Errata for J. Scott Long, 1997, Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage.

April 2, 2001

1 Errata corrected in Second Printing

Pg. 9, line 2: " xxv to xxvii" should be " xxvii to xxx".

Pg. 20, Table 2.2: In the footer it should read: "t is a t-test of β".

Pg. 109, Table 4.4: Replace the labels "Row" with "Row %".

Pg. 143, line 6: Replace "as chi-square with $K(J-1)$" with "as chi-square with $K(J-2)$".

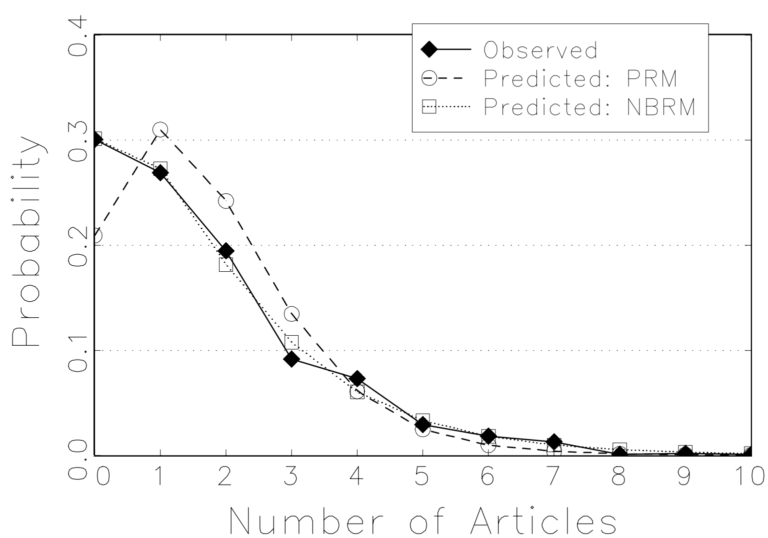
Pg. 177, line 16: Replace "between strongly agreeing and all other categories" with "between strongly disagreeing and all other categories,".

Pg. 226, immediately after Equation 8.6: Replace "$\hat{\mu} = \mathrm{x}\hat{\beta}$" with "$\hat{\mu} = \exp(\mathrm{x}\hat{\beta})$".

Pg. 229, line 10: Replace "expected probability" with "expected productivity".

Pg. 229, line 13: Replace "1.76 to 1.43" with "1.79 to 1.43".

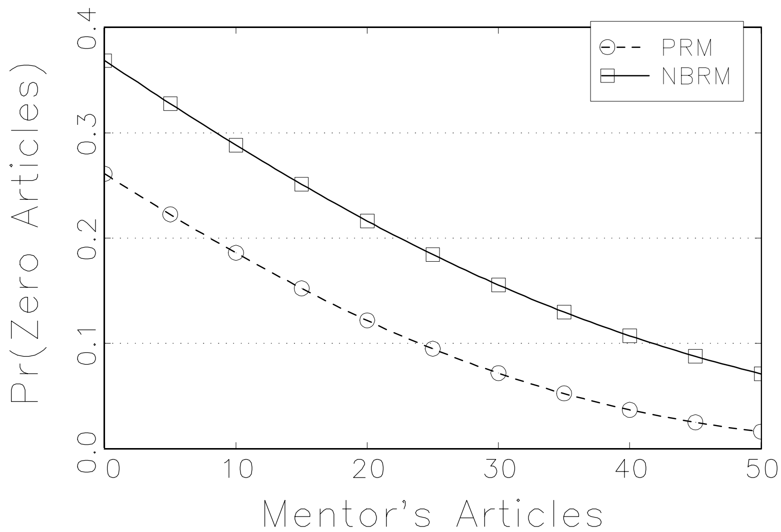
Page 229: Replace Figure 8.4 with:



Page 236-237: Replace first paragraph in Section 8.3.3 with:

It is important to test for overdispersion if you use the PRM. Even with the correct specification of the mean structure, estimates from the PRM when there is overdispersion are inefficient with standard errors that are biased downward (Cameron and Trivedi 1986). If software is available to estimate the NBRM, a one-tailed z-test of $H_0\colon \alpha = 0$ can be used to test for overdispersion, since when α is zero the NBRM reduces to the PRM. Or, a LR test can be computed. If $\ln L_{\mathrm{PRM}}$ is the log likelihood from the PRM and $\ln L_{\mathrm{NBRM}}$ is the log likelihood from the NBRM, then $G^2 = 2(\ln L_{\mathrm{NBRM}} - \ln L_{\mathrm{PRM}})$ is a test of $H_0\colon \alpha = 0$. To test at the p level of significance, a critical value of X^2_{2p} should be used since α must be positive. Cameron and Trivedi (1990) present several tests based on the residuals from the PRM that do not require estimation of the NBRM..

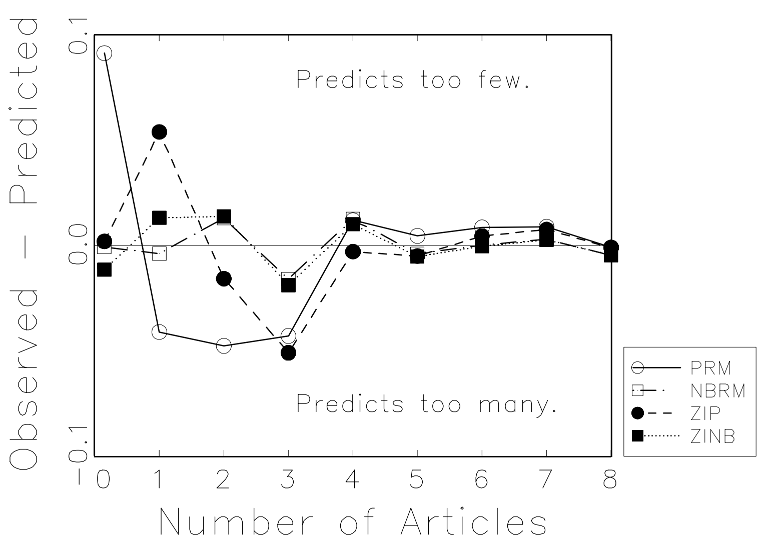
Page 238: Replace Figure 8.8 with:



Page 247: Replace paragraph starting " Figure 8.9 plots..." with:

> Figure 8.9 plots the difference between the observed proportions for each count and the mean probability from the four models. We see immediately that the major failure of the PRM is in predicting the number of zeros, with an under prediction of about .1. The ZIP does much better at predicting zeros, but has poor predictions for counts one through three. The NBRM predicts the zeros very well and also has much better predictions for the counts from one to three. The ZINB slightly over predicts zeros and under predicts ones, with similar predictions to the NBRM for other counts. Overall, the NBRM model provides the most accurate predictions, which are slightly better than those for the ZINB.

Page 248: Replace Figure 8.9 with:



Page 249: Replace first paragraph starting " Overall, these tests,... "

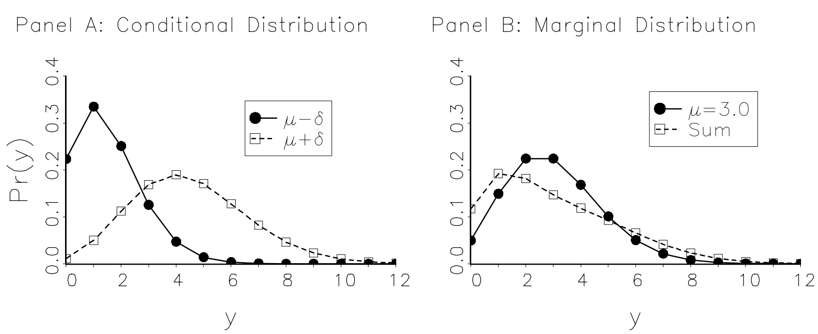
Overall, these tests provide evidence that the ZINB models fits the data best. However, when fitting a series of models without any theoretical rationale, it is easy to overfit the data. In our example, the most compelling evidence for the ZINB is that it makes substantive sense. Within science, there are scientists who for structural reasons will not publish. Other scientist do not publish only by chance in a given period. This is the basis of the zero inflated models. The NB version of the model seems preferable since it is likely that there are unobserved sources of heterogeneity that differentiate the scientists. Overall, the ZINB makes substantive sense and fits the data well. However, the NBRM also fits extremely well.

Pg. 264, last equation: Replace "$\ln\left(\frac{y}{1+y}\right)$" with "$\ln\left(\frac{y}{1-y}\right)$".

Pg. 266: Replace "Page 56" with "Page 55".

Pg. 269, regarding Page 143: Replace entire paragraph with "Excluding the intercept, $\boldsymbol{\beta}_m$ has K coefficients for each of the $J-1$ binary logits, for a total of $(J-1)K$ coefficients; $\boldsymbol{\beta}$ has K coefficients excluding the intercept. Therefore, we are imposing $(J-1)K - K = K(J-2)$ constraints."

Page 272: Replace figure with:



2 Errata to be corrected in fourth printing

Page 38, Table 3.2: For K618 the estimate of β^{S_x} is -.015 (not -.115).

Page 79, bullet 2 at top: Replace "holding all other variables constant." with "holding all other variables at their means."

Page 91, last equation. Remove square from $\hat{\sigma}^2_{\hat{\beta}_k}$. That is, replace

$$W = \sum_{k=1}^{2} \frac{\hat{\beta}_k^2}{\hat{\sigma}^2_{\hat{\beta}_k}} = \sum_{k=1}^{2} \left(\frac{\hat{\beta}_k}{\hat{\sigma}^2_{\hat{\beta}_k}} \right)^2 = \sum_{k=1}^{2} z^2_{\hat{\beta}_k}$$

with

$$W = \sum_{k=1}^{2} \frac{\hat{\beta}_k^2}{\hat{\sigma}^2_{\hat{\beta}_k}} = \sum_{k=1}^{2} \left(\frac{\hat{\beta}_k}{\hat{\sigma}_{\hat{\beta}_k}} \right)^2 = \sum_{k=1}^{2} z^2_{\hat{\beta}_k}$$

Page 108, line 5 in section Examples of Count Measures: 0's and 1's have been switch. Sentence should read: "They show that the model is more effective at predicting 1's (80% are predicted correctly) than 0's (55% are predicted correctly).

Pages 143 and 144, points 2, 3 and 4 have errors in the dimensions of the matrices:

1. No changes to this point.

2. *Estimate the covariance between* $\underline{\widehat{\boldsymbol{\beta}}}_m$ *and* $\underline{\widehat{\boldsymbol{\beta}}}_\ell$, where the underline indicates that the constant has been removed from the vector. Define

$$w_{im\ell} = \widehat{\pi}_\ell(\mathbf{x}_i) - \widehat{\pi}_m(\mathbf{x}_i)\,\widehat{\pi}_\ell(\mathbf{x}_i)$$

and let $\mathbf{W}_{m\ell}$ be a $N \times N$ diagonal matrix whose ith element is $w_{im\ell}$. Let $\mathbf{X}$ be the $N \times (K+1)$ matrix with 1's in the first column and the independent variables in the remaining columns. Brant shows that covariances among estimates from the different binary equations, $\widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_m, \underline{\widehat{\boldsymbol{\beta}}}_\ell\right)$, are estimated by deleting the first row and column of:

$$\left(\mathbf{X}'\mathbf{W}_{mm}\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{W}_{m\ell}\mathbf{X}\right)\left(\mathbf{X}'\mathbf{W}_{\ell\ell}\mathbf{X}\right)^{-1}$$

3. *Combine all estimates.* Define $\widehat{\boldsymbol{\beta}}^* = \left(\underline{\widehat{\boldsymbol{\beta}}}'_1\ \underline{\widehat{\boldsymbol{\beta}}}'_2 \cdots \underline{\widehat{\boldsymbol{\beta}}}'_{J-1}\right)'$ and

$$\widehat{Var}\left(\widehat{\boldsymbol{\beta}}^*\right) = \begin{pmatrix} \widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_1\right) & \cdots & \widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_1, \underline{\widehat{\boldsymbol{\beta}}}_{J-1}\right) \\ \vdots & \ddots & \vdots \\ \widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_{J-1}, \underline{\widehat{\boldsymbol{\beta}}}_1\right) & \cdots & \widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_{J-1}\right) \end{pmatrix}$$

The diagonal elements $\widehat{Var}\left(\underline{\widehat{\boldsymbol{\beta}}}_m\right)$ are the covariance matrices from each binary regression. The off-diagonal elements were defined in step 2.

4. *Construct the Wald test of* H_0: $\underline{\boldsymbol{\beta}}_1 = \cdots = \underline{\boldsymbol{\beta}}_{J-1}$. This hypothesis corresponds to: H_0: $\mathbf{D}\boldsymbol{\beta}^* = \mathbf{0}$, where

$$\mathbf{D} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & -\mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{I} \end{pmatrix}$$

$\mathbf{I}$ is a $K \times K$ identity matrix and $\mathbf{0}$ is a $K \times K$ matrix of zeros. (*Verify that this matrix results in the appropriate linear combination for the hypothesis being tested.*) The Wald statistic takes the standard form discussed in Chapter 4:

$$W = \left(\mathbf{D}\widehat{\boldsymbol{\beta}}^{*}\right)' \left[\mathbf{D}\,\widehat{Var}\left(\widehat{\boldsymbol{\beta}}^{*}\right)\mathbf{D}'\right]^{-1} \left(\mathbf{D}\widehat{\boldsymbol{\beta}}^{*}\right)$$

with $(J-2)K$ degrees of freedom.

5. No changes.

Page 185, line 9: Change subscript to ϕ_m. That is, replace

$$\Pr\left(y = m \mid \mathbf{x}\right) = \exp\left(\mathbf{x}\boldsymbol{\beta}\phi_m\right) / \sum_{j=1}^{J} \exp\left(\mathbf{x}\boldsymbol{\beta}\phi_m\right)$$

with:

$$\Pr\left(y = m \mid \mathbf{x}\right) = \exp\left(\mathbf{x}\boldsymbol{\beta}\phi_m\right) / \sum_{j=1}^{J} \exp\left(\mathbf{x}\boldsymbol{\beta}\phi_j\right)$$

Page 224, Second equation from bottom remove extra $\exp(\beta_0)$ That is, replace

$$E\left(y \mid \mathbf{x}, x_k\right) = \exp\left(\beta_0\right)\exp\left(\beta_0\right)\exp\left(\beta_1 x_1\right)\cdots\exp\left(\beta_k x_k\right)\cdots\exp\left(\beta_K x_K\right)$$

with

$$E\left(y \mid \mathbf{x}, x_k\right) = \exp\left(\beta_0\right)\exp\left(\beta_1 x_1\right)\cdots\exp\left(\beta_k x_k\right)\cdots\exp\left(\beta_K x_K\right)$$

Page 278, Kaufman references: Replace "*Social Science Research*" with "*Social Science Quarterly*".

Page 233, 2nd line after [8.15]: Add a space. That is, replace "What must happen toν to..." with "What must happen to ν to..."

Page 239, 2nd line after [8.19]: delete "zero". That is, replace "...given that $y > 0$ zero is computed..." with "...given that $y > 0$ is computed...".

Page 244, 2nd paragraph: Replace "For the both the..." with "For both the...".

Page 244, last equation: Replace

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma} \mid \mathbf{y}, \mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{N} \Pr\left(y_i \mid \mathbf{x}_i, \mathbf{z}_i\right)$$

with

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma} \mid \mathbf{y}, \mathbf{X}, \mathbf{Z}) = \prod_{i=1}^{N} \Pr\left(y_i \mid \mathbf{x}_i, \mathbf{z}_i\right)$$