A negative binomial model for time series of counts

Rongning Wu
Baruch College, The City University of New York

(Joint work with Richard A. Davis)

Twelfth North American Meeting of New Researchers
July 28–July 31, 2009

Outline

- Introduction
  - Time series of counts
  - Parameter-driven generalized linear models
- Parameter Estimation
  - Negative binomial regression models
- Empirical study
- Summary

Introduction

- Time series of counts

Introduction

- Time series of counts
- Generalized linear models (GLMs)
  - Response variable Y_t, covariate $\mathbf{x}_t$, & parameter $\boldsymbol{\beta}$.
  - Random component: distribution of response

$$p(Y_t) = \exp\left[\theta_t Y_t - b(\theta_t) + c(Y_t)\right].$$

  - Systematic component: link function $f(\cdot)$ s.t. $f(\mu_t) = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta}$, or function $g(\cdot)$ s.t. $\theta_t = g(\mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta})$.

Introduction

- Time series of counts
- Generalized linear models (GLMs)
  - Response variable Y_t, covariate $\mathbf{x}_t$, & parameter $\boldsymbol{\beta}$.
  - Random component: distribution of response

$$p\left(Y_t\right) = \exp\left[\theta_t Y_t - b\left(\theta_t\right) + c\left(Y_t\right)\right].$$

  - Systematic component: link function $f(\cdot)$ s.t. $f(\mu_t) = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta}$, or function $g(\cdot)$ s.t. $\theta_t = g(\mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta})$.
- Incorporation of serial dependence

Introduction

- Parameter-driven GLMs for time series
  - Observed time series $\{Y_t\}$, latent process $\{\alpha_t\}$, covariate $\mathbf{x}_t$, & parameter $\boldsymbol{\beta}$.
  - Observation equation:

$$p\,(Y_t \mid \alpha_t) \overset{\text{indep}}{\sim} \exp\left[\theta_t Y_t - b\,(\theta_t) + c\,(Y_t)\right].$$

  - States (latent variables) come into the model via $\theta_t = g(\mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + \alpha_t)$.
  - State equation: $\alpha_t = s\left(\boldsymbol{\alpha}^{(t-1)}, \zeta_t\right)$, where ζ_t is a random noise at time t.

Estimation—negative binomial regression models

- Setup
  - $Y_t \mid \alpha_t \overset{\text{indep}}{\sim} \text{NegBin}\,(r, p_t)$:

$$p\,(Y_t = y_t \mid \alpha_t) = \binom{y_t + r - 1}{r - 1} p_t^r \,(1 - p_t)^{y_t}\,,$$

and $\log \frac{1-p_t}{p_t} = \mathbf{x}_{nt}^{\mathrm{T}} \boldsymbol{\beta} + \alpha_t$.

Estimation—negative binomial regression models

- Setup
  - $Y_t \mid \alpha_t \overset{\text{indep}}{\sim} \text{NegBin}(r, p_t)$:

$$p(Y_t = y_t \mid \alpha_t) = \binom{y_t + r - 1}{r - 1} p_t^r (1 - p_t)^{y_t},$$

    and $\log \frac{1-p_t}{p_t} = \mathbf{x}_{nt}^{\mathrm{T}} \boldsymbol{\beta} + \alpha_t$.
  - Define $\{\epsilon_t = e^{\alpha_t}\}$ and center $\{\alpha_t\}$ s.t. $\{\epsilon_t\}$ has mean one. Then,

$$E(Y_t) = r \exp(\mathbf{x}_{nt}^{\mathrm{T}} \boldsymbol{\beta}).$$

Estimation—negative binomial regression models

- Setup
  - $Y_t \mid \alpha_t \overset{\text{indep}}{\sim} \text{NegBin}(r, p_t)$:

$$p(Y_t = y_t \mid \alpha_t) = \binom{y_t + r - 1}{r - 1} p_t^r (1 - p_t)^{y_t},$$

    and $\log \frac{1-p_t}{p_t} = \mathbf{x}_{nt}^{\mathrm{T}} \boldsymbol{\beta} + \alpha_t$.
  - Define $\{\epsilon_t = e^{\alpha_t}\}$ and center $\{\alpha_t\}$ s.t. $\{\epsilon_t\}$ has mean one. Then,

$$E(Y_t) = r \exp(\mathbf{x}_{nt}^{\mathrm{T}} \boldsymbol{\beta}).$$

- Primary goal: estimate $\boldsymbol{\beta}$

Estimation—negative binomial regression models

- GLM estimation
  - To estimate $\boldsymbol{\beta}$, ignore $\{\alpha_t\}$ and maximize resulting pseudo-likelihood.

Estimation—negative binomial regression models

- GLM estimation
  - To estimate $\boldsymbol{\beta}$, ignore $\{\alpha_t\}$ and maximize resulting pseudo-likelihood.
  - Classical linear model with time series noise: $Y_t = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + Z_t$.

Estimation—negative binomial regression models

- GLM estimation
  - To estimate $\boldsymbol{\beta}$, ignore $\{\alpha_t\}$ and maximize resulting pseudo-likelihood.
  - Classical linear model with time series noise: $Y_t = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + Z_t$.
- Asymptotic properties of GLM estimator $\widehat{\boldsymbol{\beta}}_n$?
  - Consistency and asymptotic normality
- Extension to one-parameter exponential family setup

Estimation—negative binomial regression models

- Assumptions on latent process
  - $\{\epsilon_t\}$ is stationary & strongly mixing.
  - Exists constant $\lambda > 0$ such that $E(|\epsilon_t|^{\lambda+4}) < \infty$.
  - Mixing coefficient $\alpha(m)$ satisfies $\sum_{m=1}^{\infty} \alpha(m)^{(\lambda+2)/\lambda} < \infty$.

Estimation—negative binomial regression models

- Assumptions on latent process
  - $\{\epsilon_t\}$ is stationary & strongly mixing.
  - Exists constant $\lambda > 0$ such that $E(|\epsilon_t|^{\lambda+4}) < \infty$.
  - Mixing coefficient $\alpha(m)$ satisfies $\sum_{m=1}^{\infty} \alpha(m)^{(\lambda+2)/\lambda} < \infty$.
- Assumptions on covariate $\mathbf{x}_{nt}$
  - Exists a sequence of nonsingular matrices M_n s.t.

$$M_n^{\mathrm{T}} \left\{ \sum_{t=1}^{n} \frac{\mathbf{x}_{nt}\mathbf{x}_{nt}^{\mathrm{T}} r e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0}}{(1 + e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})^2} \right\} M_n \to \Omega_{11},$$

$$M_n^{\mathrm{T}} \left\{ \sum_{t=1}^{n} \frac{\mathbf{x}_{nt}\mathbf{x}_{nt}^{\mathrm{T}} r e^{2\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0}}{(1 + e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})^2} \right\} M_n \to \Omega_{12},$$

and

$$M_n^{\mathrm{T}} \left\{ \sum_{t=1}^{n} \frac{\mathbf{x}_{nt}\mathbf{x}_{n,t+k}^{\mathrm{T}} r^2 e^{\left(\mathbf{x}_{nt}^{\mathrm{T}} + \mathbf{x}_{n,t+k}^{\mathrm{T}}\right)\boldsymbol{\beta}_0}}{(1 + e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})(1 + e^{\mathbf{x}_{n,t+k}^{\mathrm{T}}\boldsymbol{\beta}_0})} \right\} M_n \to W_k$$

uniformly in $|k| < n$ as $n \to \infty$.

Estimation—negative binomial regression models

- Assumptions on covariate $\mathbf{x}_{nt}$ (cont'd)
  - For each $k < 0$

$$M_n^{\mathrm{T}} \left\{ \sum_{t=1}^{-k} \frac{\mathbf{x}_{nt}\mathbf{x}_{n,t+k}^{\mathrm{T}} e^{\left(\mathbf{x}_{nt}^{\mathrm{T}}+\mathbf{x}_{n,t+k}^{\mathrm{T}}\right)\boldsymbol{\beta}_0}}{(1+e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})(1+e^{\mathbf{x}_{n,t+k}^{\mathrm{T}}\boldsymbol{\beta}_0})} \right\} M_n \to 0,$$

    & LHS is uniformly bounded in $k \in (-n, 0)$ as $n \to \infty$
  - For each $k > 0$

$$M_n^{\mathrm{T}} \left\{ \sum_{t=n-k+1}^{n} \frac{\mathbf{x}_{nt}\mathbf{x}_{n,t+k}^{\mathrm{T}} e^{\left(\mathbf{x}_{nt}^{\mathrm{T}}+\mathbf{x}_{n,t+k}^{\mathrm{T}}\right)\boldsymbol{\beta}_0}}{(1+e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})(1+e^{\mathbf{x}_{n,t+k}^{\mathrm{T}}\boldsymbol{\beta}_0})} \right\} M_n \to 0$$

    & LHS is uniformly bounded in $k \in (0, n)$ as $n \to \infty$.

Estimation—negative binomial regression models

- Assumptions on covariate $\mathbf{x}_{nt}$ (cont'd)
  - For each $k < 0$

$$M_n^{\mathrm{T}} \left\{ \sum_{t=1}^{-k} \frac{\mathbf{x}_{nt}\mathbf{x}_{n,t+k}^{\mathrm{T}} e^{\left(\mathbf{x}_{nt}^{\mathrm{T}}+\mathbf{x}_{n,t+k}^{\mathrm{T}}\right)\boldsymbol{\beta}_0}}{(1+e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})(1+e^{\mathbf{x}_{n,t+k}^{\mathrm{T}}\boldsymbol{\beta}_0})} \right\} M_n \to 0,$$

    & LHS is uniformly bounded in $k \in (-n, 0)$ as $n \to \infty$
  - For each $k > 0$

$$M_n^{\mathrm{T}} \left\{ \sum_{t=n-k+1}^{n} \frac{\mathbf{x}_{nt}\mathbf{x}_{n,t+k}^{\mathrm{T}} e^{\left(\mathbf{x}_{nt}^{\mathrm{T}}+\mathbf{x}_{n,t+k}^{\mathrm{T}}\right)\boldsymbol{\beta}_0}}{(1+e^{\mathbf{x}_{nt}^{\mathrm{T}}\boldsymbol{\beta}_0})(1+e^{\mathbf{x}_{n,t+k}^{\mathrm{T}}\boldsymbol{\beta}_0})} \right\} M_n \to 0$$

    & LHS is uniformly bounded in $k \in (0, n)$ as $n \to \infty$.
- $\mathbf{x}_{nt}$ that satisfies above conditions: $\mathbf{f}(t/n)$, $\cos(2\pi t/12)$, stationary processes, ...

Theorem

Suppose $Y_1, \ldots, Y_n$ are observations from the true model with parameter $\boldsymbol{\beta}_0$. Then, under aforementioned assumptions,

$$\widehat{\boldsymbol{\beta}}_n \xrightarrow{P} \boldsymbol{\beta}_0$$

and

$$M_n^{-1}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) \xrightarrow{d} N\left(\mathbf{0}, \Omega_1^{-1} + \Omega_1^{-1}\Omega_2\Omega_1^{-1}\right),$$

where $\Omega_1 = \Omega_{11} + \Omega_{12}$ & $\Omega_2 = \Omega_{12}\gamma_\epsilon(0) + \sum_{k=-\infty}^{\infty} W_k \gamma_\epsilon(k)$.

Empirical study—Polio data

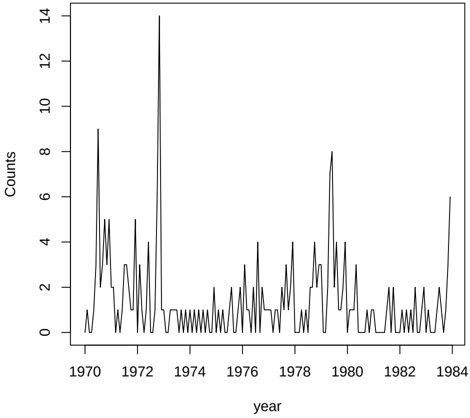

Figure: Monthly number of poliomyelitis cases in USA; 1970–1983

Empirical study—Polio data

- $Y_t|\alpha_t \overset{\text{indep}}{\sim} \text{NegBin}\,(r, p_t)$
- $\log(r(1-p_t)/p_t) = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + \alpha_t$. Then $\mu_t = e^{\mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta}}$.
- $\mathbf{x}_t = (1, t'/1000, \cos(2\pi t'/12), \sin(2\pi t'/12), \cos(2\pi t'/6), \cos(2\pi t'/6))^{\mathrm{T}}$

Empirical study—Polio data

- $Y_t|\alpha_t \overset{\text{indep}}{\sim} \text{NegBin}\,(r, p_t)$
- $\log(r(1-p_t)/p_t) = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + \alpha_t$. Then $\mu_t = e^{\mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta}}$.
- $\mathbf{x}_t =$
  $(1, t'/1000, \cos(2\pi t'/12), \sin(2\pi t'/12), \cos(2\pi t'/6), \cos(2\pi t'/6))^{\mathrm{T}}$
- Approach of Benjamin et al. (2003) yields $\hat{r} = 2$

Empirical study—Polio data

Table: Analyses by standard negative binomial and Poisson GLMs

	Poisson			Negative binomial		
Covariate	$\hat{\beta}_{Po}$	s.e.	Asy. s.e.	$\hat{\beta}_{NB}$	s.e.	Asy. s.e.
Intercept	0·207	0·075	0·205	0·209	0·100	0·167
Trend	-4·799	1·399	4·115	-4·354	1·970	3·311
$\cos(2\pi t'/12)$	-0·149	0·097	0·157	-0·143	0·134	0·156
$\sin(2\pi t'/12)$	-0·532	0·109	0·168	-0·504	0·144	0·165
$\cos(2\pi t'/6)$	0·169	0·098	0·122	0·168	0·136	0·144
$\sin(2\pi t'/6)$	-0·432	0·101	0·125	-0·422	0·138	0·146

Empirical study—Polio data

- Examine the Pearson residuals for existence of latent process.
- In both cases, PACF plot of residuals indicates an AR(1) latent process.

- Examine the Pearson residuals for existence of latent process.
- In both cases, PACF plot of residuals indicates an AR(1) latent process.

- Suppose $\alpha_t = \phi\alpha_{t-1} + Z_t, Z_t \sim \text{IID N}(0, \sigma^2)$.
- Asymptotic standard errors are readily computed.

Empirical study—Polio data

- Examine the Pearson residuals for existence of latent process.
- In both cases, PACF plot of residuals indicates an AR(1) latent process.

- Suppose $\alpha_t = \phi\alpha_{t-1} + Z_t, Z_t \sim \text{IID N}(0, \sigma^2)$.
- Asymptotic standard errors are readily computed.

- Has the incidence of polio been decreasing since 1970?

Empirical study—Polio data

Table: Analyses by parameter-driven negative binomial and Poisson GLMs

	Poisson		Negative binomial	
Covariate	$\hat{\beta}_{Po}$	s.e.	$\hat{\beta}_{NB}$	s.e.
Intercept	0·090	0·141	0·106	0·177
Trend	-3·600	2·751	-3·467	3·375
$\cos(2\pi t'/12)$	-0·098	0·143	-0·109	0·129
$\sin(2\pi t'/12)$	-0·478	0·154	-0·488	0·140
$\cos(2\pi t'/6)$	0·190	0·121	0·182	0·122
$\sin(2\pi t'/6)$	-0·355	0·122	-0·365	0·123

Empirical study—Polio data

- Examine the Pearson residuals to compare two models

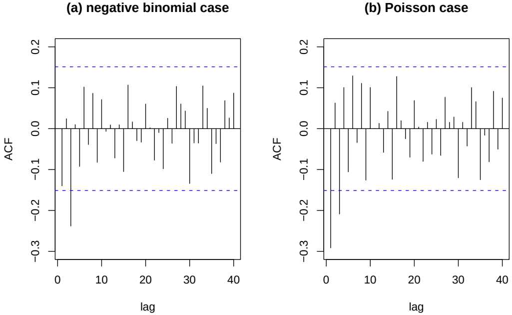

Figure: ACF plot of Pearson residuals

- *p*-value from Ljung–Box test of randomness:
  negative binomial case, 0·14550; Poisson case, 0·00097.

Summary

- Studied time series of negative binomial counts
- Established consistency & asymptotic normality of GLM estimator
- Results can be extended to one-parameter exponential family setup
- Applied methods to real data