Errata for *Applied Regression Analysis and Generalized Linear Models, Third Edition* (Sage, 2016)

John Fox

updated: 2023-09-26

Please note that some (or all) of these errors may be corrected in your printing of the book. Some of these errors (e.g., to cross-references and page references) apparently were introduced in the final typesetting of the book, and should not have been possible in a LaTeX document.

I'm particularly grateful to Likan Zhan, who has contributed a number of the errata; rather than thanking him individually for each, I've simply noted his name in parentheses following his corrections.

1. p. 56, Figure 4.1: The caption should read "$X^{(0)}$ is $\log_e(X)$" *not* "$X^{(0)}$ is $\log_e(p)$." (Likan Zhan)

2. p. 79, Exercise 4.4: The second line of the equation defining the Yeo-Johnson family of transformation is missing a minus-sign. The equation should read:

$$X \to X^{[p]} \equiv \begin{cases} (X+1)^{(p)} & \text{for } X \geq 0 \\ -(1-X)^{(2-p)} & \text{for } X < 0 \end{cases}$$

   (I'm grateful to Phil Hoon Oh for pointing out this error.)

   As well, the exercise is more effective if in part (a) the range of X is between (say) -4 and 4 rather than -10 and 10, and if the vertical axis is cut off at (say) -20. Similarly, in part (b), it's better to let X range from (say) 0.5 to 5 rather than from 0.1 to 10.

3. p. 118: Below the first equation, "the conditional mean of Y is a linear function of X" should read "the conditional mean of Y is a linear function of X_1". (I thank Kristoph Steikert for finding this error.)

4. p. 173, Table 8.3: The H_0 given for $SS(\beta|\alpha, \gamma)$ should read simply "all $\beta_k = 0$ ($\mu_{\cdot k} = \mu_{\cdot k'}$)" *not* "all $\beta_k = 0$ ($\mu_{\cdot k} = \mu_{\cdot k'}$) | no interaction." As well, the vertical alignment of the lines for $SS(\alpha|\beta)$ and $SS(\beta|\alpha)$ is poor: each of these SSs should be moved down one line in the table (to appear on the same lines as "all ...). (Likan Zhan)

5. p. 199: The exercise labeled 8.10 (d) is actually 8.10 (b).

6. p. 217: The full-model likelihood, given in an unnumbered equation near the middle of the page, should be denoted L_1, not L. (Likan Zhan)

7. p. 217: The exponent in the second expression for the likelihood ratio, in an unnumbered equation below the middle of the page, is incorrect; it should be $n/2$, not $2/n$. The equation should read

$$\frac{L_0}{L_1} = \left(\frac{\mathbf{e}_0'\mathbf{e}_0}{\mathbf{e}'\mathbf{e}}\right)^{-n/2} = \left(\frac{\mathbf{e}'\mathbf{e}}{\mathbf{e}_0'\mathbf{e}_0}\right)^{n/2}$$

   (I'm grateful to Benjamin Rogers for pointing out this error.)

8. p. 222: The "more conventional" confidence interval for β_1 is in error; it should read:

$$B_1 - t_{a/2,\,n-3}\frac{S_E}{\sqrt{\sum x_{i1}^{*2}(1-r_{12}^2)}} \leq \beta_1 \leq B_1 + t_{a/2,\,n-3}\frac{S_E}{\sqrt{\sum x_{i1}^{*2}(1-r_{12}^2)}}$$

The same errors occur in Exercise 9.12 on p. 238.

9. p. 235: "When the number of IVs in $\mathbf{Z}$ is $k+1$, $\mathbf{b}_{\text{2SLS}} = \mathbf{b}_{\text{2SLS}}$" should read, "When the number of IVs in $\mathbf{Z}$ is $k+1$, $\mathbf{b}_{\text{2SLS}} = \mathbf{b}_{\text{IV}}$".

10. p. 254: On line 3, "the vector $B_1\mathbf{x}_1^*$ is also the orthogonal projection of $\widehat{y}^*$ onto $\mathbf{x}_1^*$" should read "the vector $B\mathbf{x}_1^*$ is also the orthogonal projection of $\widehat{\mathbf{y}}^*$ onto $\mathbf{x}_1^*$". That is, B_1 should be B, and $\widehat{\mathbf{y}}^*$ should be boldface. Later in the same paragraph, the vector $\widehat{\mathbf{y}}^*$ should again be boldface. (My thanks to Che Zhihua for pointing out the substitution of B_1 for B.)

11. p. 261: The hint to Exercise 10.7 should read "Use $\widehat{\mathbf{y}}^* = B_1\mathbf{x}_1^* + B_2\mathbf{x}_2^*$." That is, $\widehat{\mathbf{y}}^*$ is a vector and so should be in boldface.

12. p. 292: The equation

$$\mathbf{u} \equiv \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1}\mathbf{L}\mathbf{b}$$

should read

$$\mathbf{u} \equiv [\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}']^{-1}\mathbf{L}\mathbf{b}$$

(Likan Zhan)

13. p. 319: Lines 1–2 of the last paragraph: "... because the model in Equation 12.5, specifying a linear relationship" should read "... because the model in Equation 12.15, specifying a linear relationship". (I'm grateful to Andrew Swift for reporting this error.)

14. p. 335: There are two errors and one possibly misleading statement in Exercise 12.3. (I'm grateful to Peter Dalgaard for bringing these problems to my attention.)

   - In part (a), the equation for the likelihood should read

   $$L(\boldsymbol{\beta}, \sigma_\varepsilon^2) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right]$$

   That is, the equation should have $\boldsymbol{\Sigma}^{-1}$ and not $\boldsymbol{\Sigma}$ in the exponent.

   - In part (b), The MLE of σ_ε^2 should be

   $$\widehat{\sigma}_\varepsilon^2 = \frac{\sum (E_i w_i)^2}{n}$$

   not

   $$\widehat{\sigma}_\varepsilon^2 = \frac{\sum (E_i / w_i)^2}{n}$$

   - Part (c) would be clearer as, "The MLE of $\boldsymbol{\beta}$ is equivalent to minimizing the weighted sum of squares $\sum w_i^2 E_i^2$."

15. p. 358: Last three lines of the first paragraph: The sentence "To make generalized variance-inflation factors comparable across dimensions, Fox and Monette suggest reporting $\text{GVIF}^{p/2}$ ...." should read "To make generalized variance-inflation factors comparable across dimensions, Fox and Monette suggest reporting $\text{GVIF}^{\frac{1}{2p}}$ ...." (Thank you to Peter Dalgaard for reporting this error.)

16. p. 390: The left curly brace in the first line of equation for $\mathcal{V}(\boldsymbol{\beta})$ is misplaced; the equation should read

$$\mathcal{V}(\boldsymbol{\beta}) = \left\{ \sum \frac{\exp(-\mathbf{x}_i'\boldsymbol{\beta})}{[1+\exp(-\mathbf{x}_i'\boldsymbol{\beta})]^2} \mathbf{x}_i \mathbf{x}_i' \right\}^{-1}$$

17. p. 415: In Exercise 14.12, the reference to Equation 14.2.1 should be to Equation 14.20 (on p. 393).

18. p. 432–433 (and correcting a previous version of this erratum): The description of the negative-binomial regression model for count data is incorrect. The following text replaces the section on the negative-binomial model. Exercises 15.1 and 15.2 on pp. 464–465 are also affected, and corrected versions of these exercises appear below. I'm grateful to Peter Dalgaard for alerting me to this problem.

The Negative-Binomial Model

There are several routes to regression models for counts that use the negative-binomial distribution (see, e.g., Cameron & Trivedi, 1998, Section 4.2; Long, 1997, Section 8.3; and McCullagh & Nelder, 1989, Section 6.2.3). The following approach (adapted from Cameron & Trivedi, 1998, Section 4.2.2) begins with the Poisson regression model and adds random errors to it based on the gamma distribution to account for overdispersion.

Recall that in the Poisson GLM with the log link, the observed count Y_i for observation i follows a Poisson distribution with parameter μ_i, which is the expected count,

$$p(y_i|\mu_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$$

where $\mu_i = \exp(\alpha + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}) = e^{\eta_i}$. Individuals who share the same values of the xs have the same linear predictor η and hence the same expected count. I explicitly show that the distribution of Y_i is conditional on μ_i because I'm about to complicate the model for μ_i.

To accommodate unmodeled heterogeneity, we can add an unobserved error ε_i to the linear predictor, obtaining

$$\mu_i = \exp(\eta_i + \varepsilon_i) = e^{\eta_i} e^{\varepsilon_i} = \mu_i^* \delta_i$$

where $\mu_i^* \equiv e^{\eta_i}$ and the multiplicative error $\delta_i \equiv e^{\varepsilon_i}$. If, for all observations, $\varepsilon_i = 0$ and thus $\delta_i = 1$, then the expected count $\mu_i = \mu_i^*$, returning us to the Poisson regression model, but if the errors differ, then there will be additional—that is, extra-Poisson—variation in the observed counts Y_i.

I have yet to specify a distribution for the δs, but let us suppose now that δ_i is gamma-distributed with shape parameter ψ and scale parameter $\omega = 1/\psi$.[18] This in effect defines a one-parameter gamma family in which the scale parameter is the inverse of the shape, and so $E(\delta_i) = \omega\psi = 1$ and $V(\delta_i) = \omega^2\psi = \omega$. Then the observed count Y_i follows a beta distribution[19]

$$p(y_i) = \frac{\Gamma(y_i + \psi)}{y_i!\Gamma(\psi)} \times \frac{\mu_i^{y_i}\psi^{\psi}}{(\mu_i + \psi)^{y_i+\psi}} \qquad (15.4)$$

with expected value $E(Y_i) = \mu_i$ and variance $V(Y_i) = \mu_i + \mu_i^2/\psi = \mu_i + \omega\mu_i^2$.

Unless the gamma scale parameter ω is small, therefore, the variance of Y increases more rapidly with the mean than the variance of a Poisson variable, and also potentially more rapidly than in the quasi-Poisson model, where the dependence of the variance of Y on μ is linear rather than quadratic. Making the expected value of Y_i a random variable incorporates additional variation among observed counts for observations that share the same values of the explanatory variables and consequently have the same linear predictor η. Because this model incorporates an individual-specific random term, it may be thought of as a simple mixed-effects model (see Chapters 23 and 24).

With the gamma scale parameter ω fixed to a known value, the negative-binomial distribution is an exponential family (in the sense of Equation 15.15 in Section 15.3.1), and a GLM based on this distribution can be fit by iterated weighted least squares (as developed in the next section). If instead—and as is typically the case—the value of ω is unknown and must therefore be estimated from the data, standard methods for GLMs based on exponential families do not apply. We can, however, obtain estimates of both the regression coefficients and ω by the method of maximum likelihood.

Applied to Ornstein's interlocking-directorate regression and using the log link, the negative-binomial GLM produces results very similar to those of the quasi-Poisson model (as the reader may wish to verify). The estimated scale parameter for the negative-binomial model is $\widehat{\omega} = 1.312$, with standard error $\mathrm{SE}(\widehat{\omega}) = 0.143$; we have, therefore, strong evidence that the conditional variance of the number of interlocks increases more rapidly than its expected value.[20]

Exercise 15.1. Testing overdispersion: Recall that ω is the scale parameter for the gamma component of the negative-binomial regression model (see pages 432–433). When $\omega = 0$, the negative-binomial model reduces to the Poisson regression model (why?), and consequently a test of H_0: $\omega = 0$ against the one-sided alternative hypothesis H_a: $\omega > 0$ is a test of overdispersion. A Wald test of this hypothesis is obtained by dividing $\widehat{\omega}$ by its standard error. We can also compute a likelihood-ratio test contrasting the deviance under the more specific Poisson regression model with that under the more general negative-binomial model. Because the negative-binomial model has one additional parameter, we refer the likelihood-ratio test statistic to a chi-square distribution with 1 degree of freedom; as Cameron and Trivedi (1998, p. 78) explain, however, the usual right-tailed p-value obtained from the chi-square distribution must be halved. Apply this likelihood-ratio test for overdispersion to Ornstein's interlocking-directorate regression.

Exercise 15.2. The error in Equation 15.4, corrected above, is repeated in part (c) of this exercise.

19. p. 442: The cross-reference for Table 15.2 is given as page 441 but this table is actually on page 421. (I'm grateful to Rosa C. Banuelos for pointing out this error.)

20. p. 442: The reference to Equation 15.3 should be to Equation 15.13. (Likan Zhan)

21. p. 444: The formula for $c(y, \phi)$ for the inverse-Gaussian family is missing a factor of 2. It should read

$$c(y, \theta) = -\frac{1}{2}[\log_e(2\pi\phi y^2) + 1/(\phi y)]$$

not

$$c(y, \theta) = -\frac{1}{2}[\log_e(\pi\phi y^2) + 1/(\phi y)]$$

22. p. 454: The estimated dispersion parameter in the displayed expression for the Pearson residuals near the bottom of the page should be denoted $\widetilde{\phi}$, not $\widehat{\phi}$, and so the expression should read

$$\frac{\widetilde{\phi}^{1/2}(Y_i - \widehat{\mu}_i)}{\sqrt{\widehat{V}(Y_i|\eta_i)}}$$

(Likan Zhan) As well, the cross-reference should be given as "Equation 15.19 on page 448," not as "Equation 15.9 on 436."

23. p. 481: Towards the bottom of the page,

$$\begin{aligned}\sigma_s \equiv C(\varepsilon_t, \varepsilon_{t-s}) &= \phi_1 E(\varepsilon_{t-1,t-s}) + \phi_2 E(\varepsilon_{t-2}\varepsilon_{t-s})\\ &= \phi_1\sigma_{s-1} + \phi_2\sigma_{s-2}\end{aligned}$$

should read

$$\begin{aligned}\sigma_s \equiv C(\varepsilon_t, \varepsilon_{t-s}) &= \phi_1 E(\varepsilon_{t-1}\varepsilon_{t-s}) + \phi_2 E(\varepsilon_{t-2}\varepsilon_{t-s})\\ &= \phi_1\sigma_{s-1} + \phi_2\sigma_{s-2}\end{aligned}$$

(Likan Zhan)

24. p. 486 & 487: The references to Equation 16.3 near the bottom of p. 486 and near the top of p. 487 should be to Equation 16.13. (Likan Zhan)

25. p. 498: There's an error in the hints for Exercise 16.5: Actually, not $\mathbf{\Sigma}_{\varepsilon\varepsilon}$ but $\mathbf{\Sigma}_{\varepsilon\varepsilon}^{-1} = (1/\sigma_\nu^2)\mathbf{\Gamma}'\mathbf{\Gamma}$, and so $\det \mathbf{\Sigma}_{\varepsilon\varepsilon} = (\sigma_\nu^2)^n(1/\det \mathbf{\Gamma})^2$.

26. p. 567: Within step 1. near the top of the page,

$$\begin{aligned}\widehat{f}_1^{(0)}(x_{i1}) &= B_1(x_{i2} - \overline{x}_2)\\ \widehat{f}_2^{(0)}(x_{i2}) &= B_2(x_{i2} - \overline{x}_2)\end{aligned}$$

should read

$$\begin{aligned}\widehat{f}_1^{(0)}(x_{i1}) &= B_1(x_{i1} - \overline{x}_1)\\ \widehat{f}_2^{(0)}(x_{i2}) &= B_2(x_{i2} - \overline{x}_2)\end{aligned}$$

(Likan Zhan)

27. p. 664. In Exercise 21.1, the bootstrap standard error of the mean is given as $\mathrm{SE}^*(\bar{Y}^*) = \frac{S}{\sqrt{n-1}}$. This result is slightly in error; the correct expression is $\mathrm{SE}^*(\bar{Y}^*) = \frac{S\sqrt{n-1}}{n}$.

28. p. 569: About a third of the way down the page, the estimated error variance should be given as $S_E^2 = \mathrm{RSS}/df_{\mathrm{res}}$, *not* $S_E^2 = \mathrm{RSS}/(n - df_{\mathrm{res}})$. (Likan Zhan)

29. p. 634: The log-likelihood for the Heckman model given in Equation 20.20 has two errors: In the first sum, $\Phi(\mathbf{z}_i'\boldsymbol{\gamma})$ should be $\Phi(-\mathbf{z}_i'\boldsymbol{\gamma})$ or equivalently $1 - \Phi(\mathbf{z}_i'\boldsymbol{\gamma})$. In the second sum, $\sqrt{\dfrac{1-\rho_{\varepsilon\delta}}{\sigma_\varepsilon}}$ should be $\sqrt{1-\rho_{\varepsilon\delta}^2}$. Thus, the equation should read:

$$\begin{aligned}\log_e L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_\varepsilon^2, \rho_{\varepsilon\delta}) &= \sum_{i=1}^{m} \log_e \Phi\left(-\mathbf{z}_i'\boldsymbol{\gamma}\right)\\ &+ \sum_{i=m+1}^{n} \log_e \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{Y_i - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} + \rho_{\varepsilon\delta}\dfrac{Y_i - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma_\varepsilon}}{\sqrt{1-\rho_{\varepsilon\delta}^2}}\right)\right]\end{aligned}$$

30. p. 713: In the table towards the bottom of the page, the ML and REML estimates of ψ_{12} are both more accurately given as 0.047 rather than as, respectively, 0.041 and 0.042. It's possible that these differences are due to a change in the software used to estimate the model. (Likan Zhan)

31. p. 716: The reference to Figure 23.4 in the last bullet item should be to Figure 23.5. (Likan Zhan)

32. p. 723: In the table towards the bottom of the page, the AIC for Model 2 should be 3605.0, *not* 360.0. (Likan Zhan)

33. p. 735: In the first bullet item, there's a missing left square bracket in the definition of $\mathbf{y}$, which should read, $\underset{(n\times 1)}{\mathbf{y}} \equiv [\mathbf{y}_1', \mathbf{y}_2', \ldots, \mathbf{y}_m']'$. (Likan Zhan)

34. p. 736: The last diagonal entry in the matrix at the top of the page should be $\mathbf{\Lambda}_m$, *not* $\mathbf{\Lambda}_2$, and consequently the equation should read

$$\sigma_\varepsilon^2 \underset{(n\times n)}{\mathbf{\Lambda}} \equiv \sigma_\varepsilon^2 \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Lambda}_m \end{bmatrix}$$

(Likan Zhan)

35. On-line Appendices, p. 57: The point at which the partial derivatives are 0 is $x_1 = x_2 = 0$, not at $x_1 = x_2 = 0.5$, as stated in the text. The remainder of the example is correct as given. (I'm grateful to Naresh Gurbuxani for pointing out this error.)