ERRATA : Negative Binomial Regression
Joseph M. Hilbe Cambridge University Press

17August, 2008

AUTHOR: Hilbe@asu.edu; jhilbe@aol.com
PUBLISHER: http://www.cambridge.org/uk/catalogue.asp?isbn=9780521857727

PREFACE, Page x: 2nd line of 3rd complete paragraph. Web address (same as on copyright page): "www.cambridge.org/XXXXX" should read www.cambridge.org/9780521857727".

INTRODUCTION, Page 2: last sentence of first paragraph: The SPSS generalized linear models procedure name is mistaken. It is GENLIN, not GLZ (Statistica).

Ch 2, Page 20, 1st paragraph, line 8: Add "family", so that the line should read: "link, variance, and family functions. The algorithm took care …"

CH 2, Page 21: line 2: First word of line reads "or". It should read "and".

CH 2, Page 22: Equation 2.6 : Should read: $X_1 = X_0 - f(X_0)/f'(X_0)$

+ : Equation 2.8 : 2nd equation has a missing "=" sign. Should read $\partial^2 L = \partial^2 L/\partial\beta\partial\beta'$

+ : Equation 2.12; replace $\partial\mathcal{L}/\partial\beta_j$ with simply $\mathcal{L}$.

CH 2, Page 23: Equation 2.14; replace denominator on left hand side term with $\partial\beta_i$

CH 2, Page 27, Table 2.1, AIC statistic note: The note to the program line defining the AIC statistic reads: "/* AIC is sometimes defined w/o η */". The last term, η, should read, n, not η. See comments below.

CH 2, Page 27, 8th line under Table 2.1 Line to amend reads: "…function; BIC represents the Baysean Information Criterion, which is usually based on the deviance value." "Baysean" should be spelled "Bayesian", and the sentence following semicolon should read: "… function; BIC represents the Bayesian Information Criterion, which was originally based on the deviance, but is now usually formulated in terms of the log-likelihood ."

CH 2, Page28, Section 2.2. Replace paragraph with the following: "In this section we discuss the derivation of the Newton-Raphson type algorithm. Until recently, the only method used to estimate the standard negative binomial model was by maximum likelihood estimation using a Newton-Raphson based algorithm. All other varieties of negative binomial are still estimated using a Newton-Raphson based routine. We shall observe in this section though, that the Iteratively Re-weighted Least Squares method we discussed in the last section, known as Fisher Scoring, is a subset of the Newton-Raphson method. We conclude by showing how the parameterization of the GLM mean, μ, can be converted to $X'\beta$."

CH 3, page 32, Table 2.4. Amend the line near bottom of table beginning with BIC, to read:

```
 BIC = DEV – (dof)ln(n) /// or  -2LL + ln(n)*k (Stata version);
k=# predictors
```

CH 3, page 40, 1st line under equation 2.53. "caste" should be spelled "cast".

CH 3, page 47, equation 3.23: There should be an equal sign between Z_i and the other terms

CH 3, page 47, code at bottom of page. First line of code, “predict xb” should read, “predict mu” - and the second line deleted “gen mu=exp(xb)” OR: it can read “predict xb, xb”, and leave the second line as is.

CH 2, page 48, Equation 3.24. The y in the numerator should have a bar over it, Indicating that it is the mean of y.

CH 2, page 48, code in middle of page. List line with leading dot. “n” should read “_n”. The line should now appear as: display “LM value =” chival _n “p-value =”

CH 4, Page 52, second to last line of program code on page. Line should read: . gen xb = 1 + .5*x1 - .75*x2 + .25*x3

CH 4, Page 63: last line of code in Table 4.1: It should read: "w = 1/sqrt(sc)" instead of "w = μ/sqrt"

CH 4, Page 73, 6th line of Stata commands: should read: . di 20.02131 + .3*20.02312^2 /* mean + alpha*mean^2 */

CH 5, Page77, 5th line from top: delete word `model’.

CH 5, Page 78. Table 5.1: The last two items should be numbered 24 and 25, and not 22 and 22.

CH 5, Page 82: 3rd line from top. Delete “would be” and substitute, “is”.
4th line from top. Delete “; only to limit *r* as positive.” Period at end.

CH 5, Page 82: Equations 5.19 and 5.20: The “choose” functions should not have a division sign between top and bottom terms. See Equation 5.29 as an example. $\binom{y + r - 1}{r - 1}$

CH 5, Page 94: value of variance under the word “hence”: In the list of formulae there are two equations with a left hand side of $V(\mu)$. The 2nd of these two equations should read $V'(\mu)$ instead of $V(\mu)$. Therefore $V'(\mu) = 1 + 2\alpha\mu$

CH 5, Page 95: table 5.5, 1st line in loop: Should read: $w = (\mu/\alpha\mu)+(y-\mu)\{\alpha\mu/(1+2\alpha\mu+\alpha^2+\mu^2)\}$

CH 5, Page 96: Section 5.5, line 6: Should be “two-parameter”, not “one-parameter”. The line should read: “Poisson variance and μ^2/ν the two-parameter gamma distribution variance. We …”

CH 5, page 96: Section 5.5, Paragraph 3, line 2. Delete comma between “version” and “of”.

CH 6, Page 112: Formula for the negative binomial variance is mistaken. Should read: $\mu + \alpha\mu^2$
The formula as it currently reads is missing α in the 2nd term.

CH 6, Page 119: Example 3, first line: “log” rather than “data”
Should read: “These data come from the 1912 Titanic survival log. It consists…”

CH 6, Page 126: 3rd sentence under Table 6.14. Insert word "the" as follows:

"…household income, Age and education are not contributory to the model. Recall…"

CH 6, Page 128: line under first table display on top of page: Delete the "?" which is at the end of the line.

CH 7, Page 136: Paragraph 2, line 4, first word: Word "variance" should read "assumptions". Line 4 of 2nd paragraph should start out as "assumptions of the Poisson distribution. Other models are …"

CH 7, Page 139, line 3 under table of parameter estimates, 1st word of line: Word "restricted" should be changed to "expected". The line should read, in part: "expected for a geometric model. Recall that unless …"

CH 7, Page 143, 2nd last line of last full paragraph. Replace term "Figure 7.1" with "Figure 7.2".

CH 7, Page 153, last two lines on page. Should read: -2{(-60322.021) – (-60258.97)}
126.102

CH 7, Page 154, second line top paragraph: Should read: Chi2tail(1,126.102) = 2.921e-29.

CH 7, Page 155, 7th-6th line from bottom: Yang, Hardin, and Addy (2006), not 2007.

CH 7, Page 163: Wald statistic in bottom output (header):
The Wald chi2(2) is now reported as a Likelihood Ratio test (LR test) in Stata's *ztnb* command. The output was created in a program by the same name I wrote before Stata offered the command. Published on the SSC site, Stata Corp used it as the basis of the new command, with LR rather than Wald, which came automatically with all ML estimations.

CH 8, Page 171, bottom-page model output. Negative binomial age3 coefficient should read: .023721. The decimal point was inadvertently dropped.

CH 8, Page 172, AIC Formula should read: $AIC_{hurdle} = ((AIC_{ZT} * (1-N_{>0}/N) + AIC_{binary}$

CH 8, Page 174, Final full paragraph, final 4 lines: The three terms predicting zero counts are parameterized in terms of μ rather than $x\beta$, as shown in the two formulae mid-page. To be consistent, the discussion should be in terms of $x\beta$ in both places. The last 4 lines should read:
"zero count are; (1) logistic inverse link, i.e. $1/(1+\exp(-x\beta))$, the prediction
that y==1, (2) 1- $[1/(1+\exp(-x\beta))$, and (3) the negative …"

CH 9, Page 180, Section 9.1.1, 2nd paragraph, line 3: "lower" should read "higher", comma to semicolon: Therefore, the line should read: "wish to extend C to any higher value in the observed distribution; the value to"

CH 9, Page 181, Header for bottom table: Should read: "POISSON: DROPPED VALUES 0-3"

CH 9, Page 193, line 2 under top table of parameter estimates: "selection" instead of "selected". The sentence should therefore read: "to the selection corrected Poisson".

CH 10, Pages 213-225, panel models using the *progabide* data. It is preferable to use i(id) and t(t) options. Both make sense, but using the above two options are more sensical and result in a better interpreted model.

CH 10, Page 221: Stata command at bottom of page. Delete `t(id)' since time panel not applicable for structure.

CH10, Page 223, reword from 4th line of text:"… decreases by a 1-unit increase in power of the previous diagonal. One might … The depiction here is true for this example, but not for all AR levels or structures."

CH 10, Page 226, Section 10.5, line 1: "also" should read "sometimes". Line should read: "Multilevel models are sometimes called hierarchical models, particularly …". For clarification purposes, I add an additional sentence after the first. The first three sentences of the paragraph should then read: "Multilevel models are sometimes called hierarchical models, particularly in educational and social science research. However, the majority of statisticians now tend to draw a distinction between multilevel and hierarchical models, primarily because of the manner in which the methods define order of levels, or nesting. Regardless, the idea behind multilevel modeling is to model the dependence that exists …"
NOTE: Chapter 10 has been rewritten and can be downloaded from:
http://www.statistics.com/hilbe/nbr.php

Pages 85-90 Graphs
The lack of color in the book makes it difficult to determine which line is associated with a specific mean and/or alpha. For Figures 5.1 – 5.6, which show different means for the same value of alpha (α), the values of the mean from top to bottom at count 0, are: [0.5, 1, 2, 5, 10]. For Figures 5.7 - 5.11, the values of alpha for a specific mean, from top to bottom at count 0, are [3, 1.5, 1, .67, .33, 0]. The BOOK COVER is the same as Figure 5.10.

CH 10: GEE models using *progabide* data
The version of the Stata software used for developing GEE models using the *progabide* data did not recognize that the stationary, nonstationary, and unstructured correlation structures are infeasible for this data; i.e. the algorithm did not recognize that the correlation structure was not positive definite. The correlation matrix for these specific models are not reliable. The software has since been amended to display an error message when the matrix is infeasible, and will not produce a table of estimates or a post-estimation correlation matrix. Stata versions 9.2 and higher work as they should; I am not certain at which lower version the *xtgee* command was amended. I suggest that the reader ignore the inherent unreliability of these specific model results, reading the text and its interpretation as if appropriate convergence was achieved – as it appears in the output. The pedagogical value of the discussion is nevertheless valid. The correlation values produced, together with parameter estimates and standard errors, appear to be reasonable, and can be used with value in demonstrating how the models are to be estimated and evaluated. The "stationary 4" correlation structure is feasible for this data, unlike other stationary correlation structures. It can be developed using the command: `. xtgee seizures time timeXprog, fam(poi) i(id) t(t) corr(stat 4) force`
It is important to use the *force* option when time intervals are not all equal. If in fact the time intervals are equal, do not use the *force* option.