambda\epsilon p}{p+\lambda\epsilon q}\right)^{i-1} \left(\frac{\lambda\epsilon}{p+\lambda\epsilon q}\right) \left(\frac{p}{p+\lambda\epsilon q}\right)^{\alpha+i-1} \\
&= \lambda(i/p)\epsilon + o(\epsilon) \\
\mathsf{P}[Y_{t+\epsilon} = i \mid Y_t = i] &= \left(\frac{p-\lambda\epsilon p}{p+\lambda\epsilon q}\right)^i \left(\frac{p}{p+\lambda\epsilon q}\right)^{\alpha+i} \\
&= (1-i\lambda\epsilon)\big(1-(\alpha+i)\lambda\epsilon q/p\big) + o(\epsilon) \\
&= 1 - (\lambda/p)[i(1+q)+\alpha q]\epsilon + o(\epsilon)
\end{aligned}$$

and hence the generator for the $Y_t \sim \mathsf{bNB}(\alpha,p,\lambda)$ process is

$$\begin{aligned}
\mathfrak{A}\phi(i) &= \lim_{\epsilon\to 0} \mathsf{E}[\phi(Y_{t+\epsilon}) - \phi(Y_t) \mid Y_t = i]/\epsilon \\
&= [\phi(i+1)-\phi(i)]\lambda(q/p)(\alpha+i) + [\phi(i-1)-\phi(i)]\lambda(i/p) \qquad (5) \\
&= \sum Q_{ij}\ \phi(j), \quad \text{where} \quad Q_{ij} = \begin{cases} (\lambda q/p)(\alpha+i) & j = i+1 \\ -(\lambda/p)[i(1+q)+\alpha q] & j = i \\ (\lambda/p)i & j = i-1 \end{cases}
\end{aligned}$$

In (Wolpert 2011) it is shown that $\epsilon Y_{t/\epsilon}$ converges to a gamma distributed diffusion; the locality of the bNB generator (5) (*i.e.*, its dependence only on $\phi(y \pm 1)$) helps explain this. It also reveals the "branching" nature of the process— which increases from immigration at rate $\lambda_0 \equiv \lambda(q/p)\alpha$ and births at rate $\lambda_+ \equiv \lambda(q/p)$, and decreases by deaths at rate $\lambda_- \equiv \lambda/p$.

3 Regression & Non-stationarity

In some applications (rock-falls, for example) we may expect that some aspect of the process $\{Y_t\}$ (and hence some of the parameters α, p, and λ) may vary over time, or depend on some exogenous explanatory variables $\{X_t\}$. Perhaps the most interesting is probably when α varies as α_t, or depends on a vector $\vec{X}_t$ through a log-linear regression model $\alpha_t = \exp(X_t\gamma)$ for some regression coefficient vector γ. This allows both mean and variance to depend log-linearly on X_t. The only effect this has on the likelihood expressions of (4) is that each appearance of "α" is replaced by an "α_j," the average value of $\alpha_t = \exp(X_t\gamma)$ over the interval $[t_{j-1}, t_j]$. Optimization or (Bayesian) integration over α is replaced with optimization or integration over γ.

4 Other NB Models

The marginal distribution $Y_t \sim \mathsf{NB}(\alpha, p)$ and autocorrelation $\rho_{st} = \exp(-\lambda|t-s|)$ don't characterize the stationary $\mathsf{bNB}(\alpha, p, \lambda)$ process. Here are three other stationary models with those same features. As shown by Wolpert and Brown (2011), however, each of these alternatives must fail either to be infinitely divisible (ID), time reversible (TR), or Markovian.

4.1 The Thinning NB Model

The Thinning Negative Binomial Model $\mathsf{tNB}(\alpha, p, \lambda)$ (Wolpert and Brown 2011, §1.1.2) may be presented recursively at integer times in the form of an initial value $Y_{t_0} \sim \mathsf{NB}\big(\alpha, p\big)$ for some $t_0 \in \mathbb{Z}$ and, at later times $t \in \mathbb{Z}$, an update step:

$$\begin{aligned} &Y_t = \xi_t + \zeta_t, \quad \text{where, for } \rho = \exp(-\lambda), \\ &\xi_t \sim \mathsf{BB}\big(Y_{t-1};\ \rho\alpha,\ (1-\rho)\alpha\big), \quad \zeta_t \sim \mathsf{NB}\big((1-\rho)\alpha,\ p\big) \end{aligned} \tag{6}$$

as the sum of a beta-binomially distributed quantity ξ_t and an independent negative binomially distributed ζ_t. A beta binomial variable $\xi \sim \mathsf{BB}(n; \alpha, \beta)$ may be viewed hierarchically as a Binomial $\xi \sim \mathsf{Bi}(n, \theta)$ for a beta-distributed $\theta \sim \mathsf{Be}(\alpha, \beta)$. It has pmf

$$\mathsf{P}[\xi = k] = \binom{n}{k} \frac{\Gamma(\alpha+\beta)\ \Gamma(\alpha+k)\ \Gamma(\beta+n-k)}{\Gamma(\alpha+\beta+n)\quad \Gamma(\alpha)\quad \Gamma(\beta)}, \quad 0 \le k \le n.$$

This reduces to the uniform $\xi \sim \mathsf{Un}(0, \ldots, n)$ for $\alpha = \beta = 1$, so the special case of $Y_t \sim \mathsf{NB}(2, p)$ with correlation $\rho = \frac{1}{2}$ is particular easy to simulate. Like the Branching Negative Binomial Process of Section (2), the Thinning Negative Binomial Process of Eqn (6) is a stationary time-reversible Markov process with $Y_t \sim \mathsf{NB}(\alpha, p)$ marginal distributions and autocorrelation $\mathsf{Corr}(Y_s, Y_t) = e^{-\lambda|s-t|}$; unlike the earlier process, it does not have infinitely divisible marginals of order 3 or more. Once again one can construct an explicit likelihood function involving gamma and hypergeometric functions (${}_3F_2(\vec{a}, \vec{b}; z)$, this time) to support inference.

4.2 The random measure NB model

The random measure Negative Binomial model $\mathsf{rmNB}(\alpha, p, \lambda)$ (Wolpert and Brown 2011, §1.2.2) may be written in the form

$$Y_t = \mathcal{N}(G_t) \tag{7}$$

for (symmetric in $t \pm x$) sets $G_t \subset \mathbb{R}^2$ of the form

$$G_t = \left\{(x, y) : \ x \in \mathbb{R}, \ 0 \le y < \alpha\lambda e^{-2\lambda|t-x|}\right\}$$

where $\mathcal{N}(dx\,dy)$ assigns independent random variables

$$\mathcal{N}(G) \sim \mathsf{NB}\big(|G|, \ p\big)$$

to disjoint Borel sets $G \subset \mathbb{R}^2$ of finite Lebesgue measure $|G|$. One can show that each $Y_t \sim \mathsf{NB}(\alpha, p)$ with autocorrelation $|G_s \cap G_t| = \exp(-\lambda|s-t|)$. Each component Y_t of Y_T at any finite collection of times $T = \{t_0 < \cdots < t_J\}$ can be written as the sum of some subset of a fixed collection of $(J+1)(J+2)/2$ independent negative binomial random variables, so rmNB is (TR) and (ID), but I don't know any simple closed-form expression for the likelihood function. It follows from Wolpert and Brown (2011) that it can't be Markov.

4.3 The Continuously Thinned NB Model

Wolpert (2009) describes a general method for constructing Markov processes with specified ID univariate marginal distributions by "continuous thinning". For the NB distribution, from (6) with $\rho = (1 - \lambda\epsilon)$ for small $\epsilon > 0$,

$$\begin{aligned}
\mathsf{P}[Y_{t+\epsilon} = j \mid Y_t = i] &= \sum_{k=0}^{i\wedge j} \mathsf{P}[\xi = k, \ \zeta = j - k] \\
&= \sum_{k=0}^{i\wedge j} \left\{ \frac{i! \ \Gamma(\alpha) \ \Gamma(\alpha + k - \alpha\lambda\epsilon) \ \Gamma(i - k + \alpha\lambda\epsilon)}{(i-k)! \ k! \ \Gamma(\alpha + i) \ \Gamma(\alpha - \alpha\lambda\epsilon) \ \Gamma(\alpha\lambda\epsilon)} \right\} \\
&\qquad\qquad \times \left\{ \frac{\Gamma(\alpha\lambda\epsilon + j - k)}{\Gamma(\alpha\lambda\epsilon) \ (j-k)!} \ p^{\alpha\lambda\epsilon}(1-p)^{j-k} \right\} \\
&= \sum_{k=0}^{i\wedge j} S_{ik} \times T_{jk}
\end{aligned}$$

where for $k < i$ and $k < j$, respectively, the summand factors are

$$\begin{aligned}
S_{ik} &= \frac{i!\ \Gamma(\alpha)\ \Gamma(\alpha+k-\alpha\lambda\epsilon)\ \Gamma(i-k+\alpha\lambda\epsilon)}{(i-k)!\ k!\ \Gamma(\alpha+i)\ \Gamma(\alpha-\alpha\lambda\epsilon)\ \Gamma(\alpha\lambda\epsilon)} \\
&= \alpha\lambda\epsilon\frac{i!\ \Gamma(\alpha)\ \Gamma(\alpha+k)\ \Gamma(i-k)}{(i-k)!\ k!\ \Gamma(\alpha+i)\ \Gamma(\alpha)} + o(\epsilon) \\
&= \left\{\frac{\alpha\ \lambda\ i!\ \Gamma(\alpha+k)}{\Gamma(\alpha+i)\ k!\ (i-k)}\right\}\epsilon\ + o(\epsilon) \\
T_{jk} &= \frac{\Gamma(\alpha\lambda\epsilon+j-k)}{\Gamma(\alpha\lambda\epsilon)\ (j-k)!}\ p^{\alpha\lambda\epsilon}(1-p)^{j-k} \\
&= \left\{\frac{\alpha\lambda\Gamma(j-k)\ (1-p)^{j-k}}{(j-k)!}\right\}\epsilon + o(\epsilon)
\end{aligned}$$

and hence their product is $S_{ik}T_{jk} = O(\epsilon^2)$ for $k < (i \wedge j)$. For $k = i < j$,

$$\begin{aligned}
S_{ii} &= \frac{\Gamma(\alpha)\ \Gamma(\alpha+i-\alpha\lambda\epsilon)}{\Gamma(\alpha+i)\ \Gamma(\alpha-\alpha\lambda\epsilon)} \\
&= \frac{\Gamma(\alpha)\ \Gamma(\alpha+i)[1-\alpha\lambda\epsilon\psi(\alpha+i)]}{\Gamma(\alpha+i)\ \Gamma(\alpha)[1-\alpha\lambda\epsilon\psi(\alpha)]} + o(\epsilon) \\
&= 1 - \alpha\lambda[\psi(\alpha+i)-\psi(\alpha)]\epsilon + o(\epsilon)
\end{aligned}$$

where $\psi(z)$ is Gauss' digamma function (Abramowitz and Stegun 1964, §6.3). For $k = j < i$,

$$T_{jj} = p^{\alpha\lambda\epsilon} = 1 + (\alpha\lambda\log p)\epsilon + o(\epsilon).$$

It follows that the transition probability is

$$\begin{aligned}
\mathsf{P}[Y_{t+\epsilon} = j \mid Y_t = i] &= \sum_{k=0}^{i\wedge j} S_{ik} \times T_{jk} = S_{i(i\wedge j)}T_{j(i\wedge j)} + o(\epsilon) \\
&= \begin{cases} \epsilon\alpha\lambda\ \left\{\frac{(1-p)^{j-i}}{j-i}\right\} + o(\epsilon) & i < j \\ \epsilon\alpha\lambda\ \left\{\frac{i!\ \Gamma(\alpha+j)}{\Gamma(\alpha+i)\ j!\ (i-j)}\right\} + o(\epsilon) & i > j \\ 1 - \epsilon\alpha\lambda[\psi(\alpha+i)-\psi(\alpha)-\log p] + o(\epsilon) & i = j \end{cases}
\end{aligned}$$

and so for bounded functions $\phi : \mathbb{N}_0 \to \mathbb{R}$,

$$\begin{aligned}
\mathsf{E}\big[\phi(Y_{t+\epsilon}) - \phi(Y_t)\big] &= \mathfrak{A}\phi(Y_t)\epsilon + o(\epsilon), \quad \text{where} \\
\mathfrak{A}\phi(i) &= \sum_{0\le j<i} \alpha\lambda[\phi(j)-\phi(i)]\left\{\frac{\Gamma(\alpha+j)\ i!}{\Gamma(\alpha+i)\ j!\ (i-j)}\right\} \qquad (8) \\
&+ \sum_{i<j<\infty} \alpha\lambda[\phi(j)-\phi(i)]\left\{\frac{(1-p)^{j-i}}{j-i}\right\} \\
&= \sum Q_{ij}\ \phi(j)
\end{aligned}$$

where, for $i, j \in \mathbb{N}_0$,

$$Q_{ij} = \begin{cases} \alpha\lambda\Gamma(\alpha+j)\; i!\big/\left[\Gamma(\alpha+i)\; j!\; (i-j)\right] & j < i \\ -\alpha\lambda[\psi(\alpha+i) - \psi(\alpha) - \log p] & j = i \\ \alpha\lambda(1-p)^{j-i}\big/\, (j-i) & j > i \end{cases} \tag{9}$$

For each $\epsilon \to 0$ let $Y_j \sim \mathsf{tNB}(\alpha, p, \lambda)$ with $\rho = 1 - \lambda\epsilon = \exp(-\lambda\epsilon) + o(\epsilon)$ be the thinning process of (6) at integer times $j \in \mathbb{Z}$. Construct a process Z_t^ϵ at times $t \in \epsilon\mathbb{Z}$ by $\left\{Z_{j\epsilon}^\epsilon = Y_j,\ j \in \mathbb{Z}\right\}$. Extend the definition of Z_t^ϵ to all $t \in \mathbb{R}$ by setting $Z_t^\epsilon = Z_{\epsilon\lfloor t/\epsilon\rfloor}^\epsilon = Y_{\lfloor t/\epsilon\rfloor}$, *i.e.*, by letting Z_t^ϵ be constant between (possible) jumps. Each Z_t^ϵ is a stationary Markov process with univariate marginal distribution $Z_t^\epsilon \sim \mathsf{NB}(\alpha, p)$ and with autocorrelation $\exp(-\lambda|t-s|) + o(\epsilon)$. Let Z_t be the limit as $\epsilon \to 0$. More formally:

There exists a unique $\mathbb{N}_0$-valued stationary Markov process Z_t with generator $\mathfrak{A}$ of (8), for which

$$\phi(Z_t) - \phi(Z_s) - \int_s^t \mathfrak{A}\phi(Z_u)\, du$$

is a martingale on $t > s$ for each bounded ϕ. This process also has $Z_t \sim \mathsf{NB}(\alpha, p)$ univariate marginal distributions, and is Markov, with autocorrelation $\exp(-\lambda|t-s|)$.

I've never seen this process in print, it seems to be new. It can't be the same as the $\mathsf{tNB}(\alpha, p, \lambda)$ process, because the conditional distribution of Y_2 given Y_0 for $\mathsf{tNB}(\alpha, p, \lambda)$ isn't the same as that of Y_1 given Y_0 for $\mathsf{tNB}(\alpha, p, 2\lambda)$. It can't be the same as the bNB, since the generators differ— recall (5). Since it's Markov, it can't coincide with the random measure $\mathsf{rmNB}(\alpha, p, \lambda)$ process.

I don't know its transition pmf $P_{ij}(t) = \mathsf{P}[X_s + t = j \mid X_s = i]$ (and hence its likelihood function), but solving the Kolmogorov forward and backward equations for the generator may lead to it if it's needed:

$$\frac{\partial}{\partial t} P_{ik}(t) = \sum_j P_{ij}(t)\; Q_{jk} \tag{10a}$$

$$= \sum_j Q_{ij}\; P_{jk}(t) \tag{10b}$$

and where $Q = \dot{P}(0)$ is given in (9), so formally $P(t) = \exp(tQ)$.

References

ABRAMOWITZ, M. and STEGUN, I. A., eds. (1964). *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, *Applied Mathematics Series*, vol. 55. National Bureau of Standards, Washington, D.C. Reprinted in paperback by Dover (1974); on-line at http://www.math.sfu.ca/~cbm/aands/.

EDWARDS, C. B. and GURLAND, J. (1961). A class of distributions applicable to accidents. *J. Am. Stat. Assoc.* **56** 503–517.

MCKENZIE, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* **21** 645–650.

MCKENZIE, E. (1986). Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Probab.* **18** 679–705.

PHATARFOD, R. M. and MARDIA, K. V. (1973). Some results for dams with Markovian inputs. *J. Appl. Probab.* **10** 166–180.

STEUTEL, F. W., VERVAAT, W. and WOLFE, S. J. (1983). Integer-valued branching processes with immigration. *Adv. Appl. Probab.* **15** 713–725.

WOLPERT, R. L. (2009). Busy diner: A Markov AR(1)-like gamma process. Research notes, originally 2009-04-19.

WOLPERT, R. L. (2011). Negative binomial & gamma processes. Research notes, 2011-08-30, Draft 35.

WOLPERT, R. L. and BROWN, L. D. (2011). Markov infinitely-divisible stationary time-reversible integer-valued processess. Discussion Paper 2011-11, Duke University Department of Statistical Science. Submission AOP1105-034 to Ann. Prob. (2011-05-30).