

Remarks and corrections for first printing of: Generalized Latent Variable Modeling

Skrondal, A. and Rabe-Hesketh, S. (2004). Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models. Chapman & Hall/CRC.

[Datasets and do-files](#)

Major corrections

- **P. 35, line 1-6**

This is incorrect for the complementary log-log link because the Gumbel distribution is asymmetric. The text should read:

It follows that

$$\begin{aligned} E(y_i|\nu_i) &= \Pr(y_i = 1|\nu_i) = \Pr(y_i^* > 0|\nu_i) = \Pr(\nu_i + \epsilon_i > 0) \\ &= \Pr(-\epsilon_i \leq \nu_i) = F(\nu_i). \end{aligned}$$

Here $F=g^{-1}$ is the cumulative distribution function; the standard normal for the probit, the logistic for the logit, and the (asymmetric) Gumbel for the complementary log-log link.

- **P. 41, Equation (2.24)**

The second brace under the equation should only cover the sum because the argument of the exponential is $-H_i(t_i)$.

- **P. 43, Display 2.1**

It should be $\tau_0=0$, not $\tau_0=1$.

- **P. 45, between equations (2.27) and (2.28)**

r subscripts are missing in “ $y_{ir} = 0$ when $r < s$ and $y_{ir} = 1$ when $r = s$ ”

It should be “the s th interval” instead of “the k th interval”.

- **P. 70**

There are some errors in the text. The corrected page can be downloaded [here](#).

We thank Dirk Knol for making us aware of some of the corrections on this page.

- **P. 74**

There are some errors in the text. The corrected page can be downloaded [here](#).

We thank Jeroen Vermunt for pointing out mistakes on this page.

- **P. 176, third equation**

The expression for c_{jr}^{MC} should be:

$$c_{jr}^{MC} = \frac{\prod_i g(y_{ij}|\mathbf{Q}^k \mathbf{d}_{jr}; \boldsymbol{\vartheta}^k)}{\sum_r \prod_i g(y_{ij}|\mathbf{Q}^k \mathbf{d}_{jr}; \boldsymbol{\vartheta}^k)},$$

- **P. 176, fifth equation**

The expression for $c_{r_m}^{GH}$ should be:

$$c_{r_m}^{GH} = \frac{p_{r_m} \prod_i g(y_{ij}|\mathbf{Q}^k \mathbf{a}_r; \boldsymbol{\vartheta}^k)}{\sum_r p_{r_m} \prod_i g(y_{ij}|\mathbf{Q}^k \mathbf{a}_r; \boldsymbol{\vartheta}^k)},$$

We thank Pengsheng Ni for making us aware of the two corrections on this page.

- **P. 210**

There are some errors in the text. The corrected page can be downloaded [here](#)

- **P. 211**

There are some errors in the text. The corrected page can be downloaded [here](#)

- **P. 307, bulleted list of variables**

- [Time] should be defined as calendar time (year-1975)
- [Age] should be defines as wife’s age (years-35)
- [Agesq] should be defined as [Age] squared

We thank Graham Dunn for making us aware of the errors.

- **P. 352**

There are some errors for the BIC’s in Table 11.2 and the surrounding text. The corrected page can be downloaded [here](#). We thank Dirk Knol for pointing this out.

- **P. 354**

There are some errors for the BIC’s in Table 11.4 and the surrounding text. The corrected page can be downloaded [here](#). We thank Dirk Knol for pointing this out.

- **P. 409, Table 13.7**

For model \mathcal{M}_0 : Estimate -0.37 (0.34) should be for class 3 whereas estimate -0.73 (0.23) should be for class 2. We thank Dirk Knol for pointing this out.

- **P. 413, Table 13.10**

There are errors in the estimates. For the latent class models, estimates are those obtained if effect coding is used instead of dummy coding. See [here](#) for the correct table. Thanks to Jay Magidson for pointing out the errors.

- **P. 431, line 10 (displayed equation)**

The probability should be $\Pr(\eta_j = 1 | r_j = 0)$, i.e. η_j instead of c_j .

Minor corrections

- **P. 23, last displayed equation**

Should be

$$\Pr(y_i | \mu_i) = \binom{n}{y_i} \mu_i^{y_i} (1 - \mu_i)^{(1-y_i)n},$$

(thanks to Dirk Knol for pointing this out)

- **P. 40, lines 6-7**

Should be “Let T_i be a random variable representing the duration for a unit i from becoming at risk until it experiences the event.”

- **P. 66, line 26**

Should be “... Skrondal and Laake, 1999)”

- **P. 81, lines -5, -6**

α_j should not be bold

- **P. 112, line 4**

Should be “ ϱ^c are regression parameters with $\varrho^1 = \mathbf{0}$ ”

- **P. 124, log link**

Replace “apart from the intercept” by “for all covariates that do not appear in the random part of the model”

- **P. 124, line -5**

Replace “Section 4.28” by “equation (4.28)”

- **P. 125, line -13**

Should be

$$\frac{(16/15)^2}{\pi^2/3} \approx 0.35$$

(thanks to Dirk Knol for pointing this out)

- **P. 165, equation (6.11)**

Should be a_{r_1} and $a_{r_{M_l}}$ in last line:

$$\approx \sum_{r_1=1}^{R_1^{(l)}} p_{r_1} \cdots \sum_{r_{M_l}=1}^{R_{M_l}^{(l)}} p_{r_{M_l}} \prod g^{(l-1)}(\mathbf{y}_{(l-1)} | a_{r_1}, \dots, a_{r_{M_l}}, \mathbf{v}^{([l+1]^+)})$$

- **P. 240, line -1**

Should be $\text{Var}_{\mathbf{y}}$ in second term:

$$\text{Var}_{\mathbf{y}} \left(\tilde{\zeta}_{(2)}^{\text{ML}} | \zeta_{(2)}, \mathbf{X}_{(2)}; \hat{\boldsymbol{\theta}} \right) = \text{Var}_{\mathbf{y}} \left(\tilde{\zeta}_{(2)}^{\text{ML}} - \zeta_{(2)} | \mathbf{X}_{(2)}; \hat{\boldsymbol{\theta}} \right) = \hat{\boldsymbol{\theta}}/n_j$$

- **P. 286, Table 9.1**

SE for $\hat{\psi}$ should be 0.217 (thanks to Tor Haldorsen for pointing this out)

- **P. 290, Table 9.3**

Est for [History], 1-Spec. should be 0.20 (thanks to Scott Long for pointing this out)

- **P. 293, line 2**

Replace $\hat{\beta}_i$ by $-\hat{\beta}_i$ (thanks to Dirk Knol for pointing this out)

- **P. 308, last two displayed equations**

Should be

$$\Pr(y_{tj}=1 | \mathbf{x}_{tj}, y_{t-1,j}=0, \zeta_j) = \frac{\exp(\mathbf{x}'_{tj}\boldsymbol{\beta} + \zeta_j)}{1 + \exp(\mathbf{x}'_{tj}\boldsymbol{\beta} + \zeta_j)},$$

and

$$\Pr(y_{tj}=1 | \mathbf{x}_{tj}, y_{t-1,j}=1, \zeta_j) = \frac{\exp(\mathbf{x}'_{tj}\boldsymbol{\beta} + \gamma + \lambda\zeta_j)}{1 + \exp(\mathbf{x}'_{tj}\boldsymbol{\beta} + \gamma + \lambda\zeta_j)}.$$

(thanks to Dirk Knol for pointing this out)

- **P. 316, line -9 to -8**

Should be “ k for polling district (level 4)” (thanks to William Saylor for pointing this out)

- **P. 329, line -19**

Replace “four thresholds” by “three thresholds” (thanks to Dirk Knol for pointing this out)

- **P. 329, line -18**

Replace “ $\kappa_{1i}, \dots, \kappa_{4i}$ ” by “ $\kappa_{i1}, \kappa_{i2}, \kappa_{i3}$ ” (thanks to Dirk Knol for pointing this out)

- **P. 344-345, x-axis labels of Figures 10.8 and 10.9**

Should be “Life satisfaction $-\eta_j$ ”

- **P. 353, displayed equation**

Should be

$$g(y_i; \mu_i) = \frac{y_i!}{y_i!(8 - y_i)!} \mu_i^{y_i} (1 - \mu_i)^{8 - y_i}.$$

(thanks to Dirk Knol for pointing this out)

- **P. 358, line -18**

Should be: “probabilities 0.02, 0.15, 0.45, 0.22, 0.09 and 0.09, respectively...” (thanks to Dirk Knol for pointing this out)

- **P. 359, Caption of Figure 11.2**

Should be “...for components $c=1, \dots, 6$ when...” (thanks to Dirk Knol for pointing this out)

- **P. 370, line -14**

Replace “Hall (1997)” by Hall (2000)”

- **P. 407, Table 13.6**

Deviance for model with 3 classes should be 23.68 instead of 23.58 (thanks to Dirk Knol for pointing this out)

- **P. 382, line 7**

Should be “21 subjects ...”

- **P. 433, line -6**

After “... count variable.”, insert “This variable is the product of self-reported drinking frequency (the number of days in the past two weeks with any drinking) and drinking intensity (the average number of drinks on a day with any drinking), rounded to the nearest integer.”

- **P. 461, lines 4-5**

Replace “Hall, D. B. 1997” by “Hall, D. B. 2000”

- **P. 476, line 16-18**

Reference should be:

Rabe-Hesketh, S. and Skrondal, A. (2005). *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press.

Remarks

- **P. 43, line -9 to -8**

Note that omitting the offset $\ln(d_{ir})$ leads to an equivalent model if a dummy variable is used for each risk set. We thank Harvey Goldstein for pointing this out.

- **P. 99, line 3**

Another useful reference here is:

Rijmen, F., Tuerlinckx, F., De Boeck, P. and Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods* **8**, 185-205.

- **P. 113, line -17**

Important omitted references on robustness are:

Neuhaus, J. M., Hauck, W. W. and Kalbfleisch, J. D. (1992). The effects of mixture distribution misspecification when fitting mixed-effects logistic models. *Biometrika* **79**, 755-762.

Heagerty, P. J. and Kurland, B. F. (2001). Misspecified maximum likelihood estimates and generalized linear mixed models. *Biometrika* **88**, 973-985.

- **P. 199, line -5**

Add reference:

Robins, J. M., Rotnitzky, A. G. and Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association* **90**, 106-121.

- **P. 216**

There is of course plenty of software for linear mixed models, including xtmixed in Stata, PROC MIXED in SAS, lme in S-PLUS and R, Linear Mixed Models in SPSS, MLwiN

- **P. 216**

From Stata 10 there are commands xtmeologit and xtmeopoisson implementing adaptive quadrature for random effects logistic and Poisson regression. and HLM.

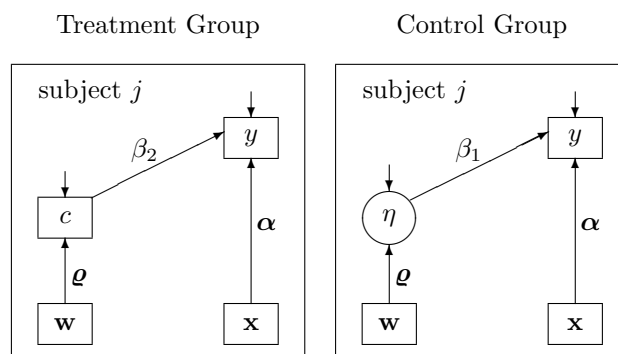
- **Section 10.2**

A useful reference for latent growth curve models with continuous responses is

McArdle, J. J. (1988). Dynamic but structural equation modeling with repeated measures data. In J. R. Nesselrode and R. B. Cattell (Eds.). *Handbook of Multivariate Experimental Psychology*, Volume II. New York: Plenum, pp. 561-614.

- **P. 432**

Readers might find a path diagram for the CACE model useful:



- **Notation**

Covariates in structural model for discrete latent variable sometimes referred to as \mathbf{v}_j (e.g. Section 4.3) and sometimes as \mathbf{w}_j (e.g. Section 14.4) – should be \mathbf{w}_j everywhere.

- **References**

For some of the software manuals the publisher is missing

We would appreciate to be informed of further remarks and corrections; please email Sophia Rabe-Hesketh: sophiarh@berkeley.edu.