The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
1-year subscription	$ 79	1-year subscription	$115
2-year subscription	$155	2-year subscription	$225
3-year subscription	$225	3-year subscription	$329
3-year subscription (electronic only)	$210	3-year subscription (electronic only)	$210
1-year student subscription	$ 48	1-year student subscription	$ 79
1-year university library subscription	$ 99	1-year university library subscription	$135
2-year university library subscription	$195	2-year university library subscription	$265
3-year university library subscription	$289	3-year university library subscription	$395
1-year institutional subscription	$225	1-year institutional subscription	$259
2-year institutional subscription	$445	2-year institutional subscription	$510
3-year institutional subscription	$650	3-year institutional subscription	$750

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The Stata Journal (2012)
12, Number 4, pp. 736–747

Modeling underdispersed count data with generalized Poisson regression

Tammy Harris
Institute for Families in Society
Department of Epidemiology and Biostatistics
University of South Carolina
Columbia, SC
harris68@mailbox.sc.edu

Zhao Yang
Quintiles, Inc.
Morrisville, NC
tonyyangsxz@gmail.com

James W. Hardin
Institute for Families in Society
Department of Epidemiology and Biostatistics
University of South Carolina
Columbia, SC
jhardin@sc.edu

Abstract. We present motivation and new Stata commands for modeling count data. While the focus of this article is on modeling data with underdispersion, the new command for fitting generalized Poisson regression models is also suitable as an alternative to negative binomial regression for overdispersed data.

Keywords: st0279, gpoisson, Poisson, count data, overdispersion, underdispersion

1 Introduction

We compare the effectiveness of regression models when dealing with underdispersed count data, and we introduce supporting Stata programs. Poisson regression analysis is widely used to model response variables comprising count data. The Poisson model assumes equidispersion, that is, that the mean and variance are equal. In practice, equidispersion is rarely reflected in data. In most situations, the variance exceeds the mean. This occurrence of extra-Poisson variation is known as overdispersion (see, for example, Dean [1992]). In situations for which the variance is smaller than the mean, data are characterized as underdispersed. Modeling underdispersed count data using inappropriate models can lead to overestimated standard errors and misleading inference. While there exist various approaches such as negative binomial distributions and other mixtures of Poisson (Yang et al. 2007) for modeling overdispersed count data, there are few models for underdispersed count data.

Winkelmann and Zimmermann (1994) proposed a generalized event count model that is appropriate for both overdispersed and underdispersed count data. Consul and Jain (1973) discussed the generalized Poisson (GP) distribution, which can also accommodate both overdispersed and underdispersed count data. The properties of the GP distribution are discussed in Consul (1989), Lerner, Lone, and Rao (1997), and Tuenter

(2000). Application of the GP regression model has been illustrated in Wang and Famoye (1997) for household fertility data and Cui, Kim, and Zhu (2006) for mapping quantitative trait loci.

Herein, we illustrate how to model underdispersed count data using the Poisson, the GP, and the quasi-Poisson (QP) regression models. This article is organized as follows. In section 2, we review appropriate count-data regression models. In section 3, Stata syntax is presented for the new command. A graphical illustration and real-world data example are contained in section 4, followed by a simulation study in section 5. Finally, a summary and conclusions are presented in section 6. Software, enhanced from Hardin and Hilbe (2012), was developed to fit GP models.

2 The models

For a random variable Y_i, we have a response vector $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$, where n is the sample size and Y_i and Y_j are independent and identically distributed for any $i \neq j$.

2.1 Poisson model

The most commonly used regression model for count data is the Poisson regression model, where covariates are included in the model via an invertible link function describing the relationship of the linear predictor $x_i\beta = \eta_i$ to the expected value of the responses θ_i:

$$f(y_i; \theta_i) = \frac{\theta_i^{y_i} e^{-\theta_i}}{y_i!}, y_i = 0, 1, 2, \ldots, \quad \theta_i > 0$$

The expected outcome in terms of the inverse of the log link function is given by $\theta_i = \exp(\mathbf{x}_i\beta)$, where $\mathbf{x}_i$ is a covariate vector and β is a vector of regression parameters to be estimated.

2.2 GP model

For equidispersed or for possibly overdispersed or underdispersed count data $\mathbf{Y}$, we may consider a regression model based on the GP distribution. This model assumes the response variable Y_i has probability mass function (PMF)

$$f(y_i; \theta_i, \delta) = \frac{\theta_i(\theta_i + \delta y_i)^{y_i - 1} e^{-\theta_i - \delta y_i}}{y_i!}, \quad y_i = 0, 1, 2, \ldots$$

where $\theta_i > 0$ and $\max(-1, -\theta_i/4) < \delta < 1$. The mean and variance of the GP random variable Y_i are given by

$$\mu_i = E(Y_i) = \frac{\theta_i}{1-\delta}, \quad \text{Var}(Y_i) = \frac{\theta_i}{(1-\delta)^3} = \frac{1}{(1-\delta)^2} E(Y_i) = \phi E(Y_i)$$

The term $\phi = 1/(1-\delta)^2$ plays the role of a dispersion factor. Clearly, when $\delta = 0$, there is equidispersion, and the GP distribution reduces to the usual Poisson distribution with parameter θ_i. Further, when $\delta > 0$, we have overdispersion in the model; when $\delta < 0$, we have underdispersion. In the current discussion, we are concerned with underdispersion, $\delta < 0$.

The associated log likelihood ($\mathcal{L}$) for the GP model is given by

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{n} \mathcal{L}(\theta_i, \delta; y_i) = \sum_{i=1}^{n} \ln L(\theta_i, \delta; y_i) \\ &= \sum_{i=1}^{n} \{\ln \theta_i + (y_i - 1)\ln(\theta_i + \delta y_i) - (\theta_i + \delta y_i) - \ln y_i!\}\end{aligned}$$

Consul and Famoye (1992) and Consul (1989) illustrated that covariates can be introduced into a regression model via the relationship

$$\log \frac{\theta_i}{1-\delta} = \sum_{r=1}^{p} x_{ir}\beta_r$$

where x_{ir} is the ith observation of the rth covariate, p is the number of covariates in the model, and β_r is the rth regression parameter.

2.3 QP model

Lastly, we consider the QP regression model for possibly overdispersed or underdispersed count data $\mathbf{Y}$. QP regression models are framed as generalized linear models. For the Poisson regression model, the expectation of Y_i is equal to the mean of the distribution, $E(Y_i) = \theta_i$. In generalized linear models terminology, the variance is equal to θ_i/δ. When fitting a Poisson model, $\delta = 1$. However, the variance of Y_i for the QP regression model is $\mathrm{Var}(Y_i) = \delta\theta_i$, where δ is assumed to be unknown. This generalization implies a quasi-likelihood (Wedderburn 1974). Hence, we assume for this QP model that $Y_i \sim$ QP(θ_i, δ).

3 The gpoisson command

3.1 Syntax

The basic syntax of `gpoisson` is equivalent to that of the `poisson` command (see [R] **poisson**):

`gpoisson` *depvar* [*indepvars*] [*if*] [*in*] [*weight*] [`, noconstant`
`exposure(`*varname_e*`)` `offset(`*varname_o*`)` `constraints(`*constraints*`)`
`collinear` `vce(`*vcetype*`)` `level(`#`)` `irr` `nocnsreport` *display_options*
maximize_options `coeflegend`]

After `gpoisson` estimation, `gpoisson postestimation` is available. Dialog boxes (see [R] **db**) are available for `gpoisson` and `predict` after `gpoisson`.

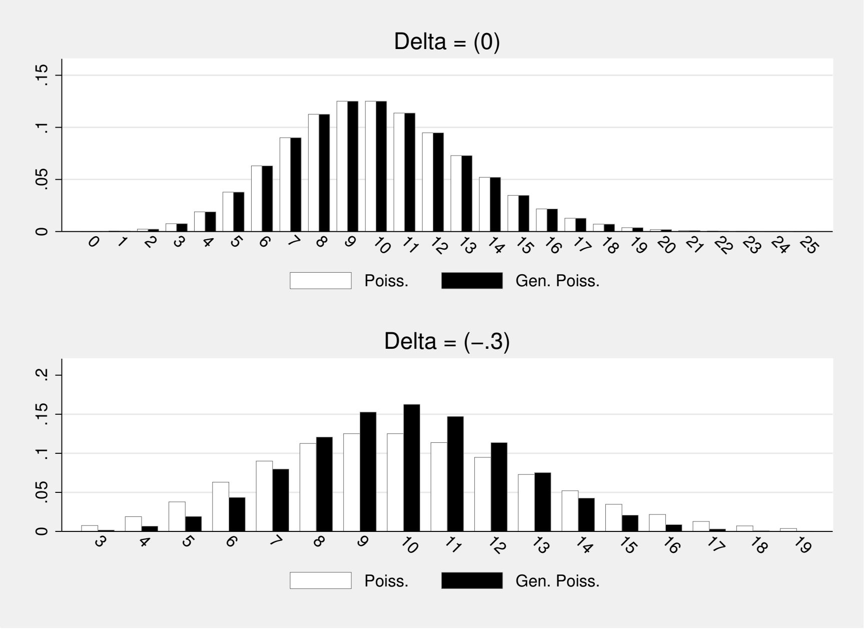
4 Example

In this section, we first present a graphical illustration of the Poisson and GP distributions. In the graphs, we juxtapose the distributions when equivalent and when there is significant underdispersion. Such a presentation illustrates the heavier tails of the Poisson distribution, which then leads to the discoveries of a substantial simulation study in section 5.

Second, we present a real-world data example using information collected on U.S. firms. These data are notable for exhibiting moderate underdispersion and for having relatively few zero outcomes.

4.1 A graphical comparison of Poisson and GP

To illustrate the difference in distributions used for calculating probabilities and likelihoods, we present below two graphs of the GP and Poisson PMFs. In each bar graph, the means are the same. The top figure is for $\delta = 0.0$, which means that the distributions are the same. As $|\delta|$ increases, the two distributions become more distinct. The bottom figure is for $\delta = -0.3$. In that bar graph, one can see that the Poisson PMF is not as peaked and has thicker tails.



Thus, as δ differs from zero, so does the GP distribution differ from the Poisson. In the subsequent section, we investigate coverage probabilities and power affected by the differences in distributions.

4.2 A real-world data analysis example

Of 126 U.S. firms that were targets of tender offers, Jaggia and Thosar (1993) studied the determinants of the number of bids received. Using relationships outlined by the authors, we model the number of bids received (over a 52-week period following the initial bid) as a function of defensive actions of the firm in these terms: whether there were defensive actions (lawsuits) from the firm's management, `leglrest`; whether the firm is under any real restructuring, `realrest`; whether the firm proposed a change in ownership, `finrest`; whether the firm invited a friendly third-party bid, `whtknght`; the bid premium (bid price divided by the price 14 days before the bid was issued), `bidprem`; the percentage of stock held by institutions, `insthold`; the size (book value of assets) of the firm measured in billions of dollars, `size`; the square of the size of the firm, `sizesq`; and whether the Department of Justice intervened, `regulatn`.

The results of fitting a GP model to the outcomes are given by

```
. use jaggia
. generate sizesq = size*size
. gpoisson numbids leglrest realrest finrest whtknght bidprem insthold size
> sizesq regulatn, irr nolog
Generalized Poisson regression                     Number of obs   =        126
                                                   LR chi2(9)      =      39.90
Dispersion     = -.1812696                         Prob > chi2     =     0.0000
Log likelihood = -181.12051                        Pseudo R2       =     0.0992

------------------------------------------------------------------------------
     numbids |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    leglrest |   1.281231    .163009     1.95   0.051     .9984595    1.644086
    realrest |   .8556441   .1395632    -0.96   0.339     .6215192    1.177963
     finrest |   1.157229   .2034714     0.83   0.406     .8198899    1.633364
    whtknght |   1.717354   .2377103     3.91   0.000     1.309301    2.252581
     bidprem |   .4588859   .1492794    -2.39   0.017     .2425503    .8681757
    insthold |   .6673574   .2434747    -1.11   0.268     .3264459    1.364287
        size |   1.208949   .0610227     3.76   0.000     1.095072    1.334667
      sizesq |    .991915   .0026968    -2.99   0.003     .9866436    .9972147
    regulatn |   .9985279    .133861    -0.01   0.991     .7678025    1.298586
       _cons |   2.877161   1.329555     2.29   0.022     1.163104    7.117215
-------------+----------------------------------------------------------------
 /atanhdelta |  -.1832951   .0640189                     -.3087699   -.0578203
-------------+----------------------------------------------------------------
       delta |  -.1812696   .0619154                     -.2993176   -.0577559
------------------------------------------------------------------------------
Likelihood-ratio test of delta=0:  chi2(1) =     7.66       Prob>=chi2 = 0.0028
```

where `/atanhdelta` refers to the inverse hyperbolic tangent function of δ shown below:

$$\texttt{/atanhdelta} = \tanh^{-1}(\delta) = \frac{1}{2}\ln\left(\frac{1+\delta}{1-\delta}\right)$$

The coefficient was surprisingly positive (associated incidence-rate ratio [IRR] surprisingly greater than 1) on the indicator of whether the firm had filed lawsuits in defense of the initial bid. That is, those firms who had filed lawsuits to protect against the initial bid received additional bids at 1.28 times the rate of those firms who did not file lawsuits against the initial bid; this result is marginally significant because the associated p-value is 0.051. The invitation to a friendly bidder to participate leads to a rate of subsequent bids that is 1.7 times the rate of accumulated bids of those firms who do not invite a friendly bidder. The size of the firm also matters, where results indicate that the rate of bids increases at first but then decreases as the firm gets larger. As expected, the rate of bids decreased as the bid premium increased (IRR = 0.46). Finally, there is evidence of significant underdispersion.

5 Simulation study

We compare the coverage probabilities and power when data are underdispersed for Poisson, GP, and QP regression models. The model used in our simulations was

$$\log \mu_i = \log \left(\frac{\theta_i}{1-\delta} \right) = 0.5 - 0.25x_1 - 0.25x_2 + 0x_3 \tag{1}$$

where we synthesize x_1 from a Bernoulli(0.5) distribution, x_2 from a Bernoulli(0.6) distribution, and x_3 from a Bernoulli(0.4) distribution. However, only x_1 and x_2 are used in the generation of the linear predictor (effectively assuming a zero coefficient for x_3). For each covariate coefficient in (1), the coverage probabilities of the confidence intervals were computed. Therefore, the proportion of time that the intervals contain the true parameter estimates was calculated using the Poisson, GP, and QP regression models.

Samples of size $n = 30, 40, 50, 60, 70, 80, 90, 100, 1000$ were generated for a series of dispersion parameter values. Simulation results for each size are based on 10,000 replications for each situation. Coverage probability results are summarized in tables 1a, 1b, and 1c; power for covariate significance is summarized in table 2. When modeling underdispersed count data, the GP and QP models have greater power to determine significance of the covariates than does the Poisson regression model. Furthermore, the GP regression model is the most effective model (of the three models considered) in establishing the significance of covariates over a wide range of underdispersion. In the case of equidispersion ($\delta = 0$), all models performed well, with the GP model having slightly elevated type I error rates over the QP model.

Table 1a. Coverage probabilities (10,000 replications) for the true parameter estimates [from (1)] using the Poisson, GP, and QP regression models; $\delta = 0$ indicates equidispersion, $\delta < 0$ indicates underdispersion. Values in table are percentages.

			Dispersion parameter (δ)				
n	Model	x_i	0.0	−0.1	−0.2	−0.3	−0.4
30	Poisson	x_1	95.21	97.30	98.09	99.11	99.42
		x_2	95.38	96.98	98.15	99.11	99.28
		x_3	95.21	97.14	97.98	98.91	99.51
	GP	x_1	92.11	92.49	91.70	92.17	92.36
		x_2	91.83	92.02	91.53	92.09	91.93
		x_3	92.33	92.46	91.83	91.68	92.63
	QP	x_1	94.09	94.51	94.02	94.31	94.48
		x_2	94.18	94.04	93.92	94.30	94.15
		x_3	94.35	94.44	93.78	93.83	94.39
40	Poisson	x_1	94.98	97.10	98.18	99.03	99.37
		x_2	94.84	96.80	98.08	98.81	99.51
		x_3	95.28	96.57	98.08	98.84	99.39
	GP	x_1	92.92	93.01	92.67	92.64	92.65
		x_2	92.71	92.65	92.75	92.74	93.00
		x_3	93.22	92.47	92.70	92.81	92.37
	QP	x_1	94.37	94.38	94.26	94.43	94.26
		x_2	94.28	94.10	94.10	94.35	94.54
		x_3	94.80	93.88	94.41	94.29	94.47
50	Poisson	x_1	94.87	96.58	98.12	98.87	99.37
		x_2	94.56	96.97	98.03	98.84	99.44
		x_3	95.03	96.99	98.38	99.10	99.46
	GP	x_1	93.02	92.97	93.51	93.12	92.76
		x_2	92.86	93.23	93.19	93.26	93.38
		x_3	93.30	93.45	93.54	93.60	93.15
	QP	x_1	94.25	94.07	94.61	94.49	94.08
		x_2	94.00	94.38	94.51	94.31	94.80
		x_3	94.39	94.51	94.79	94.70	94.54

Table 1b. Coverage probabilities (10,000 replications) for the true parameter estimates [from (1)] using the Poisson, GP, and QP regression models; $\delta = 0$ indicates equidispersion, $\delta < 0$ indicates underdispersion. Values in table are percentages.

			Dispersion parameter (δ)				
n	Model	x_i	0.0	−0.1	−0.2	−0.3	−0.4
60	Poisson	x_1	95.12	96.83	98.00	98.88	99.29
		x_2	95.25	96.72	98.04	98.83	99.38
		x_3	94.95	97.02	98.13	98.82	99.41
	GP	x_1	93.53	93.46	93.05	93.45	93.58
		x_2	93.92	93.18	93.35	93.15	93.54
		x_3	93.46	93.79	93.21	93.33	93.14
	QP	x_1	94.68	94.65	94.21	94.63	94.42
		x_2	94.71	94.01	94.44	94.31	94.30
		x_3	94.51	94.76	94.23	94.58	94.25
70	Poisson	x_1	94.96	97.13	98.21	98.99	99.39
		x_2	95.31	96.86	97.95	99.01	99.40
		x_3	95.10	96.85	98.01	99.10	99.46
	GP	x_1	94.07	94.06	93.86	93.65	93.82
		x_2	94.15	93.78	93.19	93.85	93.66
		x_3	93.91	93.86	93.44	94.08	93.67
	QP	x_1	94.76	94.81	94.94	94.36	94.50
		x_2	94.89	94.67	94.01	94.79	94.93
		x_3	94.66	94.64	94.34	94.78	94.94
80	Poisson	x_1	95.15	97.28	98.16	99.00	99.46
		x_2	94.89	97.00	98.07	98.89	99.21
		x_3	95.22	96.97	98.11	98.99	99.41
	GP	x_1	94.20	94.16	94.01	94.04	93.66
		x_2	94.07	94.18	93.57	93.89	93.47
		x_3	94.19	94.21	93.82	93.94	93.44
	QP	x_1	94.92	95.09	94.89	94.90	94.50
		x_2	94.68	94.92	94.49	94.63	94.35
		x_3	94.83	94.83	94.62	94.54	94.30

Table 1c. Coverage probabilities (10,000 replications) for the true parameter estimates [from (1)] using the Poisson, GP, and QP regression models; $\delta = 0$ indicates equidispersion, $\delta < 0$ indicates underdispersion. Values in table are percentages.

			Dispersion parameter (δ)				
n	Model	x_i	0.0	−0.1	−0.2	−0.3	−0.4
90	Poisson	x_1	95.33	97.10	97.92	98.74	99.37
		x_2	95.45	96.77	98.11	99.00	99.40
		x_3	95.11	96.91	98.18	98.98	99.54
	GP	x_1	94.36	94.08	93.77	94.12	94.44
		x_2	94.32	93.93	94.11	94.10	94.03
		x_3	94.38	94.24	94.01	93.94	94.06
	QP	x_1	94.83	94.81	94.44	94.73	95.10
		x_2	94.98	94.52	94.55	94.70	94.62
		x_3	94.91	94.80	94.77	94.70	94.85
100	Poisson	x_1	95.02	96.99	98.08	98.77	99.36
		x_2	94.95	96.68	98.03	98.86	99.37
		x_3	95.00	97.31	98.27	99.00	99.30
	GP	x_1	93.98	94.34	94.20	94.08	94.18
		x_2	94.20	94.02	94.18	94.07	94.31
		x_3	94.29	94.57	94.18	94.21	94.09
	QP	x_1	94.51	94.84	94.70	94.75	94.76
		x_2	94.72	94.51	94.80	94.81	95.01
		x_3	94.82	95.21	95.01	95.01	94.83
1000	Poisson	x_1	95.46	96.71	97.91	98.94	99.30
		x_2	95.21	96.99	98.23	98.88	99.39
		x_3	95.08	96.97	98.13	98.81	99.52
	GP	x_1	95.33	94.99	94.74	94.98	94.83
		x_2	94.99	95.11	94.69	94.87	95.02
		x_3	95.04	94.99	95.14	94.99	95.04
	QP	x_1	95.39	95.05	94.87	94.97	94.75
		x_2	95.07	95.15	94.68	95.20	94.98
		x_3	95.07	95.04	95.18	94.98	95.20

Table 2. Power (from 10,000 replications) in determining significance of covariates $[E(Y) \sim \exp(4 + Bx_1 - 0.25x_2 + 0x_3)$, where $x_1 \sim \text{Uniform}(0.5)]$ when coefficients are assessed using Poisson, GP, and QP regression models. Values in table are percentages.

		Model		
n	B coef.	Poisson	GP	QP
30	0.00	0.94	7.66	5.83
	0.02	1.87	10.87	8.65
	0.04	5.52	20.22	16.81
	0.06	12.31	33.77	29.15
	0.08	23.06	50.02	45.01
	0.10	38.81	66.51	61.67
	0.12	54.90	79.85	75.62
	0.14	72.14	89.94	87.18
	0.16	84.62	95.45	94.17
	0.18	92.53	98.16	97.44
	0.20	96.59	99.19	98.95
	0.22	98.47	99.73	99.62
	0.24	99.59	99.92	99.89
60	0.00	1.06	6.37	5.75
	0.02	3.27	12.59	11.18
	0.04	11.15	29.31	27.13
	0.06	28.95	54.35	51.68
	0.08	54.55	77.27	75.04
	0.10	77.28	92.24	91.31
	0.12	91.45	97.89	97.56
	0.14	97.37	99.44	99.37
	0.16	99.47	99.94	99.93
	0.18	99.93	99.97	99.97
	0.20	99.99	100.00	100.00
	0.22	100.00	100.00	100.00
	0.24	100.00	100.00	100.00

6 Discussion and conclusions

In this article, we introduced a supporting Stata program and illustrated the effectiveness of three Poisson regression models (Poisson, GP, and QP) when dealing with underdispersed count data. Underdispersion is a less explored occurrence in modeling count data. Underdispersion can lead to overestimated standard errors and misleading conclusions if handled inappropriately. We simulated underdispersed count data and compared inferences from three Poisson regression models by using synthesized covariates and different equidispersion and underdispersion parameters.

7 References

Consul, P. C. 1989. *Generalized Poisson Distributions: Properties and Applications*. New York: Dekker.

Consul, P. C., and F. Famoye. 1992. Generalized Poisson regression model. *Communications in Statistics—Theory and Methods* 21: 89–109.

Consul, P. C., and G. C. Jain. 1973. A generalization of the Poisson distribution. *Technometrics* 15: 791–799.

Cui, Y., D.-Y. Kim, and J. Zhu. 2006. On the generalized Poisson regression mixture model for mapping quantitative trait loci with count data. *Genetics* 174: 2159–2172.

Dean, C. B. 1992. Testing for overdispersion in Poisson and binomial regression models. *Journal of the American Statistical Association* 87: 451–457.

Hardin, J. W., and J. M. Hilbe. 2012. *Generalized Linear Models and Extensions*. 3rd ed. College Station, TX: Stata Press.

Jaggia, S., and S. Thosar. 1993. Multiple bids as a consequence of target management resistance: A count data approach. *Review of Quantitative Finance and Accounting* 3: 447–457.

Lerner, B., A. Lone, and M. Rao. 1997. On generalized Poisson distributions. *Probability and Mathematical Statistics* 17: 377–385.

Tuenter, H. J. H. 2000. On the generalized Poisson distribution. *Statistica Neerlandica* 54: 374–376.

Wang, W., and F. Famoye. 1997. Modeling household fertility decisions with generalized Poisson regression. *Journal of Population Economics* 10: 273–283.

Wedderburn, R. W. M. 1974. Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika* 61: 439–447.

Winkelmann, R., and K. F. Zimmermann. 1994. Count data models for demographic data. *Mathematical Population Studies* 4: 205–221.

Yang, Z., J. W. Hardin, C. L. Addy, and Q. H. Vuong. 2007. Testing approaches for overdispersion in Poisson regression versus the generalized Poisson model. *Biometrical Journal* 49: 565–584.

About the authors

Tammy Harris is a PhD candidate in the Department of Epidemiology and Biostatistics and an affiliated researcher in the Institute for Families in Society at the University of South Carolina in Columbia, SC.

Zhao Yang is a senior statistician at Quintiles, Inc. in Morrisville, NC.

James W. Hardin is an associate professor in the Department of Epidemiology and Biostatistics and an affiliated faculty in the Institute for Families in Society at the University of South Carolina in Columbia, SC.