A COM-Poisson Mixed Model for Clustered Count Data

Darcy Steeg Morris, U.S. Census Bureau
Kimberly Sellers, Georgetown University and U.S. Census Bureau

Conference on Multivariate Count Analysis
July 5, 2018

Disclaimer

This presentation is intended to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

Outline

Introduction

COM-Poisson Distribution and Regression

COM-Poisson Mixed Model

COM-Poisson-Normal Model

COM-Poisson-Conjugate Model

Analysis: Simulated Data

Analysis: Epilepsy Data

Discussion

Clustered Count Data

- Count data may exhibit over- or under-dispersion.
- Positive correlation between responses is one cause of over-dispersion (Hilbe 2008).
- Clustered data has inherent correlation within a cluster.
- Can account for the correlation by incorporating random effects in a count model.

Random Intercept Poisson Model

- Poisson distribution assumes equi-dispersion.
- Random effects loosen this assumption to capture additional variability induced by the correlation of measurements within a cluster.
- **Model Assumptions**:

$$y_{ij}|\alpha_i \sim \mathrm{Poi}(\lambda^*_{ij})$$

$$\log\left(\lambda^*_{ij}\right) = \beta_1 x_{ij1} + \cdots + \beta_p x_{ijp} + \alpha_i$$

$$\alpha_i \sim g(\alpha_i|\boldsymbol{\theta})$$

where y_{ij} is the count outcome for cluster $i = 1, \ldots, N$ at occurrence $j = 1, \ldots, J_i$; $x_{ij1}, \ldots, x_{ijp}$ are the p covariates for cluster i at occurrence j; and α_i is the cluster-specific random intercept.

Random Intercept Poisson Model

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}) = \prod_{i=1}^{N} \int \left[\prod_{j=1}^{J_i} \frac{e^{-\lambda_{ij}^*} \lambda_{ij}^{*y_{ij}}}{y_{ij}!} \right] g(\alpha_i | \boldsymbol{\theta}) \; d\alpha_i$$

- Random intercept distributional assumption:
  1. $g(\alpha_i|\boldsymbol{\theta}) = \mathrm{N}(\mu, \sigma^2) \Rightarrow$ Poisson-normal model, or
  2. $g(u_i|\boldsymbol{\theta}) =$gamma$(a, c)$, where $u_i = e^{\alpha_i} \Rightarrow$ Poisson-gamma model.
- Assumption (1) requires numerical integration.
- Assumption (2) reduces to a tractable form of the density.

Why use COM-Poisson?

- Dispersion may also exist from underlying count process mechanism (mean-variance relationship) - not adequately modeled by cluster random effects (Booth et. al. 2003, Molenberghs et. al. 2007, 2010).
- The Conway-Maxwell-Poisson (COM-Poisson) distribution is a flexible count distribution that allows for under- and over-dispersion.
- COM-Poisson model for clustered data allows modeling of additional variability due to (1) the within-cluster correlation and (2) the dispersion from the underlying count process.
- Marginal COM-Poisson models have been proposed (Khan and Jowaheer 2013, Choo-Wosoba et. al. 2016).
- We study a COM-Poisson mixed (i.e. conditional) model (Choo-Wosoba and Datta 2018, Choo-Wosoba et. al. 2018).

Outline

COM-Poisson Distribution

- The COM-Poisson pmf takes the form (Shmueli et. al., 2005)

$$P(Y = y \mid \lambda, \nu) = \frac{\lambda^y}{(y!)^\nu Z(\lambda, \nu)}, \quad y = 0, 1, 2, \ldots$$

  for a random variable Y, where $Z(\lambda, \nu) = \sum_{s=0}^{\infty} \frac{\lambda^s}{(s!)^\nu}$ is a normalizing constant.

- Dispersion parameter $\nu \geq 0$:

$$\begin{aligned} \nu = 1 &\quad \Rightarrow \quad \text{equi-dispersion} \\ \nu > 1 &\quad \Rightarrow \quad \text{under-dispersion} \\ \nu < 1 &\quad \Rightarrow \quad \text{over-dispersion} \end{aligned}$$

- Special Cases: Poisson ($\nu = 1$), geometric ($\nu = 0$, $\lambda < 1$) and Bernoulli ($\nu \to \infty$ with probability $\frac{\lambda}{1+\lambda}$).

COM-Poisson Distribution: Mean

Moments are not of closed form, but mean can be approximated:

$$E(Y) = \lambda \frac{\partial \log Z(\lambda, \nu)}{\partial \lambda} \approx \lambda^{1/\nu} - \frac{\nu - 1}{2\nu}$$

for $\nu \leq 1$ or $\lambda > 10^{\nu}$ (Shmueli et. al., 2005). Or more generally?

COM-Poisson Regression

- Sellers and Shmueli (2010) extend the COM-Poisson distribution to regression.
- Allows varying λ for each observation i.
- **Model Assumptions**:

$$y_i \sim \mathrm{CMP}(\lambda_i, \nu)$$
$$\log(\lambda_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$$

- Indirectly models the relationship between the mean and the linear predictor.

COM-Poisson Regression

- Likelihood:

$$L(\beta, \nu) = \prod_{i=1}^{N} \frac{\lambda_i^{y_i}}{(y_i!)^{\nu} Z(\lambda_i, \nu)}$$

- Loglikelihood:

$$\log L(\beta, \nu) = \sum_{i=1}^{N} y_i \log \lambda_i - \nu \sum_{i=1}^{N} \log y_i! - \sum_{i=1}^{N} \log Z(\lambda_i, \nu)$$

- Sellers and Shmueli (2010): maximum likelihood estimation.
- Guikema and Coffelt (2008): Bayesian estimation of a re-parameterized version.

Outline

Random Intercept COM-Poisson Model

- Extend Sellers and Shmueli (2010) COM-Poisson regression model to include a random effect.
- **Model Assumptions**:

$$y_{ij}|\alpha_i \sim \mathrm{CMP}(\lambda_{ij}^*, \nu)$$

$$\log\left(\lambda_{ij}^*\right) = \log\left(u_i \lambda_{ij}\right) = \beta_1 x_{ij1} + \cdots + \beta_p x_{ijp} + \alpha_i$$

$$\alpha_i \sim g(\alpha_i|\boldsymbol{\theta})$$

- α_i assumed to capture all within-cluster correlation so that

$$y_{ij} \perp y_{ik}|\alpha_i \text{ for } j \neq k.$$

Random Intercept COM-Poisson Model

$$L(\boldsymbol{\beta}, \nu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \int \left[\prod_{j=1}^{J_i} \frac{\lambda_{ij}^{*y_{ij}}}{(y_{ij}!)^{\nu}} \frac{1}{Z(\lambda_{ij}^{*}, \nu))} \right] g(\alpha_i|\boldsymbol{\theta})\ d\alpha_i$$

- Random intercept distributional assumption:
  1. $g(\alpha_i|\boldsymbol{\theta}) = N(\mu, \sigma^2) \Rightarrow$ COM-Poisson-normal model, or
  2. $g(u_i|\boldsymbol{\theta}) \propto u_i^{a-1} Z^{-c}(u_i, \nu) \Rightarrow$ COM-Poisson-conjugate model.
- Assumption (1) and (2) BOTH require numerical integration.

Outline

CMP-normal Model: Loglikelihood

- CMP-normal loglikelihood involves an intractable integral:

$$
\begin{aligned}
\log L(\boldsymbol{\beta}, \nu, \mu, \sigma^2) &= \log \left(\prod_{i=1}^{N} \int \left[\prod_{j=1}^{J_i} \overbrace{\frac{(\lambda_{ij}^*)^{y_{ij}}}{(y_{ij}!)^{\nu}} \frac{1}{Z(\lambda_{ij}^*, \nu)}}^{f(y_{ij}|\alpha_i)} \right] \overbrace{\left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\alpha_i - \mu)^2}{2\sigma^2}} \right]}^{g(\alpha_i)} d\alpha_i \right) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{J_i} y_{ij} \log(\lambda_{ij}) - \nu \sum_{i=1}^{N} \sum_{j=1}^{J_i} \log(y_{ij}!) - \sum_{i=1}^{N} \log(\sigma\sqrt{2\pi}) \\
&\quad + \sum_{i=1}^{N} \log \left(\int e^{\alpha_i \sum_{j=1}^{J_i} y_{ij} - \frac{(\alpha_i - \mu)^2}{2\sigma^2}} \left(\prod_{j=1}^{J_i} Z(e^{\alpha_i} \lambda_{ij}, \nu) \right)^{-1} d\alpha_i \right)
\end{aligned}
$$

CMP-normal Model: MLE

Obtain maximum likelihood estimates in `R` (with help `Rcpp`!) using:

1. numerical integration (the `integrate` function) to obtain an approximation of the marginal loglikelihood,
2. optimization (the `nlminb` function) to maximize the approximate marginal loglikelihood.

Maximum likelihood estimates of the CMP-normal model can similarly be obtained in SAS® using the NLMIXED procedure (Morris et. al. 2017).

Outline

COM-Poisson Conjugate (Kadane et. al. 2005)

- Conjugate prior for COM-Poisson ("extended bivariate gamma"):

$$h(\lambda, \nu) = \lambda^{a-1} e^{-\nu b} Z^{-c}(\lambda, \nu)\ \kappa(a, b, c)$$

where $\kappa(a, b, c)$ is the integration constant.

- Associated conditional distribution of λ:

$$h(\lambda|\nu) = \lambda^{a-1} Z^{-c}(\lambda, \nu)\ \kappa(a, c)$$

where $\kappa(a, c)$ is the integration constant.

CMP Conditional Conjugate: Special Cases

- Conditional conjugate distribution $h(\lambda|\nu)$ special cases:
  1. $\nu = 1 \Rightarrow \text{gamma}(a, c)$,
  2. $\nu = 0 \Rightarrow \text{beta}(a, c+1)$, and
  3. $\nu \to \infty \Rightarrow \frac{a}{c-a}F(2a, 2(c-a)), c > a \equiv \frac{\lambda}{1+\lambda} \sim \text{beta}(a, c-a)$.
- COM-Poisson conjugate relationship special cases:
  1. $\nu = 1 \Rightarrow$ Poisson-gamma,
  2. $\nu = 0 \Rightarrow$ geometric-beta, and
  3. $\nu \to \infty \Rightarrow$ Bernoulli-beta.

CMP Conditional Conjugate

$h(\lambda|\nu, a, c)$ Shiny App

CMP Conjugate: Parameter Constraints

- Joint conjugate $h(\lambda, \nu)$: $\kappa^{-1}(a, b, c)$ is finite when (Kadane et. al. 2005)

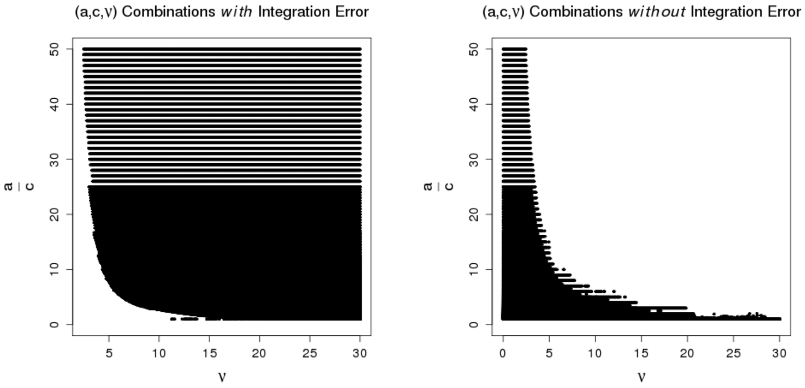
$$\frac{b}{c} > \log(\lfloor a/c \rfloor !) + (a/c - \lfloor a/c \rfloor) \log(\lfloor a/c \rfloor + 1)$$

- Conditional conjugate $h(\lambda | \nu)$:

$$\kappa^{-1}(a, c) = \int \lambda^{a-1} Z^{-c}(\lambda, \nu) d\lambda = \int \lambda^{a-1} \left[\sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\nu} \right]^{-c} d\lambda$$

divergent when $a > c$ and ν large.

Empirical Study of Parameter Constraints



$\kappa^{-1}(a, c)$ evaluated over $a \in (.1, 5), c \in (.1, 5)$ by .1 and $\nu \in (0, 30)$ by .05 for $a > c$.

CMP-conjugate Model: Loglikelihood

- *Poisson-gamma model*: conjugate distribution $\Rightarrow$ closed form.
- Unfortunately CMP-conjugate loglikelihood involves intractable integrals (violates strong conjugacy, Molenberghs et. al. 2010):

$$
\begin{aligned}
\log L(\boldsymbol{\beta}, \nu, a, c) &= \log \left(\prod_{i=1}^{N} \int \left[\prod_{j=1}^{J_i} \overbrace{\frac{(\lambda_{ij}^{*})^{y_{ij}}}{(y_{ij}!)^{\nu}} \frac{1}{Z(\lambda_{ij}^{*}, \nu)}}^{f(y_{ij} | u_i)} \right] \overbrace{\left[u_i^{a-1} Z^{-c}(u_i, \nu)\ \kappa(a, c) \right]}^{g(u_i)} du_i \right) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{J_i} y_{ij} \log(\lambda_{ij}) - \nu \sum_{i=1}^{N} \sum_{j=1}^{J_i} \log(y_{ij}!) + \sum_{i=1}^{N} \log\left(\kappa(a, c)\right) \\
&\quad + \sum_{i=1}^{N} \log \left(\int_{u_i} u_i^{a-1+\sum_{j=1}^{J_i} y_{ij}} \left(Z^{c}(u_i, \nu) \prod_{j=1}^{J_i} Z(u_i \lambda_{ij}, \nu) \right)^{-1} du_i \right)
\end{aligned}
$$

CMP-conjugate Model: MLE

Obtain maximum likelihood estimates in `R` (with help `Rcpp`!) from using:

1. numerical integration (the `integrate` function) to obtain an approximation of the marginal loglikelihood,
2. optimization (the `nlminb` function) to maximize the approximate marginal loglikelihood.

Outline

Simulated Data Generating Process

- $N = 100$ clusters & $J_i = 5 \; \forall \; i$ (500 observations) & 50 replications.
- **Distributional Assumptions**:

$$y_{ij}|u_i \sim f(y_{ij}|\lambda^*_{ij}, \{\nu\})$$

$$\log\left(\lambda^*_{ij}\right) = \beta_1 x_i + \alpha_i$$

$$x_i \sim \mathrm{N}(0, .1) \text{ and } \alpha_i \sim \mathrm{N}(.5, .5) \text{ and } \beta_1 = .5$$

where $f(y_{ij}|\lambda^*_{ij}, \{\nu\})$ is:

- $\text{Poi}(\lambda^*_{ij}) \equiv \text{CMP}\left(\lambda^*_{ij}, 1\right)$
- $\text{Bern}\left(\frac{\lambda^*_{ij}}{1+\lambda^*_{ij}}\right) \equiv \text{CMP}\left(\lambda^*_{ij}, \infty\right)$
- $\text{geom}\left(\frac{1}{1+\lambda^*_{ij}}\right) \cong \text{CMP}\left(\lambda^*_{ij}, 0\right)$
- $\text{CMP}\left(\lambda^*_{ij}, 5\right)$
- $\text{CMP}\left(\lambda^*_{ij}, .75\right)$

Simulated Data: Mean Link

Recall

$$E(Y) = \lambda^* \frac{\partial \log Z(\lambda^*, \nu)}{\partial \lambda^*}$$

- For Poisson data:

$$E(Y) = \lambda^* \quad \equiv \quad E(Y) = \lambda^* \frac{\partial \log\left(e^{\lambda^*}\right)}{\partial \lambda^*} = \lambda^*$$

- For Bernoulli data:

$$E(Y) = p = \frac{\lambda^*}{1+\lambda^*} \quad \equiv \quad E(Y) = \lambda^* \frac{\partial \log\left(1+\lambda^*\right)}{\partial \lambda^*} = \frac{\lambda^*}{1+\lambda^*}$$

- For geometric data:

$$E(Y) = \frac{1-p}{p} = \lambda^* \quad \not\equiv \quad E(Y) = \lambda^* \frac{\partial \log\left(1-\lambda^*\right)^{-1}}{\partial \lambda^*} = \frac{\lambda^*}{1-\lambda^*}$$

Simulated Data: Misspecification

- Fit 4 models: Poisson-normal, NB-normal, CMP-normal, CMP-conjugate.
- Sources of misspecification.
  - Random effect distribution.
    - *Special Cases*: CMP-conjugate.
    - *COM-Poisson Data*: CMP-conjugate.
  - Link to linear predictor.
    - *Special Cases*: CMP-normal/conjugate for geometric data.
    - *COM-Poisson Data*: Poisson-normal, NB-normal.

Simulation Study Results: Special Cases

Special Case Simulated Data *Mean* Estimates

Simulated Dataset	Est.	Model: Poisson	NB	CMP-normal	CMP-conjugate
Poisson	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 1.02$	$\hat{\nu} = 0.99$
	Var.	$\hat{\sigma}^2 = 0.49$	$\hat{\sigma}^2 = 0.49$	$\hat{\sigma}^2 = 0.51$	$\hat{a} = 2.12, \hat{c} = 1.01$
	min AIC	**0.96**	**0.72**	**0.96**	0.12
	max ℓ	0.00	0.04	**0.76**	0.20
Bernoulli**	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 37.9$	$\hat{\nu} = 34.8$
	Var.	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.53$	$\hat{a} = 7.26, \hat{c} = 11.75$
	min AIC	0.00	0.00	**1.00**	**1.00**
	max ℓ	0.00	0.00	**0.55**	**0.45**
Geometric	Disp		$\hat{k} = 1.01$	$\hat{\nu} = 0.02$	$\hat{\nu} = 0.02$
	Var.	$\hat{\sigma}^2 = 0.67$	$\hat{\sigma}^2 = 0.45$	$\hat{\sigma}^2 = 0.04$	$\hat{a} = 6.70, \hat{c} = 3.33$
	min AIC	0.00	**0.98**	0.22	**0.34**
	max ℓ	0.00	**0.64**	0.02	**0.34**

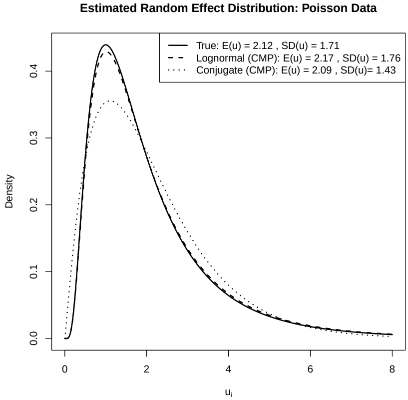
Note: min AIC is the proportion of replications where AIC $\leq$ min(AIC) + 2 and max ℓ is proportion of replications where ℓ is largest.

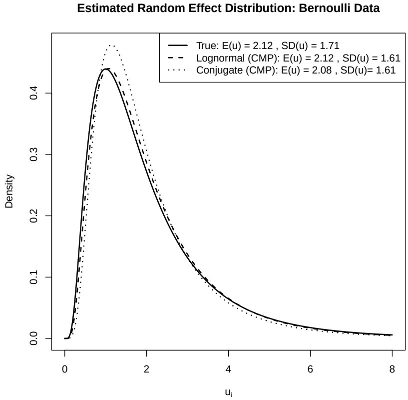
** The random intercept logistic model results/estimates for the simulated Bernoulli data are: $\hat{\sigma}^2 = 0.45$, min AIC = **1.00**, and max $\ell = 0.00$.

Simulation Study Results: Special Cases

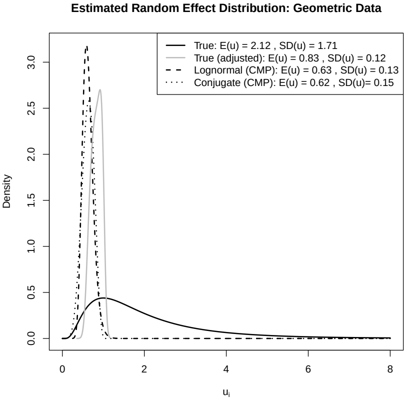
- COM-Poisson models have better/comparable model fit to special cases.
- COM-Poisson model recognizes special cases:
  - $\overline{\hat{\nu}} = 1.02, 0.99 \approx 1$ for Poisson,
  - $\overline{\hat{\nu}} = 37.9, 34.8$ is large for Bernoulli, and
  - $\overline{\hat{\nu}} = 0.02, 0.02 \approx 0$ for geometric.
- Cluster variability.
  - Captured by Poisson and NB for over-dispersed data: $\overline{\hat{\sigma}^2} > 0$. (NB recognizes geometric special case: $\overline{\hat{k}} \approx 1$.)
  - Not captured by Poisson or NB for under-dispersed data: $\overline{\hat{\sigma}^2} = 0$.
  - Captured by COM-Poisson models: $\overline{\hat{\sigma}^2} > 0$ and ...

Simulation Study Results: Special Cases





Simulation Study Results: Special Cases



COM-Poisson Simulated Data *Mean* Estimates

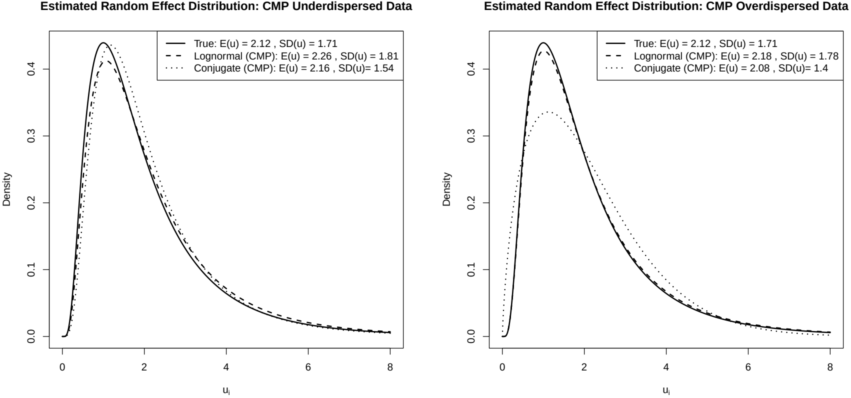
Simulated Dataset	Est.	Poisson	NB	CMP-normal	CMP-conjugate
CMP (under)	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 5.05$	$\hat{\nu} = 5.08$
	Var.	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.50$	$\hat{a} = 6.00, \hat{c} = 8.83$
	min AIC	0.00	0.00	**1.00**	**0.97**
	max ℓ	0.00	0.00	**0.58**	**0.42**
CMP (over)	Disp		$\hat{k} = 0.00$	$\hat{\nu} = 0.77$	$\hat{\nu} = 0.74$
	Var.	$\hat{\sigma}^2 = 0.76$	$\hat{\sigma}^2 = 0.75$	$\hat{\sigma}^2 = 0.51$	$\hat{a} = 1.76, \hat{c} = 0.56$
	min AIC	0.17	0.19	**0.93**	0.19
	max ℓ	0.00	0.10	**0.79**	0.12

Note: min AIC is the proportion of replications where AIC $\leq$ min(AIC) + 2 and max ℓ is proportion of replications where ℓ is largest.

Simulation Study Results: COM-Poisson Data

- COM-Poisson models outperform for both cases of intermediate levels of over- and under-dispersion.
- COM-Poisson models recognize over- and under-dispersion:
  - $\overline{\hat{\nu}} = 5.05, 5.08 \approx 5.00 > 1$ for COM-Poisson (under), and
  - $\overline{\hat{\nu}} = 0.77, 0.74 \approx 0.75 < 1$ for COM-Poisson (over).
- Cluster variability.
  - Not captured by Poisson or NB for under-dispersed data: $\overline{\hat{\sigma}^2} = 0$.
  - Captured by COM-Poisson models: $\overline{\hat{\sigma}^2} > 0$ and ...

Simulation Study Results: COM-Poisson Data



United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Simulation Study Results: Model Comparisons

Simulated Data Best Model by AIC

Simulated Dataset	Model: Poisson	NB	CMP-normal	CMP-conjugate
Poisson	X	X	X	
Bernoulli			X	**X**
Geometric		X		**X**
CMP (under)			X	**X**
CMP (over)			X	

Note: bolded are misspecified models.

- But what if random effect distribution is misspecified for -normal models?

$$u_i \sim \mathrm{gamma}(1.54, 1.37)$$

assuming same mean and variance as $u_i \sim \log \mathrm{N}(.5, .5)$.

Simulated Data: Misspecification

- Fit 4 models: Poisson-normal, NB-normal, CMP-normal, CMP-conjugate.
- Sources of misspecification.
  - Random effect distribution.
    - *Special Cases*: All except CMP-conjugate for Poisson data.
    - *COM-Poisson Data*: All.
  - Link to linear predictor.
    - *Special Cases*: CMP-normal/conjugate for geometric data.
    - *COM-Poisson Data*: Poisson-normal, NB-normal.

Simulation Study Results: Special Cases

Special Case Simulated Data *Mean* Estimates

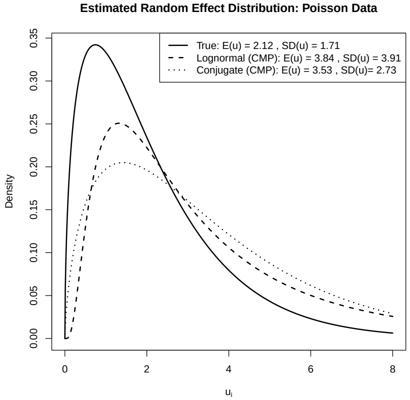
Simulated Dataset	Est.	Model: Poisson	NB	CMP-normal	CMP-conjugate
Poisson	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 1.03$	$\hat{\nu} = 1.00$
	Var.	$\hat{\sigma}^2 = 0.68$	$\hat{\sigma}^2 = 0.68$	$\hat{\sigma}^2 = 0.71$	$\hat{a} = 1.67, \hat{c} = 0.48$
	min AIC	0.28	0.08	0.22	**0.88**
	max ℓ	0.00	0.00	0.12	**0.88**
Bernoulli**	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 36.0$	$\hat{\nu} = 35.6$
	Var.	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.96$	$\hat{a} = 4.47, \hat{c} = 6.50$
	min AIC	0.00	0.00	**1.00**	**1.00**
	max ℓ	0.00	0.00	**0.69**	**0.31**
Geometric	Disp		$\hat{k} = 1.03$	$\hat{\nu} = 0.00$	$\hat{\nu} = 0.00$
	Var.	$\hat{\sigma}^2 = 0.90$	$\hat{\sigma}^2 = 0.66$	$\hat{\sigma}^2 = 0.04$	$\hat{a} = 5.62, \hat{c} = 1.59$
	min AIC	0.00	**0.97**	0.00	**0.45**
	max ℓ	0.00	**0.76**	0.00	**0.24**

Note: min AIC is the proportion of replications where AIC $\leq$ min(AIC) + 2 and max ℓ is proportion of replications where ℓ is largest.

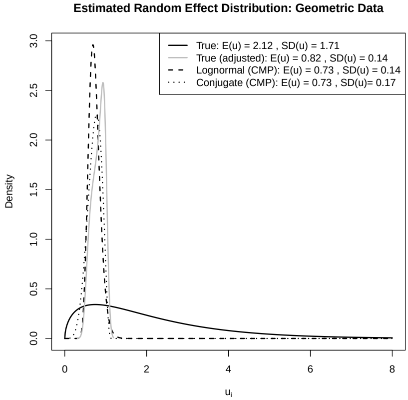
** The random intercept logistic model results/estimates for the simulated Bernoulli data are: $\hat{\sigma}^2 = 0.85$, min AIC $=$ **1.00**, and max $\ell = 0.00$.

United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Simulation Study Results: Special Cases



Simulation Study Results: Special Cases



United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Simulation Study Results: COM-Poisson Data

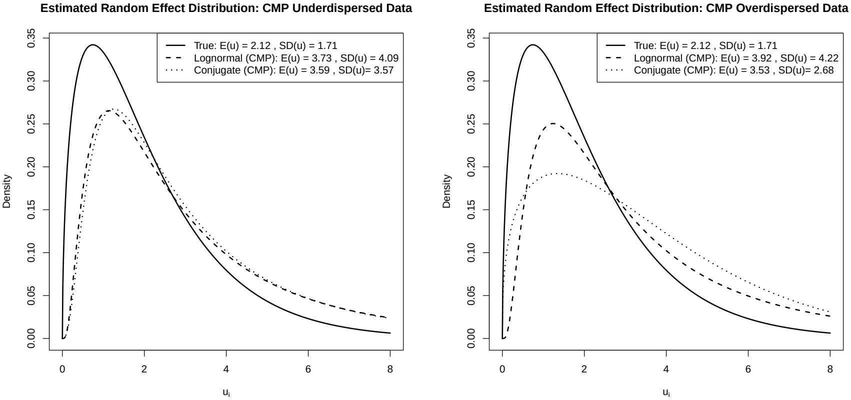
COM-Poisson Simulated Data *Mean* Estimates

Simulated Dataset	Est.	Model: Poisson	NB	CMP-normal	CMP-conjugate
CMP (under)	Disp.		$\hat{k} = 0.00$	$\hat{\nu} = 5.10$	$\hat{\nu} = 5.09$
	Var.	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.00$	$\hat{\sigma}^2 = 0.79$	$\hat{a} = 4.01, \hat{c} = 5.14$
	min AIC	0.00	0.00	**1.00**	**1.00**
	max ℓ	0.00	0.00	**0.21**	**0.79**
CMP (over)	Disp		$\hat{k} = 0.02$	$\hat{\nu} = 0.78$	$\hat{\nu} = 0.76$
	Var.	$\hat{\sigma}^2 = 1.14$	$\hat{\sigma}^2 = 1.14$	$\hat{\sigma}^2 = 0.77$	$\hat{a} = 1.38, \hat{c} = 0.23$
	min AIC	0.06	0.09	0.23	**0.86**
	max ℓ	0.00	0.06	0.14	**0.80**

Note: min AIC is the proportion of replications where AIC $\leq$ min(AIC) + 2 and max ℓ is proportion of replications where ℓ is largest.

United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Simulation Study Results: COM-Poisson Data



United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Simulation Study Results: Model Comparisons

Simulated Data Best Model by AIC

Simulated Dataset	Model: Poisson	Model: NB	Model: CMP-normal	Model: CMP-conjugate
Poisson	~~X~~	~~X~~	~~X~~	X
Bernoulli			**X**	**X**
Geometric		**X**		**X**
CMP (under)			**X**	**X**
CMP (over)			~~X~~	**X**

Note: bolded are misspecified models.

Outline

Epilepsy Data & Model (Diggle et. al. 1994)

- Number of seizures measured for 59 epileptic patients in an 8-week baseline period followed by 4 consecutive 2-week treatment periods.
- **Outcome Variable**, y_{ij}: number of seizures for subject i in time period j.
- **Covariates**:
  - x_{ij1}: indicator of a period after baseline (weeks $8-16$).
  - x_{ij2}: indicator of receipt of progabide (vs. placebo).
  - T_{ij}: length of time period t.
- **Model**: Diggle et. al. (1994) fit a random intercept Poisson regression model.

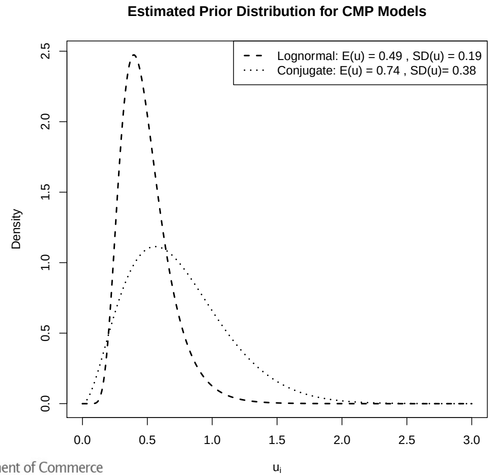
$$\log E(y_{ij}|u_i) = \beta_0 + x_{ij1}\beta_1 + x_{ij2}\beta_2 + x_{ij1}x_{ij2}\beta_3 + \log(T_{ij}) + u_i$$

Epilepsy Data Estimates

	Model			
Parameter	Poisson	NB	CMP-normal	CMP-conjugate
β_0 or μ	1.033	1.080	-0.781	
β_1	0.111	0.023	0.793	0.633
β_2	-0.024	0.073	0.048	0.165
β_3	-0.104	-0.310	-0.176	-0.198
k		0.148		
ν			0.420	0.541
σ^2	0.608	0.661	0.143	
a, c				2.93, 2.79
AIC	2031.4	1789.5	**1754.0**	1776.6
$-\ell$	1010.6	888.7	**871.0**	882.3

Epilepsy Data Results

- CMP models have the best fit.
- All models indicate subject variability: $\hat{\sigma}^2 > 0$ and ...



Epilepsy Data Results

- Additional over-dispersion is evident in CMP and NB:
  - $\nu < 1$ ($\hat{\nu} = 0.420, 0.541$) in COM-Poisson, and
  - $k > 0$ ($\hat{k} = 0.148$) in negative binomial.
- Both the negative binomial and COM-Poisson models can account for variability beyond the subject-specific random effect, however the COM-Poisson captures additional over-dispersion in a way that the negative binomial model cannot.
- Findings consistent with Molenberghs et. al. (2010).

Outline

Discussion

- COM-Poisson regression model can be extended to include random effects to model clustered data.
- The flexibility of the COM-Poisson mixed model allows modeling of variability in the count outcome beyond that induced by within-cluster correlation.
- Assuming the conditional conjugate distribution for random effects allows further flexibility.
- Framework naturally allows random slopes, mixed modeling of the dispersion parameter, etc.
- To do: Mean parametrization?!

Thank you!

darcy.steeg.morris@census.gov

References

- Shmueli, Minka, Kadane, Borle and Boatwright. (2005). "A Useful Distribution for Fitting Discrete Data: Revival of the Conway-Maxwell-Poisson Distribution." *Journal of the Royal Statistical Society, Series C*, 54: 127-142.
- Sellers and Shmueli. (2010). "A Flexible Regression Model for Count Data." *The Annals of Applied Statistics*, 4(2): 943-961.
- Guikema and Coffelt. (2008). "A Flexible Count Data Regression Model for Risk Analysis." *Risk Analysis*. 28(1): 213-223.
- Kadane, Shmueli, Minka, Borle and Boatwright, P. (2006). "Conjugate Analysis of the Conway-Maxwell-Poisson Distribution." *Bayesian Analysis*. 1(2): 363-374.
- Diggle, Heagerty, Liang and Zeger. (1994). *Analysis of Longitudinal Data.* Oxford: Clarendon.

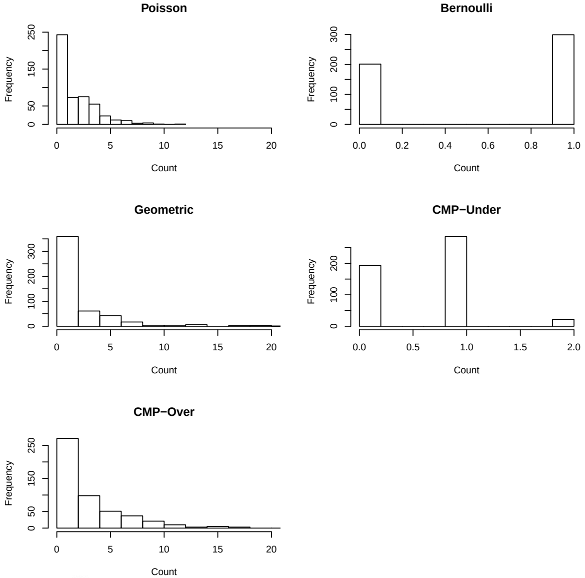
References

- Choo-Wosoba and Datta. (2018). "Analyzing Clustered Count Data with a Cluster-specific Random Effect Zero-inflated Conway-Maxwell-Poisson Distribution." *Journal of Applied Statistics*, 45(5): 799-814.
- Choo-Wosoba, Gaskins, Levy and Datta. (2018). "A Bayesian approach for analyzing zero-inflated clustered count data with dispersion: Bayesian Conway-Maxwell-Poisson." *Statistics in Medicine*, 37(1): 801-812.
- Choo-Wosoba, Levy and Datta. (2016). "Marginal Regression Models for Clustered Count Data Based on Zero-Inflated Conway-Maxwell-Poisson Distribution with Applications." *Biometrics*, 72(2): 606-618.
- Morris, Sellers and Menger. (2017). "Fitting a Flexible Model for Longitudinal Count Data Using the NLMIXED Procedure," In *SAS Global Forum Proceedings*. Cary, NC: SAS Institute.

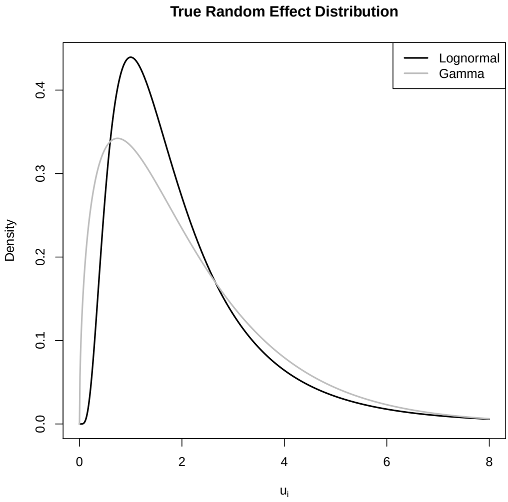
References

- Booth, Casella, Friedl and Hobert. (2003). "Negative Binomial Loglinear Mixed Models." *Statistical Modelling*, 3: 179-191.
- Molenberghs, Verbeke, Demetrio and Vieira. (2010). "A Family of Generalized Linear Models for Repeated Measures with Normal and Conjugate Random Effects." *Statistical Science*, 25(3): 325-247.
- Molenberghs, Verbeke and Demetrio. (2007). "An Extended Random-Effects Approach to Modeling Repeated, Overdispersed Count Data." *Lifetime Data Analysis*, 13: 513-531.

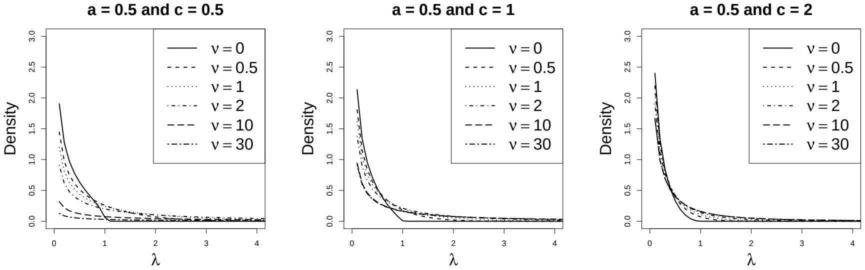
Simulated Data Histograms



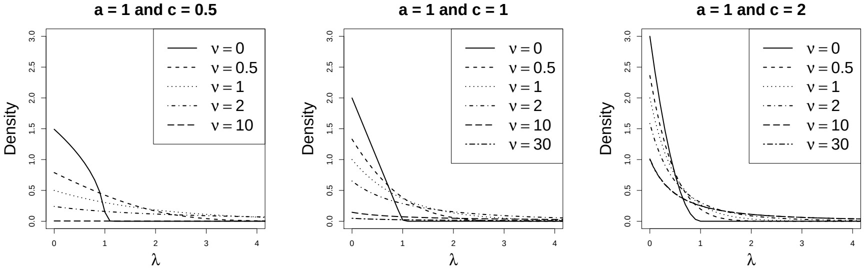
True Random Effects Distribution



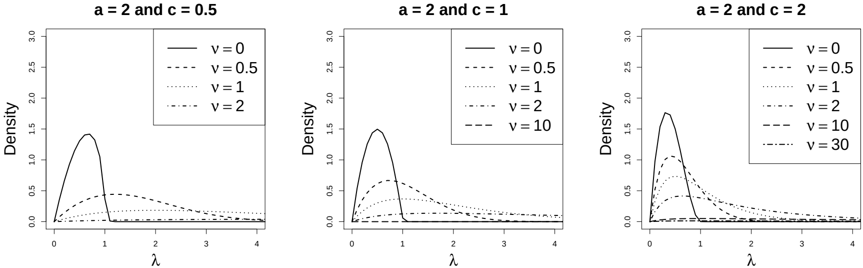
CMP Conditional Conjugate



CMP Conditional Conjugate



CMP Conditional Conjugate



CMP Conditional Conjugate