Stationary Infinitely-Divisible Markov Processes with Non-negative Integer Values

Robert L Wolpert

Department of Statistical Science
Duke University

Borrowing Strength: Theory Powering Applications–
A Conference in Honor of Larry Brown's 70th Birthday
December 2010, Philadelphia, PA

Prelude

Progress isn't made by early risers. It's made by lazy men trying to find easier ways to do something.

- Robert Heinlein, *Time Enough for Love*

Acknowledgment

This is collaborative work with Larry Brown, one part of a larger effort exploring and characterizing classes of stationary infinitely-divisible time series and processes.

It's partially supported by grants from the NSF and NASA and our collaboration has been facilitated by the generosity of the Wharton Department of Statistics and of SAMSI. Thanks!

Modeling Correlated Count Data

Q: How can we model processes or time series of *counts* X_t (*e.g.*, dial-ins to a call-center), w/serial autocorrelation?

A: Obvious idea: $X_t \sim \text{Po}(\Lambda_t)$, with random $\Lambda_t \geq 0$.

X: That sounds hard— even if $\Lambda_t \geq 0$ is Markov, X_t would only be *hidden* Markov. Any easier ideas?

A: Okay— how about Markov X_t?

Q: Great. What other properties should we impose?

Modeling Correlated Count Data

Q: How can we model processes or time series of *counts* X_t (*e.g.*, dial-ins to a call-center), w/serial autocorrelation?

A: Obvious idea: $X_t \sim \text{Po}(\Lambda_t)$, with random $\Lambda_t \geq 0$.

X: That sounds hard— even if $\Lambda_t \geq 0$ is Markov, X_t would only be *hidden* Markov. Any easier ideas?

A: Okay— how about Markov X_t?

Q: Great. What other properties should we impose?

Modeling Correlated Count Data

Q: How can we model processes or time series of *counts* X_t (*e.g.*, dial-ins to a call-center), w/serial autocorrelation?

A: Obvious idea: $X_t \sim \mathrm{Po}(\Lambda_t)$, with random $\Lambda_t \geq 0$.

X: That sounds hard— even if $\Lambda_t \geq 0$ is Markov, X_t would only be *hidden* Markov. Any easier ideas?

A: Okay— how about Markov X_t?

Q: Great. What other properties should we impose?

Modeling Correlated Count Data

Q: How can we model processes or time series of *counts* X_t (*e.g.*, dial-ins to a call-center), w/serial autocorrelation?

A: Obvious idea: $X_t \sim \text{Po}(\Lambda_t)$, with random $\Lambda_t \geq 0$.

X: That sounds hard— even if $\Lambda_t \geq 0$ is Markov, X_t would only be *hidden* Markov. Any easier ideas?

A: Okay— how about Markov X_t?

Q: Great. What other properties should we impose?

Modeling Correlated Count Data

Q: How can we model processes or time series of *counts* X_t (*e.g.*, dial-ins to a call-center), w/serial autocorrelation?

A: Obvious idea: $X_t \sim \text{Po}(\Lambda_t)$, with random $\Lambda_t \geq 0$.

X: That sounds hard— even if $\Lambda_t \geq 0$ is Markov, X_t would only be *hidden* Markov. Any easier ideas?

A: Okay— how about Markov X_t?

Q: Great. What other properties should we impose?

ID Count Models

For either **discrete** ($t \in \mathcal{T} = \mathbb{Z}$) or **continuous** ($t \in \mathcal{T} = \mathbb{R}$) time, model count data X_t such that for each $\vec{t} = (t_1, t_2, ..., t_p) \in \mathcal{T}^p$, $p \in \mathbb{N}$, the joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+^p$;
2. Markov.;
3. Stationary;
4. Infinitely-divisible (not just the marginals);
5. Time-reversible.

Why?

ID Count Models

For either **discrete** ($t \in \mathcal{T} = \mathbb{Z}$) or **continuous** ($t \in \mathcal{T} = \mathbb{R}$) time, model count data X_t such that for each $\vec{t} = (t_1, t_2, ..., t_p) \in \mathcal{T}^p$, $p \in \mathbb{N}$, the joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+^p$;
2. Markov;
3. Stationary.;
4. Infinitely-divisible (not just the marginals);
5. Time-reversible.

Why?

ID Count Models

For either **discrete** ($t \in \mathcal{T} = \mathbb{Z}$) or **continuous** ($t \in \mathcal{T} = \mathbb{R}$) time, model count data X_t such that for each $\vec{t} = (t_1, t_2, ..., t_p) \in \mathcal{T}^p$, $p \in \mathbb{N}$, the joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+^p$;
2. Markov;
3. Stationary;
4. Infinitely-divisible (not just the marginals).;
5. Time-reversible.

Why?

ID Count Models

For either **discrete** ($t \in \mathcal{T} = \mathbb{Z}$) or **continuous** ($t \in \mathcal{T} = \mathbb{R}$) time, model count data X_t such that for each $\vec{t} = (t_1, t_2, ..., t_p) \in \mathcal{T}^p$, $p \in \mathbb{N}$, the joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+^p$;
2. Markov;
3. Stationary;
4. Infinitely-divisible (not just the marginals);
5. Time-reversible.

Why?

ID Count Models

For either **discrete** ($t \in \mathcal{T} = \mathbb{Z}$) or **continuous** ($t \in \mathcal{T} = \mathbb{R}$) time, model count data X_t such that for each $\vec{t} = (t_1, t_2, \ldots, t_p) \in \mathcal{T}^p$, $p \in \mathbb{N}$, the joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+^p$;
2. Markov;
3. Stationary;
4. Infinitely-divisible (not just the marginals);
5. Time-reversible.

Why?

Why Markov?

- To specify the distribution of a general Z_+-valued process X_t would require specifying all finite-dimensional marginal distributions

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i],\ \vec{t} \in \mathcal{T}^p,\ \vec{x} \in \mathcal{X}^p,\ p \in \mathbb{N}$$

That sounds hard.

- For Markov processes we need only to specify the marginal

$$p_t(x) = \mathsf{P}[X_t = x]$$

and transition distributions

$$q_{st}(y \mid x) = \mathsf{P}[X_t = y \mid X_s = x].$$

That sounds easier— but not easy enough.

Why Markov?

- To specify the distribution of a general Z_+-valued process X_t would require specifying all finite-dimensional marginal distributions

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i],\ \vec{t} \in \mathcal{T}^p,\ \vec{x} \in \mathcal{X}^p,\ p \in \mathbb{N}$$

  That sounds hard.

- For Markov processes we need only to specify the marginal

$$p_t(x) = \mathsf{P}[X_t = x]$$

  and transition distributions

$$q_{st}(y \mid x) = \mathsf{P}[X_t = y \mid X_s = x].$$

  That sounds easier— but not easy enough.

Why Markov?

- To specify the distribution of a general Z_+-valued process X_t would require specifying all finite-dimensional marginal distributions

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i], \ \vec{t} \in \mathcal{T}^p, \ \vec{x} \in \mathcal{X}^p, \ p \in \mathbb{N}$$

That sounds hard.

- For Markov processes we need only to specify the marginal

$$p_t(x) = \mathsf{P}[X_t = x]$$

and transition distributions

$$q_{st}(y \mid x) = \mathsf{P}[X_t = y \mid X_s = x].$$

That sounds easier— but not easy enough.

Why Markov?

- To specify the distribution of a general Z_+-valued process X_t would require specifying all finite-dimensional marginal distributions

$$p_{\vec{t}}(\vec{x}) = \mathrm{P}[X_{t_i} = x_i],\ \vec{t} \in \mathcal{T}^p,\ \vec{x} \in \mathcal{X}^p,\ p \in \mathbb{N}$$

That sounds hard.

- For Markov processes we need only to specify the marginal

$$p_t(x) = \mathrm{P}[X_t = x]$$

and transition distributions

$$q_{st}(y \mid x) = \mathrm{P}[X_t = y \mid X_s = x].$$

That sounds easier— but not easy enough.

Why Stationary?

- Under stationarity, the Markov specification of marginal

$$p(x) = \mathrm{P}[X_t = x]$$

and transition distributions

$$q_s(y \mid x) = \mathrm{P}[X_{t+s} = y \mid X_t = x]$$

are simpler because they don't depend on t.
- We'll worry later about temporal patterns;
we can use stationary processes as building-blocks.
- Simplify the problem by facing one issue at a time.

Why Stationary?

- Under stationarity, the Markov specification of marginal

$$p(x) = \mathrm{P}[X_t = x]$$

and transition distributions

$$q_s(y \mid x) = \mathrm{P}[X_{t+s} = y \mid X_t = x]$$

are simpler because they don't depend on t.
- We'll worry later about temporal patterns;
we can use stationary processes as building-blocks.
- Simplify the problem by facing one issue at a time.

Why Stationary?

- Under stationarity, the Markov specification of marginal

$$p(x) = \mathrm{P}[X_t = x]$$

and transition distributions

$$q_s(y \mid x) = \mathrm{P}[X_{t+s} = y \mid X_t = x]$$

are simpler because they don't depend on t.
- We'll worry later about temporal patterns;
we can use stationary processes as building-blocks.
- Simplify the problem by facing one issue at a time.

Why Infinitely Divisible?

- A random *vector* $\vec{Y} \in \mathbb{R}^p$ is ID if for every $n \in \mathbb{N}$ we can write

$$\vec{Y} = \vec{Y}_1^{(n)} + \cdots + \vec{Y}_n^{(n)} \qquad \text{w/iid } \vec{Y}_j^{(n)}.$$

- Calls arrive from otherwise similar individuals of different
  - Hair color;
  - Zip codes;
  - Gender;
  - Profession;
  - Experience;
  - SES;
  - Age;
  - SSN (mod 10);
- Lévy-Khinchine characterization (for counts):

$$\log \mathsf{E}\left[e^{i\omega\cdot\vec{Y}}\right] = \sum_j \left(e^{i\omega\cdot u_j} - 1\right)\ \nu_j, \qquad \{u_j\} \subset \mathbb{R}^p,\ \{\nu_j\} \subset \mathbb{R}_+$$

$$= \int_{\mathbb{Z}_+^p} \left(e^{i\omega\cdot u} - 1\right)\ \nu(du)$$

- Simplicity: All we need specify is $\nu(du)$ on $\mathbb{Z}_+^2$.

Why Infinitely Divisible?

- A random *vector* $\vec{Y} \in \mathbb{R}^p$ is ID if for every $n \in \mathbb{N}$ we can write

$$\vec{Y} = \vec{Y}_1^{(n)} + \cdots + \vec{Y}_n^{(n)} \qquad \text{w/iid } \vec{Y}_j^{(n)}.$$

- Calls arrive from otherwise similar individuals of different
  - Hair color;
  - Zip codes;
  - Gender;
  - Profession;
  - Experience;
  - SES;
  - Age;
  - SSN (mod 10);
- Lévy-Khinchine characterization (for counts):

$$\log \mathsf{E}\left[e^{i\omega\cdot\vec{Y}}\right] = \sum_j \left(e^{i\omega\cdot u_j} - 1\right)\ \nu_j, \qquad \{u_j\} \subset \mathbb{R}^p,\ \{\nu_j\} \subset \mathbb{R}_+$$

$$= \int_{\mathbb{Z}_+^p} \left(e^{i\omega\cdot u} - 1\right)\ \nu(du)$$

- Simplicity: All we need specify is $\nu(du)$ on $\mathbb{Z}_+^2$.

Why Infinitely Divisible?

- A random *vector* $\vec{Y} \in \mathbb{R}^p$ is ID if for every $n \in \mathbb{N}$ we can write

$$\vec{Y} = \vec{Y}_1^{(n)} + \cdots + \vec{Y}_n^{(n)} \qquad \text{w/iid } \vec{Y}_j^{(n)}.$$

- Calls arrive from otherwise similar individuals of different
  - Hair color;
  - Zip codes;
  - Gender;
  - Profession;
  - Experience;
  - SES;
  - Age;
  - SSN (mod 10);
- Lévy-Khinchine characterization (for counts):

$$\log \mathsf{E}\left[e^{i\omega\cdot\vec{Y}}\right] = \sum_j \left(e^{i\omega\cdot u_j} - 1\right)\ \nu_j, \qquad \{u_j\} \subset \mathbb{R}^p,\ \{\nu_j\} \subset \mathbb{R}_+$$

$$= \int_{\mathbb{Z}_+^p} \left(e^{i\omega\cdot u} - 1\right)\ \nu(du)$$

- Simplicity: All we need specify is $\nu(du)$ on $\mathbb{Z}_+^2$.

Why Infinitely Divisible?

- A random *vector* $\vec{Y} \in \mathbb{R}^p$ is ID if for every $n \in \mathbb{N}$ we can write

$$\vec{Y} = \vec{Y}_1^{(n)} + \cdots + \vec{Y}_n^{(n)} \qquad \text{w/iid } \vec{Y}_j^{(n)}.$$

- Calls arrive from otherwise similar individuals of different
  - Hair color;
  - Zip codes;
  - Gender;
  - Profession;
  - Experience;
  - SES;
  - Age;
  - SSN (mod 10);
- Lévy-Khinchine characterization (for counts):

$$\log \mathsf{E}\left[e^{i\omega\cdot\vec{Y}}\right] = \sum_j \left(e^{i\omega\cdot u_j} - 1\right)\ \nu_j, \qquad \{u_j\} \subset \mathbb{R}^p,\ \{\nu_j\} \subset \mathbb{R}_+$$

$$= \int_{\mathbb{Z}_+^p} \left(e^{i\omega\cdot u} - 1\right)\ \nu(du)$$

- Simplicity: All we need specify is $\nu(du)$ on $\mathbb{Z}_+^2$.

Why Time-Reversible?

- That surprised us too. Turns out it's needed for uniqueness.

Why Time-Reversible?

- That surprised us too. Turns out it's needed for uniqueness.

Recap

SO— what are *all* **discrete time** ($t \in \mathcal{T} = \mathbb{Z}$) and **continuous time** ($t \in \mathcal{T} = \mathbb{R}$) time series with distributions that are:

1. Supported on the non-negative integers $\mathbb{Z}_+$;
2. Markov;
3. Stationary;
4. Infinitely-divisible (not just the marginals); and
5. Time-reversible?

The usual solutions I

- Trivial:

$$X_t \equiv X_0 \sim p(x), \text{ an arbitrary ID distribution}$$

The usual solutions II

- Trivial:

$$X_t \equiv X_0 \sim p(x), \text{ an arbitrary ID distribution, or:}$$
$$X_t \overset{\text{iid}}{\sim} p(x), \text{ an arbitrary ID distribution}$$

The usual solutions III

- Well-known: First, recall “Bivariate Poisson” distribution.

Set:

$$X_1 = \zeta_1 + \zeta_{12}, \qquad X_2 = \zeta_{12} + \zeta_2$$

$$\zeta_1, \zeta_2 \overset{\text{iid}}{\sim} \mathsf{Po}\big(\lambda(1-\rho)\big) \quad \perp\!\!\!\perp \quad \zeta_{12} \sim \mathsf{Po}(\lambda\rho);$$

$$X_1,\ X_2 \sim \mathsf{Po}(\lambda), \qquad \mathsf{Cov}(X_1, X_2) = \lambda\,\rho.$$

- Equivalent recursive construction:

$$X_1 \sim \mathsf{Po}(\lambda) \qquad X_2 \mid X_1 \sim \mathsf{Bi}(X_1, \rho) + \mathsf{Po}\big(\lambda(1-\rho)\big)$$

The usual solutions III

- Well-known: First, recall "Bivariate Poisson" distribution. Set:

$$X_1 = \zeta_1 + \zeta_{12}, \qquad X_2 = \zeta_{12} + \zeta_2$$

$$\zeta_1, \zeta_2 \overset{\text{iid}}{\sim} \mathsf{Po}\big(\lambda(1-\rho)\big) \quad \perp\!\!\!\perp \quad \zeta_{12} \sim \mathsf{Po}(\lambda\rho);$$

$$X_1,\ X_2 \sim \mathsf{Po}(\lambda), \qquad \mathsf{Cov}(X_1, X_2) = \lambda\,\rho.$$

- Equivalent recursive construction:

$$X_1 \sim \mathsf{Po}(\lambda) \qquad X_2 \mid X_1 \sim \mathsf{Bi}(X_1, \rho) + \mathsf{Po}\big(\lambda(1-\rho)\big)$$

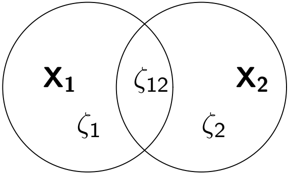
The usual solutions III

- Well-known: First, recall "Bivariate Poisson" distribution.
  Set:

$$X_1 = \zeta_1 + \zeta_{12}, \qquad X_2 = \zeta_{12} + \zeta_2$$

$$\zeta_1, \zeta_2 \overset{\text{iid}}{\sim} \mathsf{Po}\big(\lambda(1-\rho)\big) \quad \perp\!\!\!\perp \quad \zeta_{12} \sim \mathsf{Po}(\lambda\rho);$$

$$X_1,\ X_2 \sim \mathsf{Po}(\lambda), \qquad \mathsf{Cov}(X_1, X_2) = \lambda\,\rho.$$



- Equivalent recursive construction:

$$X_1 \sim \mathsf{Po}(\lambda) \qquad X_2 \mid X_1 \sim \mathsf{Bi}(X_1, \rho) + \mathsf{Po}\big(\lambda(1-\rho)\big)$$

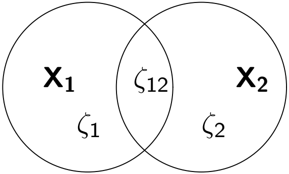
The usual solutions III

- Well-known: First, recall "Bivariate Poisson" distribution. Set:

$$X_1 = \zeta_1 + \zeta_{12}, \qquad X_2 = \zeta_{12} + \zeta_2$$

$$\zeta_1, \zeta_2 \overset{\text{iid}}{\sim} \mathsf{Po}\big(\lambda(1-\rho)\big) \quad \perp\!\!\!\perp \quad \zeta_{12} \sim \mathsf{Po}(\lambda\rho);$$

$$X_1,\ X_2 \sim \mathsf{Po}(\lambda), \qquad \mathsf{Cov}(X_1, X_2) = \lambda\,\rho.$$



- Equivalent recursive construction:

$$X_1 \sim \mathsf{Po}(\lambda) \qquad X_2 \mid X_1 \sim \mathsf{Bi}(X_1, \rho) + \mathsf{Po}\big(\lambda(1-\rho)\big)$$

The usual solutions III: Thinning Construction

- The recursive (or **Thinning**) bivariate rep'n

$$X_2 \mid X_1 \sim \mathrm{Bi}(X_1, \rho) + \mathrm{Po}\big(\lambda(1-\rho)\big)$$

leads to recursive Markov prescription (E McKenzie '85, etc.):

$$\begin{aligned} X_{t_0} &\sim \mathrm{Po}(\lambda) \\ X_{t+1} \mid \mathcal{F}_t &\sim \mathrm{Bi}(X_t, \rho) + \mathrm{Po}\big(\lambda(1-\rho)\big), \quad t \geq t_0 \end{aligned}$$

- But— is it ID? Yes. Because...

The usual solutions III: Thinning Construction

- The recursive (or **Thinning**) bivariate rep'n

$$X_2 \mid X_1 \sim \mathrm{Bi}(X_1, \rho) + \mathrm{Po}\big(\lambda(1-\rho)\big)$$

leads to recursive Markov prescription (E McKenzie '85, etc.):

$$X_{t_0} \sim \mathrm{Po}(\lambda)$$
$$X_{t+1} \mid \mathcal{F}_t \sim \mathrm{Bi}(X_t, \rho) + \mathrm{Po}\big(\lambda(1-\rho)\big), \quad t \geq t_0$$

- But— is it ID? Yes. Because...

The usual solutions III: Thinning Construction

- The recursive (or **Thinning**) bivariate rep'n

$$X_2 \mid X_1 \sim \text{Bi}(X_1, \rho) + \text{Po}\big(\lambda(1-\rho)\big)$$

leads to recursive Markov prescription (E McKenzie '85, etc.):

$$\begin{aligned} X_{t_0} &\sim \text{Po}(\lambda) \\ X_{t+1} \mid \mathcal{F}_t &\sim \text{Bi}(X_t, \rho) + \text{Po}\big(\lambda(1-\rho)\big), \quad t \geq t_0 \end{aligned}$$

- But— is it ID? Yes. Because...

The usual solutions III: Another Construction

Fix $s, u \in \mathcal{T} := \mathbb{Z}$ with $s \le u$; construct X_t as finite sum:

$$X_t = \sum_{i=u-t}^{u-s} \sum_{j=i-u+t}^{i} \zeta_{ij}, \qquad \zeta_{ij} \overset{\text{ind}}{\sim} \mathrm{Po}(\lambda_{ij}) \tag{1}$$

with intensities $\lambda_{ij} = \lambda \rho^j (1-\rho)^2$, $0 \le j < i$

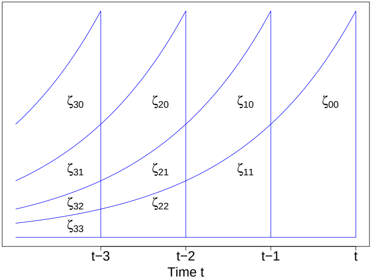


Figure: Illustration of Poisson process X_t

Is that it?

Are these three the *only* time series whose joint distribution

$$p_{\vec{t}}(\vec{x}) = \mathsf{P}[X_{t_i} = x_i]$$

is:

1. Supported on the non-negative integers $\mathbb{Z}_+$;
2. Markov;
3. Stationary;
4. Infinitely-divisible (not just the marginals);
5. Time-reversible.

?

No.

Theorem

Let X_t be a nonnegative **integer-valued** *process indexed by time $t \in \mathbb{Z}$ that is* **Markov**, **Stationary**, *(multivariate)* **Infinitely Divisible**, *and* **Time-reversible**. *Then the joint distribution of $\{X_t\}$ is one of the four possibilities:*

1. $X_t \equiv X$ *for some X with arbitrary ID dist'n on $\mathbb{Z}_+$; or*
2. $X_t \overset{\text{iid}}{\sim} \mu_0$ *for some arbitrary ID dist'n μ_0 on $\mathbb{Z}_+$; or*
3. $X_t \sim \text{Po}(\lambda)$ *with Bivariate Poisson 2-marginals.; or*
4. $X_t \sim \text{NB}(\alpha, p)$ *with Negative Trinomial 2-marginals.*

No.

Theorem

Let X_t be a nonnegative **integer-valued** *process indexed by time $t \in \mathbb{Z}$ that is* **Markov**, **Stationary**, *(multivariate)* **Infinitely Divisible**, *and* **Time-reversible**. *Then the joint distribution of $\{X_t\}$ is one of the four possibilities:*

1. $X_t \equiv X$ *for some X with arbitrary ID dist'n on $\mathbb{Z}_+$; or*
2. $X_t \overset{\text{iid}}{\sim} \mu_0$ *for some arbitrary ID dist'n μ_0 on $\mathbb{Z}_+$; or*
3. $X_t \sim \text{Po}(\lambda)$ *with Bivariate Poisson 2-marginals; or*
4. $X_t \sim \text{NB}(\alpha, p)$ *with Negative Trinomial 2-marginals.*

The Negative Trinomial distribution

- For $\alpha > 0$ and $0 \le p, q, r \le 1$ with $p + q + r = 1$, let X_0, X_1 have joint pmf:

$$\mathsf{P}[X_0 = j, X_1 = k] = \binom{-\alpha}{j,\ k} r^{\alpha}\ (-p)^j\ (-q)^k$$
$$= \frac{\Gamma(\alpha + j + k)}{\Gamma(\alpha)\ j!\ k!} r^{\alpha}\ p^j\ q^k,$$

with negative binomial univariate marginals and conditionals

$$X_0 \sim \mathsf{NB}\left(\alpha, \frac{r}{r+p}\right), \qquad X_1 \mid X_0 \sim \mathsf{NB}(\alpha + X_0, r + p)$$

- Note that (X_0, X_1) is ID...
and Stationary if $p = q \le \frac{1}{2}$, w/correlation $\rho = p/(1-p)$.

The Negative Trinomial distribution

- For $\alpha > 0$ and $0 \le p, q, r \le 1$ with $p + q + r = 1$, let X_0, X_1 have joint pmf:

$$\mathsf{P}[X_0 = j, X_1 = k] = \binom{-\alpha}{j,\ k} r^\alpha\ (-p)^j\ (-q)^k$$
$$= \frac{\Gamma(\alpha + j + k)}{\Gamma(\alpha)\ j!\ k!} r^\alpha\ p^j\ q^k,$$

with negative binomial univariate marginals and conditionals

$$X_0 \sim \mathsf{NB}\left(\alpha, \frac{r}{r+p}\right), \qquad X_1 \mid X_0 \sim \mathsf{NB}(\alpha + X_0, r + p)$$

- Note that (X_0, X_1) is ID...
and Stationary if $p = q \le \frac{1}{2}$, w/correlation $\rho = p/(1-p)$.

The Negative Trinomial distribution

- For $\alpha > 0$ and $0 \le p, q, r \le 1$ with $p + q + r = 1$, let X_0, X_1 have joint pmf:

$$\begin{aligned} \mathsf{P}[X_0 = j, X_1 = k] &= \binom{-\alpha}{j,\ \ k} r^\alpha\ (-p)^j\ (-q)^k \\ &= \frac{\Gamma(\alpha + j + k)}{\Gamma(\alpha)\ j!\ k!} r^\alpha\ p^j\ q^k, \end{aligned}$$

  with negative binomial univariate marginals and conditionals

$$X_0 \sim \mathsf{NB}\left(\alpha, \frac{r}{r+p}\right), \qquad X_1 \mid X_0 \sim \mathsf{NB}(\alpha + X_0, r + p)$$

- Note that (X_0, X_1) is ID...
  and Stationary if $p = q \le \frac{1}{2}$, w/correlation $\rho = p/(1-p)$.

Proof of Theorem

- By Stationarity & Markov property, probability generating function

$$\phi(s,t,u) = \mathrm{E}\; s^{X_0}\, t^{X_1}\, u^{X_2}$$

determines distribution of entire process;

- By Lévy-Khinchine, for some $\nu_{ijk} \geq 0$,

$$\log \phi(s,t,u) = \sum_{\mathbb{Z}_+^3} (s^i\, t^j\, u^k - 1)\, \nu_{ijk}$$

- Poisson representation: for indep. $N_{ijk} \sim \mathrm{Po}(\nu_{ijk})$,

$$X_0 = \sum iN_{i++} \quad X_1 = \sum jN_{+j+} \quad X_2 = \sum kN_{++k}$$

Proof of Theorem

- By Stationarity & Markov property, probability generating function

$$\phi(s,t,u) = \mathrm{E}\ s^{X_0}\ t^{X_1}\ u^{X_2}$$

determines distribution of entire process;

- By Lévy-Khinchine, for some $\nu_{ijk} \geq 0$,

$$\log\phi(s,t,u) = \sum_{\mathbb{Z}_+^3} (s^i\ t^j\ u^k - 1)\ \nu_{ijk}$$

- Poisson representation: for indep. $N_{ijk} \sim \mathrm{Po}(\nu_{ijk})$,

$$X_0 = \sum iN_{i++} \quad X_1 = \sum jN_{+j+} \quad X_2 = \sum kN_{++k}$$

Proof of Theorem

- By Stationarity & Markov property, probability generating function

$$\phi(s,t,u) = \mathrm{E}\ s^{X_0}\ t^{X_1}\ u^{X_2}$$

determines distribution of entire process;

- By Lévy-Khinchine, for some $\nu_{ijk} \geq 0$,

$$\log \phi(s,t,u) = \sum_{\mathbb{Z}_+^3} (s^i\ t^j\ u^k - 1)\,\nu_{ijk}$$

- Poisson representation: for indep. $N_{ijk} \sim \mathrm{Po}(\nu_{ijk})$,

$$X_0 = \sum iN_{i++} \quad X_1 = \sum jN_{+j+} \quad X_2 = \sum kN_{++k}$$

Proof of Theorem

- By Time Reversibility,

$$\nu_{ijk} = \nu_{kji}$$

- By Stationarity,

$$\nu_{ij+} = \nu_{+ij}$$

- By Markov Property and Conditioning, eventually we find just four solutions: two trivial, one Poisson, and one Negative Binomial, as claimed.

Proof Sketch of Theorem

- By Time Reversibility,

$$\nu_{ijk} = \nu_{kji}$$

- By Stationarity,

$$\nu_{ij+} = \nu_{+ij}$$

- By Markov Property and Conditioning, eventually we find just four solutions: two trivial, one Poisson, and one Negative Binomial, as claimed.

Continuous Time

- The continuous-time case turns out to be easier than the discrete one!
- The same four solutions arise: constant, iid, Poisson, and Negative Binomial.
- The Po case is a Linear Death branching process, with immigration;
- The NB case is a Linear Birth/Death branching process, with immigration.

Continuous Time

- The continuous-time case turns out to be easier than the discrete one!
- The same four solutions arise: constant, iid, Poisson, and Negative Binomial.
- The Po case is a Linear Death branching process, with immigration;
- The NB case is a Linear Birth/Death branching process, with immigration.

Continuous Time

- The continuous-time case turns out to be easier than the discrete one!
- The same four solutions arise: constant, iid, Poisson, and Negative Binomial.
- The Po case is a Linear Death branching process, with immigration;
- The NB case is a Linear Birth/Death branching process, with immigration.

What's Next?

- Stationary Markov ID Processes abound, with every ID marginal distribution— No, Ga, St, Po, NB, etc.
- Could be useful— *e.g.*, stationary Gamma processes $\Lambda_t \sim \mathsf{Ga}(\alpha, \beta)$ for point process rates $X_t \sim \mathsf{Po}(\Lambda_t)$.
- The families are rich— *e.g.*, we know of at least four distinct stationary Gamma processes Λ_t with identical marginals and covariance structures!
- Characterizing *all* of them could take us until Larry's 80th birthday... or longer... so stay tuned!

What's Next?

- Stationary Markov ID Processes abound, with every ID marginal distribution— No, Ga, St, Po, NB, etc.
- Could be useful— *e.g.*, stationary Gamma processes $\Lambda_t \sim \mathsf{Ga}(\alpha, \beta)$ for point process rates $X_t \sim \mathsf{Po}(\Lambda_t)$.
- The families are rich— *e.g.*, we know of at least four distinct stationary Gamma processes Λ_t with identical marginals and covariance structures!
- Characterizing *all* of them could take us until Larry's 80th birthday... or longer... so stay tuned!

What's Next?

- Stationary Markov ID Processes abound, with every ID marginal distribution— No, Ga, St, Po, NB, etc.
- Could be useful— *e.g.*, stationary Gamma processes $\Lambda_t \sim \text{Ga}(\alpha, \beta)$ for point process rates $X_t \sim \text{Po}(\Lambda_t)$.
- The families are rich— *e.g.*, we know of at least four distinct stationary Gamma processes Λ_t with identical marginals and covariance structures!
- Characterizing *all* of them could take us until Larry's 80th birthday... or longer... so stay tuned!

What's Next?

- Stationary Markov ID Processes abound, with every ID marginal distribution— No, Ga, St, Po, NB, etc.
- Could be useful— *e.g.*, stationary Gamma processes $\Lambda_t \sim \mathrm{Ga}(\alpha, \beta)$ for point process rates $X_t \sim \mathrm{Po}(\Lambda_t)$.
- The families are rich— *e.g.*, we know of at least four distinct stationary Gamma processes Λ_t with identical marginals and covariance structures!
- Characterizing *all* of them could take us until Larry's 80th birthday... or longer... so stay tuned!

Thanks!

More details (references, this talk in .pdf, related work) are available at

http://www.stat.duke.edu/~rlw/.

Thanks to Jim, Tony, Linda, and everyone at the Wharton Dept of Statistics. And thanks to SAMSI, the NSF, and NASA for their support. And, especially, thanks to

Larry!

Thanks!

More details (references, this talk in .pdf, related work) are available at

http://www.stat.duke.edu/∼rlw/.

Thanks to Jim, Tony, Linda, and everyone at the Wharton Dept of Statistics. And thanks to SAMSI, the NSF, and NASA for their support. And, especially, thanks to

Larry!

Thanks!

More details (references, this talk in .pdf, related work) are available at

http://www.stat.duke.edu/~rlw/.

Thanks to Jim, Tony, Linda, and everyone at the Wharton Dept of Statistics. And thanks to SAMSI, the NSF, and NASA for their support. And, especially, thanks to

Larry Brown!

Thanks!

More details (references, this talk in .pdf, related work) are available at

http://www.stat.duke.edu/~rlw/.

Thanks to Jim, Tony, Linda, and everyone at the Wharton Dept of Statistics. And thanks to SAMSI, the NSF, and NASA for their support. And, especially, thanks to

Happy Birthday, Larry!