A Flexible Zero-Inflated Model to Address Data Dispersion

Andrew M. Raim

Center for Statistical Research and Methodology
U.S. Census Bureau
`andrew.raim@census.gov`

2016 International Conference on Statistical Distributions and Applications

Joint work with **Kimberly F. Sellers** (Georgetown University)

This talk to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Overview

- Conway-Maxwell-Poisson (CMP) is a flexible count distribution that can handle both under- and overdispersion, relative to Poisson.
- Excess zeroes are frequently encountered in count datasets.
- If both a zero-generating process and a count distribution are responsible for the data, the count distribution may be either under- or overdispersed (Sellers and Shmueli, 2013).
- This motivates our study of the Zero-Inflated CMP (ZICMP) distribution and associated regression models.
- We will introduce ZICMP, study properties of the maximum likelihood estimator (MLE) and a test for equidispersion, and apply the model to a real dataset.

Literature Review

- Zero-Inflated Poisson (ZIP) is a popular regression model for count data with excess zeroes (Lambert, 1992).
- ZIP regression has been used in manufacturing (Lambert, 1992), horticulture (Hall, 2000), zoology (Zipkin et al., 2014), and criminology (Famoye and Singh, 2006).
- Zero-Inflated Negative Binomial (ZINB) regression is often used under both overdispersion and excess zeroes (Hilbe, 2011). A special case of the ZINB distribution is the Zero-Inflated Geometric (ZIG) distribution.
- These and other zero-inflated regression models are available in the `VGAM` `R` package.
- CMP is a flexible, two-parameter distribution for count data expressing under- or overdispersion (Conway and Maxwell, 1962).
- A ZICMP regression model addresses the excess zeroes and provides flexibility in modeling dispersion.

ZICMP Model

- Recall the CMP distribution. Write $Y_0 \sim \text{CMP}(\lambda, \nu)$ for Y_0 with density

$$f(y \mid \boldsymbol{\theta}) = \frac{\lambda^y}{(y!)^\nu Z(\lambda, \nu)}, \quad y = 0, 1, \dots,$$

  where $Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}$ and $\boldsymbol{\theta} = (\lambda, \nu)$.

- Suppose $S \sim \text{Ber}(p)$ and $Y_0 \sim \text{CMP}(\lambda, \nu)$ independently, and let

$$Y = S \cdot 0 + (1 - S) Y_0.$$

  We will write $Y \sim \text{ZICMP}(\lambda, \nu, p)$.

- Let $\Delta = I(Y = 1)$ and $\boldsymbol{\theta} = (\lambda, \nu, p)$; density of Y can be written as

$$f(y \mid \boldsymbol{\theta}) = \left[\frac{p \left\{ Z(\lambda, \nu) - 1 \right\} + 1}{Z(\lambda, \nu)} \right]^{\Delta} \left[\frac{(1 - p)\lambda^y}{(y!)^\nu Z(\lambda, \nu)} \right]^{1 - \Delta}.$$

ZICMP Model

- Moments of ZICMP can be computed from CMP using

$$\mathrm{E}(Y^r) = \mathrm{E}[(1-S)^r]\,\mathrm{E}(Y_0^r) = (1-p)\,\mathrm{E}(Y_0^r),$$
$$\mathrm{E}(Y_0^{r+1}) = \begin{cases} \lambda[\mathrm{E}(Y_0+1)]^{1-\nu} & r = 0, \\ \lambda \frac{\partial}{\partial\lambda}\,\mathrm{E}(Y_0^r) + \mathrm{E}(Y_0)\,\mathrm{E}(Y_0^r) & r > 0. \end{cases}$$

- Moment generating function of ZICMP can be computed from CMP using

$$\begin{aligned} \mathrm{E}(e^{tY}) &= \mathrm{E}_S\,\mathrm{E}_{Y_0|S}[e^{t(1-S)Y_0}] \\ &= \mathrm{E}_S\left[\frac{Z(\lambda e^{t(1-S)},\nu)}{Z(\lambda,\nu)}\right] \\ &= p + (1-p)\frac{Z(\lambda e^t,\nu)}{Z(\lambda,\nu)}. \end{aligned}$$

ZICMP Special Cases

- If $\nu = 1$, pdf of ZICMP(λ, ν, p) becomes Zero-Inflated Poisson,

$$Y \sim \begin{cases} 0 & \text{w.p. } p, \\ \text{Poisson}(\lambda) & \text{w.p. } 1 - p. \end{cases}$$

- If $\nu = 0$, pdf of ZICMP(λ, ν, p) becomes Zero-Inflated Geometric,

$$Y \sim \begin{cases} 0 & \text{w.p. } p, \\ \text{Geometric}(1 - \lambda) & \text{w.p. } 1 - p. \end{cases}$$

- As $\nu \to \infty$, pdf of ZICMP(λ, ν, p) becomes "Zero-Inflated Bernoulli",

$$Y \sim \begin{cases} 0 & \text{w.p. } p, \\ \text{Ber}(\frac{\lambda}{1+\lambda}) & \text{w.p. } 1 - p, \end{cases}$$

  which is actually just Ber$\left(\frac{(1-p)\lambda}{1+\lambda}\right)$. However, λ and p are not identifiable.

ZICMP Regression Model

- We will consider an indepedent sample $Y_i \sim \text{ZICMP}(\lambda_i, \nu, p_i)$, $i = 1, \ldots, n$, where

$$\log(\lambda_i) = \mathbf{x}_i^T \boldsymbol{\beta} \quad \text{and} \quad \text{logit}(p_i) \equiv \log \frac{p_i}{1 - p_i} = \mathbf{w}_i^T \boldsymbol{\zeta}.$$

- We could further model $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ through a regression if desired, e.g. with $\log(\nu_i) = \mathbf{s}_i^T \boldsymbol{\gamma}$.

- The log-likelihood for $\boldsymbol{\theta} = (\boldsymbol{\beta}, \nu, \boldsymbol{\zeta})$ is

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \Big\{ \Delta_i \log \Big[p_i Z(\lambda_i, \nu) + (1 - p_i) \Big] - \log Z(\lambda_i, \nu) \\ + (1 - \Delta_i) \Big[\log(1 - p_i) + y_i \log(\lambda_i) - \nu \log(y_i!) \Big] \Big\}.$$

ZICMP Score Function

- First derivatives of the log-density $\ell_i(\boldsymbol{\theta}) = \log f(y_i \mid \boldsymbol{\theta})$ for $\boldsymbol{\theta} = (\lambda, \nu, p)$ are

$$\begin{aligned}
\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \lambda} &= -\frac{1-p}{zp + (1-p)} \frac{\partial \log z}{\partial \lambda} \Delta + \frac{y}{\lambda}(1-\Delta) - \frac{\partial \log z}{\partial \lambda}(1-\Delta), \\
\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \nu} &= -\frac{1-p}{zp + (1-p)} \frac{\partial \log z}{\partial \nu} \Delta + (1-\Delta) \log \Gamma(y+1) - \frac{\partial \log z}{\partial \nu}(1-\Delta), \\
\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial p} &= \frac{z-1}{zp + (1-p)} \Delta - \frac{1}{1-p}(1-\Delta),
\end{aligned}$$

  using the shorthand $z = Z(\lambda, \nu)$.

- Using these expressions, we obtain the score function

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log \mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

ZICMP Information Matrix

The Fisher information matrix (FIM) for ZICMP with $\boldsymbol{\theta} = (\lambda, \nu, p)$ has form

$$\mathcal{I}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathcal{I}_{\lambda\lambda} & \mathcal{I}_{\lambda\nu} & \mathcal{I}_{\lambda p} \\ \mathcal{I}_{\lambda\nu} & \mathcal{I}_{\nu\nu} & \mathcal{I}_{\nu p} \\ \mathcal{I}_{\lambda p} & \mathcal{I}_{\nu p} & \mathcal{I}_{pp} \end{pmatrix}.$$

Denoting $\mu = E(Y) = (1-p)\lambda \frac{\partial \log z}{\partial \lambda}$, FIM components are

$$\mathcal{I}_{\lambda\lambda} = (1-p)\frac{\partial^2 \log z}{\partial \lambda^2} - \frac{p(1-p)}{zp + (1-p)}\left(\frac{\partial \log z}{\partial \lambda}\right)^2 + \frac{\mu}{\lambda^2},$$

$$\mathcal{I}_{\nu\nu} = (1-p)\frac{\partial^2 \log z}{\partial \nu^2} - \frac{p(1-p)}{zp + (1-p)}\left(\frac{\partial \log z}{\partial \nu}\right)^2,$$

$$\mathcal{I}_{pp} = \frac{1}{z}\frac{(z-1)^2}{zp + (1-p)} + \frac{1}{z}\frac{z-1}{1-p},$$

$$\mathcal{I}_{\lambda\nu} = (1-p)\frac{\partial^2 \log z}{\partial \nu \partial \lambda} - \frac{p(1-p)}{zp + (1-p)}\frac{\partial \log z}{\partial \nu}\frac{\partial \log z}{\partial \lambda},$$

$$\mathcal{I}_{\lambda p} = -\frac{1}{zp + (1-p)}\frac{\partial \log z}{\partial \lambda}, \qquad \mathcal{I}_{\nu p} = -\frac{1}{zp + (1-p)}\frac{\partial \log z}{\partial \nu}.$$

The FIM becomes singular under non-identifiability (Rothenberg, 1971).

Computational Details

- Normalizing constant $z = Z(\lambda, \nu)$ and derivatives truncated, e.g.

$$Z(\lambda, \nu) \approx \sum_{j=0}^{J} \frac{\lambda^j}{(j!)^\nu},$$

$$\frac{\partial \log z}{\partial \lambda} \approx \frac{1}{z} \sum_{j=0}^{J} \frac{j \lambda^j}{(j!)^\nu}, \quad \frac{\partial \log z}{\partial \nu} \approx -\frac{1}{z} \sum_{j=0}^{J} \frac{\log(j!) \lambda^j}{(j!)^\nu}.$$

  We take $J = 100$.

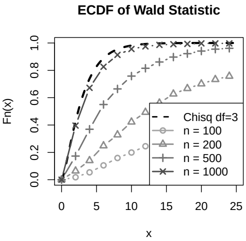
- We use the `R` function `nlminb` to maximize the likelihood subject to constraints. For the ZICMP regression model, we maximize $\log \mathcal{L}(\boldsymbol{\theta})$ subject to $\nu > 0$.

- Standard errors and confidence intervals for $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}, \hat{\nu}, \hat{\boldsymbol{\zeta}})$ are computed using approximate normality $\hat{\boldsymbol{\theta}} \sim \mathrm{N}(\boldsymbol{\theta}, \mathcal{I}_{\boldsymbol{\theta}}^{-1})$. We can estimate $\mathcal{I}_{\boldsymbol{\theta}}$ by $\mathcal{I}_{\hat{\boldsymbol{\theta}}}$.

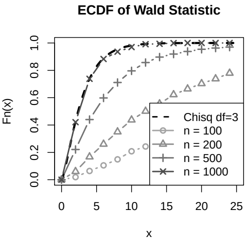
Large Sample Properties of the MLE

- We assess large sample properties of the MLE $\hat{\boldsymbol{\theta}} = (\hat{\lambda}, \hat{\nu}, \hat{p})$ through a simulation study.
- Draw an iid sample from ZICMP(λ, ν, p) where $\lambda = 2$, $p = 0.1$, $\nu \in \{0.25, 0.5, 0.75, 1, 2, 5, 10, 20, 30\}$, and $n \in \{100, 200, 500, 1000\}$.
- For each combination of parameters (λ, ν, p) and each n, $R = 1000$ samples of size n are drawn, and the MLE $\hat{\boldsymbol{\theta}}^{(r)}$ is computed on each sample, $r = 1, \ldots, R$.
- Wald statistics $W^{(r)} = (\hat{\boldsymbol{\theta}}^{(r)} - \boldsymbol{\theta})^T \mathcal{I}_{\boldsymbol{\theta}} (\hat{\boldsymbol{\theta}}^{(r)} - \boldsymbol{\theta})$ are then obtained for $r = 1, \ldots, R$.
- If $\hat{\boldsymbol{\theta}}$ follows the anticipated large sample N$(\boldsymbol{\theta}, \mathcal{I}_{\boldsymbol{\theta}}^{-1})$ distribution, the empirical CDF of $W^{(1)}, \ldots, W^{(R)}$ should approach the CDF of χ^2_3 as n becomes large.
- Recall that ZICMP(λ, ν, p) approaches a non-identifiable Zero-Inflated Bernoulli distribution with a singular FIM, which is likely to influence the W statistic.

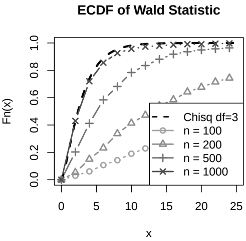
Large Sample Properties of the MLE



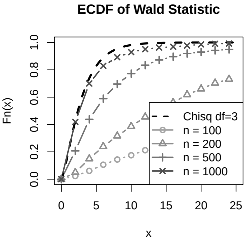
(a) $\nu = 0.25$



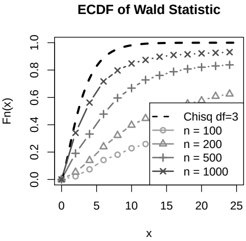
(b) $\nu = 0.5$



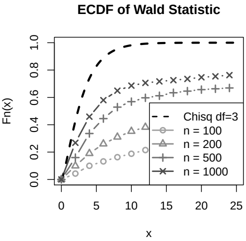
(c) $\nu = 0.75$



(d) $\nu = 1$

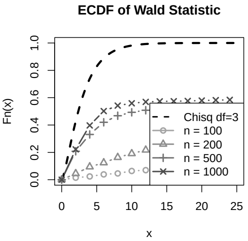


(e) $\nu = 2$

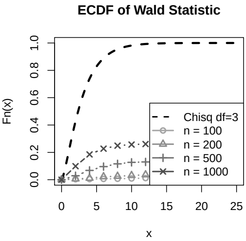


(f) $\nu = 3$

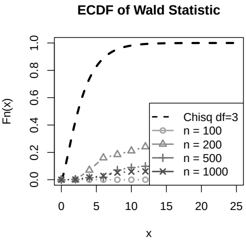
Large Sample Properties of the MLE



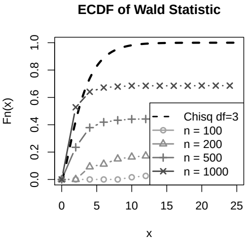
(a) $\nu = 4$



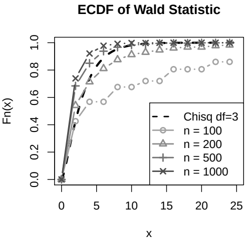
(b) $\nu = 5$



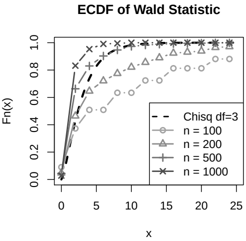
(c) $\nu = 7$



(d) $\nu = 10$



(e) $\nu = 20$



(f) $\nu = 30$

Test for Equidispersion

- Consider a level α test of $H_0 : \nu = 1$ vs $H_1 : \nu \neq 1$. Under H_0, ZICMP is restricted to ZIP.

- Let $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}, \hat{\nu}, \hat{\boldsymbol{\zeta}})$ be the unrestricted MLE, and let $\hat{\boldsymbol{\theta}}_0 = (\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\zeta}}_0)$ be the MLE under the ZIP distribution.

- The likelihood-ratio test (LRT) statistic is

$$-2 \log \Lambda = 2 \log \mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\nu}, \hat{\boldsymbol{\zeta}}) - 2 \log \mathcal{L}(\hat{\boldsymbol{\beta}}_0, \nu = 1, \hat{\boldsymbol{\zeta}}_0),$$

$$\text{where } \Lambda = \frac{\mathcal{L}(\hat{\boldsymbol{\beta}}_0, \nu = 1, \hat{\boldsymbol{\zeta}}_0)}{\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\nu}, \hat{\boldsymbol{\zeta}})}.$$

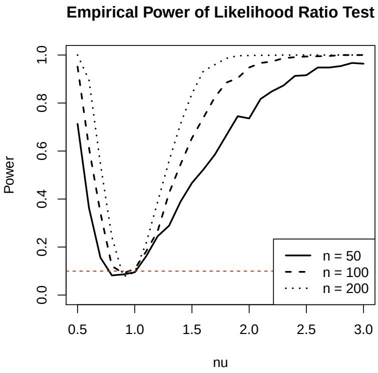
- Test procedure using the large sample distribution of the LRT $-2 \log \Lambda \sim \chi^2_1$ is

$$\text{Reject } H_0 \text{ if } -2 \log \Lambda \geq \chi^2_1(1 - \alpha),$$

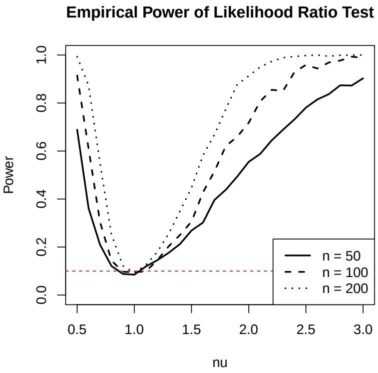
where $\chi^2_1(\xi)$ is the ξ quantile of the χ^2_1 distribution.

Empirical Power for LRT

- Draw 1000 iid samples of size $n \in \{50, 100, 200\}$ from ZICMP(λ, ν, p).
- Compute proportion of rejections for each setting.
- Choose $\alpha = 0.1$, $\lambda = 2$, $p \in \{0.01, 0.1\}$, and let ν vary.



(a) $p = 0.01$



(b) $p = 0.1$

Model Flexibility Study

- We illustrate the flexibility of ZICMP compared to several other count distributions.
- We randomly generate five datasets selected distributions; each contains 900 randomly drawn counts and 100 zeroes.
- Fit each of the competing models for each dataset.
- Compare models using AIC and a goodness-of-fit (GOF) statistic.

Model Flexibility Study

GOF Statistic

- Let $\mathrm{I}_1, \ldots, \mathrm{I}_K$ be a partition of $[0, \infty)$ and let density $g(\cdot \mid \boldsymbol{\theta})$ be a proposed model for the data.
- Let O_ℓ be the observed count on I_ℓ and $E_{\ell,\boldsymbol{\theta}}$ be the corresponding expected count under $g(\cdot \mid \boldsymbol{\theta})$, for $\ell = 1, \ldots, K$.
- To test the null hypothesis that the data are a random sample from $g(\cdot \mid \boldsymbol{\theta})$, a GOF statistic is

$$\mathrm{GOF}(\boldsymbol{\theta}) = \sum_{\ell=1}^{K} \frac{[O_\ell - E_{\ell,\boldsymbol{\theta}}]^2}{E_{\ell,\boldsymbol{\theta}}},$$

- When $\boldsymbol{\theta} \in \mathbb{R}^q$ is estimated by MLE, $\mathrm{GOF}(\hat{\boldsymbol{\theta}})$ follows a distribution between χ^2_{K-1-q} and χ^2_K under the null hypothesis (Sutradhar et al., 2008).
- Where possible, we merged the possible counts $\{0, 1, 2, \ldots\}$ into K categories $\mathrm{I}_1, \ldots, \mathrm{I}_K$ so that each $E_{\ell,\boldsymbol{\theta}} \geq 3$.

Sim Dist'n	ZIP	ZIG	ZINB	ZICMP
ZIG($p_* = 0.3$)	$\hat{\lambda}_*$: 3.272 (2.403) $\hat{p}$: 0.349 (0.507)	$\hat{p}_*$: 0.294 (0.312) $\hat{p}$: 0.113 (0.790)	$\hat{\lambda}$: 2.160 (0.198) $\hat{p}$: 0.014 (0.080) $\hat{\kappa}$: 1.359 (0.265)	$\hat{\lambda}$: 0.706 (1.407) $\hat{p}$: 0.113 (1.130) $\hat{\nu}$: 0.00 (1.108)
AIC	4434.942	3905.277	3904.460	3907.277
GOF	370.63, 6, 0.001	8.86, 11, 0.635	8.63, 10, 0.568	8.86, 10, 0.546
ZIP($\lambda = 3$)	$\hat{\lambda}_*$: 3.054 (1.984) $\hat{p}$: 0.086 (0.363)	$\hat{p}_*$: 0.264 (0.264) $\hat{p}$: 0.000 (0.698)	$\hat{\lambda}$: 3.044 (0.066) $\hat{p}$: 0.083 (0.013) $\hat{\kappa}$: 0.012 (0.019)	$\hat{\lambda}$: 2.930 (10.034) $\hat{p}$: 0.083 (0.430) $\hat{\nu}$: 0.970 (2.411)
AIC	3984.716	4379.247	3986.257	3986.564
GOF	8.87, 6, 0.181	392.34, 12, 0.001	8.88, 6, 0.180	8.99, 6, 0.174
"ZIB($\pi = 0.7$)" $=$ Ber($\pi = 0.63$)	$\hat{\lambda}_*$: 0.618 (1.493) $\hat{p}$: 0.000 (2.053)	$\hat{p}_*$: 0.618 (0.618) $\hat{p}$: 0.000 (2.058)	$\hat{\lambda}$: 0.618 (0.025) $\hat{p}$: 0.000 (0.000) $\hat{\kappa}$: 0.000 (0.000)	$\hat{\lambda}$: 1.638 (NA) $\hat{p}$: 0.005 (NA) $\hat{\nu}$: 33.325 (NA)
AIC	1834.846	2155.979	1836.846	1336.070
GOF	417.22, 2, 0.001	853.60, 3, 0.001	417.24, 1, 0.001	0.001, 1, 0.999
ZICMP ($\lambda = 8, \nu = 3$)	$\hat{\lambda}_*$: 1.505 (1.624) $\hat{p}$: 0.000 (0.707)	$\hat{p}_*$: 0.399 (0.399) $\hat{p}$: 0.0000 (1.052)	$\hat{\lambda}$: 1.505 (0.039) $\hat{p}$: 0.000 (0.001) $\hat{\kappa}$: 0.000 (0.001)	$\hat{\lambda}$: 6.513 (46.851) $\hat{p}$: 0.090 (0.585) $\hat{\nu}$: 2.721 (7.740)
AIC	2824.286	3374.160	2826.308	2701.204
GOF	104.02, 4, 0.001	624.68, 8, 0.001	104.05, 3, 0.001	6.57, 1, 0.010
ZICMP ($\lambda = 2, \nu = 0.25$)	$\hat{\lambda}_*$: 18.045 (4.483) $\hat{p}$: 0.102 (0.303)	$\hat{p}_*$: 0.054 (0.059) $\hat{p}$: 0.050 (0.325)	$\hat{\lambda}$: 18.041 (0.279) $\hat{p}$: 0.102 (0.010) $\hat{\kappa}$: 0.160 (0.010)	$\hat{\lambda}$: 1.987 (2.512) $\hat{p}$: 0.101 (0.303) $\hat{\nu}$: 0.245 (0.428)
AIC	8185.519	7597.370	6931.829	6924.219
GOF	2902.28, 21, 0.001	655.83, 45, 0.001	55.36, 36, 0.021	46.43, 34, 0.076

* Each entry in the GOF row lists three values: goodness-of-fit test statistic, degrees of freedom $K - 1 - q$, and resulting p-value.

** Very small values have been truncated to 0.001.

Analysis of Couples Data

- Loeys et al. (2012) investigated unwanted pursuit behaviors in separations between $n = 387$ couples.
- Outcome y_i is count of unwanted pursuit behaviors; 246 of 387 cases have $y_i = 0$.
- Covariates are education level x_{i1} and level of anxious attachment x_{i2}
- The dataset is overdispersed; the mean of $y_1, \ldots, y_n$ is 2.284 while the variance is 23.302.
- We compare the fit of several count regression models to this data.

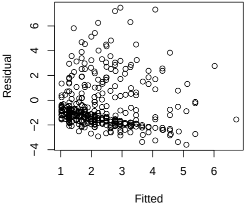
Analysis of Couples Data

	P	NB	CMP	ZIP	ZINB	ZICMP	ZIG
--- Count component ---							
(int)	0.817	0.855	−0.385	1.921	1.723	−0.160	1.770
	(0.044)	(0.155)	(0.055)	(0.044)	(0.150)	(0.077)	(0.122)
educ	−0.216	-0.353	-0.056	−0.350	−0.490	−0.068	−0.476
	(0.070)	(0.250)	(0.038)	(0.071)	(0.206)	(0.034)	(0.191)
anx	0.422	0.486	0.117	0.133	0.205	0.023	0.199
	(0.033)	(0.122)	(0.021)	(0.034)	(0.108)	(0.015)	(0.100)
--- Zero component ---							
(int)				0.673	0.340	0.418	0.422
				(0.142)	(0.210)	(0.167)	(0.159)
educ				-0.232	-0.459	-0.388	-0.416
				(0.222)	(0.297)	(0.268)	(0.271)
anx				−0.483	−0.520	−0.524	−0.503
				(0.111)	(0.147)	(0.133)	(0.135)
$\hat{\theta}$		0.194			0.821		
		(0.022)			(0.226)		
$\hat{\nu}$			0.000			0.000	
			(0.033)			(0.031)	
# params	3	4	4	6	7	7	6
log L	-1388.20	-638.96	-756.92	-802.45	-626.14	-627.17	-626.42
AIC	2782.4	1285.9	1521.84	1616.9	1266.3	1268.3	1264.8

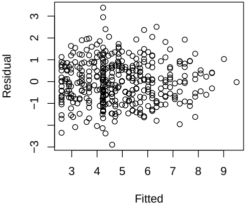
Red indicates significance at 0.05 level.

Analysis of Couples Data

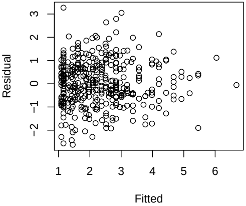
Randomized Quantile Residuals (Dunn and Smyth, 1996)



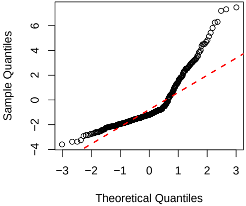
(a) Poisson



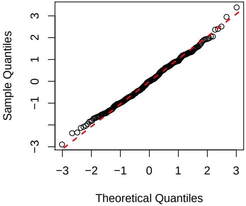
(b) ZINB



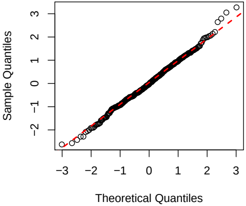
(c) ZICMP



(d) Poisson



(e) ZINB



(f) ZICMP

Conclusions and Future Work

- ZICMP regression was developed to model count data containing excess zeroes and either under- or overdispersion.
- For more details, see Sellers and Raim (2016) in CSDA.
- For very underdispersed datasets, use CMP regression to avoid identifiability issues.
- R code for ZICMP regression is available on request.
- Choo-Wosoba et al. (2016) extend ZICMP to handle longitudinal data with clustering.

`andrew.raim@census.gov`

References I

Hyoyoung Choo-Wosoba, Steven M. Levy, and Somnath Datta. Marginal regression models for clustered count data based on zero-inflated Conway-Maxwell-Poisson distribution with applications. *Biometrics*, 72(2):606–618, 2016.

R. W. Conway and W. L. Maxwell. A queuing model with state dependent service rates. *Journal of Industrial Engineering*, 12:132–136, 1962.

P. K. Dunn and G. K. Smyth. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5:236–244, 1996.

Felix Famoye and Karan P. Singh. Zero-inflated generalized Poisson regression model with an application to domestic violence data. *Journal of Data Science*, 4:117–130, 2006.

Daniel B. Hall. Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics*, 2000.

Joseph M. Hilbe. *Negative Binomial Regression*. Cambridge University Press, 2nd edition, 2011.

Diane Lambert. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14, February 1992.

References II

Tom Loeys, Beatrijs Moerkerke, Olivia DeSmet, and Ann Buysse. The analysis of zero-inflated count data: Beyond zero-inflated Poisson regression. *British Journal of Mathematical and Statistical Psychology*, 65(1):163–180, 2012.

T. J. Rothenberg. Identification in parametric models. *Econometrica*, 39:577–591, 1971.

K.F. Sellers and G. Shmueli. Data dispersion: Now you see it... now you don't. *Communications in Statistics - Theory and Methods*, 42:1–14, 2013.

Kimberly F. Sellers and Andrew M. Raim. A flexible zero-inflated model to address data dispersion. *Computational Statistics and Data Analysis*, 2016. (Available online).

Santosh C. Sutradhar, Nagaraj K. Neerchal, and Jorge G. Morel. A goodness-of-fit test for overdispersed binomial (or multinomial) models. *Journal of Statistical Planning and Inference*, 138(5):1459–1471, 2008.

Elise F. Zipkin, Jeffery B. Leirness, Brian P. Kinlan, Allan F. O'Connell, and Emily D. Silverman. Fitting statistical distributions to sea duck count data: Implications for survey design and abundance estimation. *Statistical Methodology*, 17:67–81, 2014.