Generalized Linear Probability Models in HLM

R. B. Taylor
Department of Criminal Justice
Temple University

fi=hlml15

The Problem

Up to now we have been addressing multilevel problems where we have assumed the outcome is relatively normally distributed, and has a roughly linear relationship with the predictors. This situation does not fit two types of outcomes: dichotomous outcomes, and count outcomes.

Dichotomous outcomes follow a binomial distribution.

Count outcomes can follow any of a number of different distributions (e.g., negative binomial, Poisson). HLM will allow you to model a Poisson-distributed outcome.

For each of these distributions, HLM will allow you to specify over-dispersion (BRC: 133). More on this later.

Different Varieties

The first question is to decide what distributional form your outcome takes.

Bernoulli. If your outcome is 0/1, then you will be following a Bernoulli distribution. The mean of the distribution, the proportion of cases scoring 1, will be between .01 and .99. FIGURE 1 shows a Bernoulli distribution of 100 randomly generated cases; the mean was intended to be set at .3 (30% of the cases score 1); it came out to 38%.Some possible applications:

* recruit fails out of the police academy (0), or succeeds (1)
* felon is convicted (1) or not (0)
* police officer is released due to corruption problems (1) or leaves through normal channels (0).
* prisoner is granted parole (1) or is not (0).

In regression, we would deal with such outcomes using either a PROBIT model, or a dichotomous LOGIT model. The latter is available in SPSS and is relatively easy to use; you just set it up like a regular regression.

Binomial. If your outcome is in the form of y many successes in Nij trials, then the outcome is a binomial distribution (Blalock 1979: 149). When you are thinking about an outcome that is success/failure, over a certain number of trials, and the number of trials is relatively small, and the different trials are independent, then you are following a binomial distribution. For example:

* the number of random drug tests where R tested positive
* the number of times a prosecutor plea bargained a case (y), out of all the cases processed that day.
* the number of times an officer was cited in a year (y), as a result of routine monthly inspections (Nij=12)

The binomial distribution is like the Bernoulli, it's just that the outcome is not limited to 0 or 1, because it is now Y successes out of Nij trials. The Bernoulli distribution is just a special subset of Binomial distributions.

Poisson. Poisson distributions are used for distributions of COUNT data when the outcome is an extremely rare event. The first use was to describe deaths of cavalry officers from horse kicks in the Prussian Army in the Crimean war. The distribution's variance is driven by its mean. Figure 3 shows an example of a distribution with a mean of 1. Generally, if you are modeling a rare event, and you are looking at count data, you want to use a Poisson distribution.

The Solution

In these situations, regular HLM is inappropriate. In the case of a binary outcome, for example (BRC: 117): "the random effect can only take on one of two values, and therefore cannot be normally distributed ... the level 1 random effect cannot have homogenous variance." Further (118) the predicted values need to be between 0 and 1, and cannot go above or below these.

HLM carries out a link function (BRC: 119) which restricts the predicted values to fall between (in the case of a dichotomous outcome) 0 and 1.

From this you can recover the log of the odds of success (BRC: 121, Eq. 5.7) and the predicted probability that a case will score 1 (BRC: 121 - Eq. 5.8).

In the case of count data there are two additional things to think about. First, whether or not the exposure was constant across individuals or groups. For example if you are looking at the number of re-arrests for persons released from drug courts, was everyone followed for the same number of months. Months would be the exposure variable. Here also, a different link function is used (BRC: 122 - Eq. 5.11). You are getting an event rate.

Sometimes your data may not perfectly follow a binomial or a Poisson distribution; variance may be larger than it should be according to the theoretically defined distribution. This can happen due to "undetected clustering within level-1 units or if the level-1 model is under-specified" (BRC: 131). For example, you might have extremely homogeneous subgroups within your neighborhoods (homogeneous on the outcome). If this is so, you specify extra-dispersion.

Extra Data Requirements

It is important that your raw data be available, because the program returns to it after the micro-iterations; it is important that the raw data and the ssm be in the same subdirectory and that subdirectory be the one where they were originally created.

Interpreting Output

In the case of a Bernoulli outcome, you need to exponentiate your coefficients, to get at the impacts of how many times more likely someone was to score “yes” on the outcome based on a one unit increase on X. You only need to exponentiate your level 1 coefficients. You can interpret your level 2 coefficients in the usual way.

Think for a minute about why this would be so. Remember that your level 2 predictors are working on the Level 2 portion of the outcome - your neighborhood means. These means will be PROPORTIONS, not just 0's and 1's.

References

Blalock, H. M. Jr (1979) Social statistics (2^{nd}. Edition revised). New York: McGraw Hill.

Bryk, A., Raudenbush, S., and Congdon, J. (1996). HLM: Hierarchical Linear and Nonlinear Modeling with the HLM/2L and the HLM/3L Program. Chicago: SSI.

FIGURE 1

Bernoulli Distribution (p=.3)

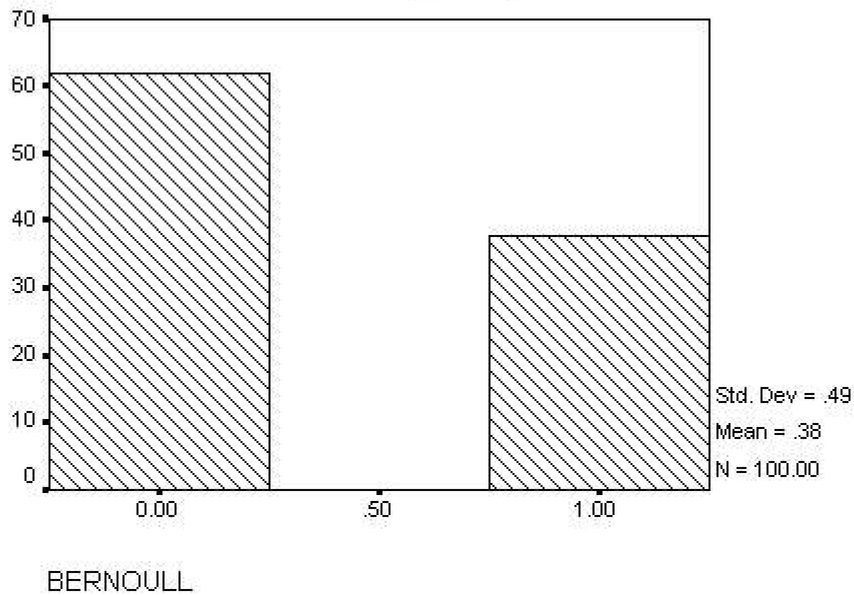


FIGURE 2

Y-av=1; NIJ=10

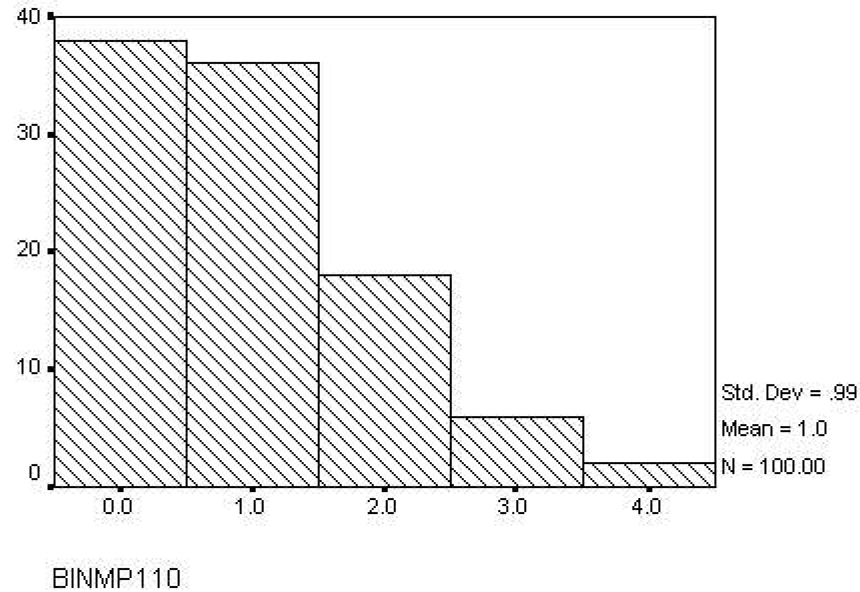
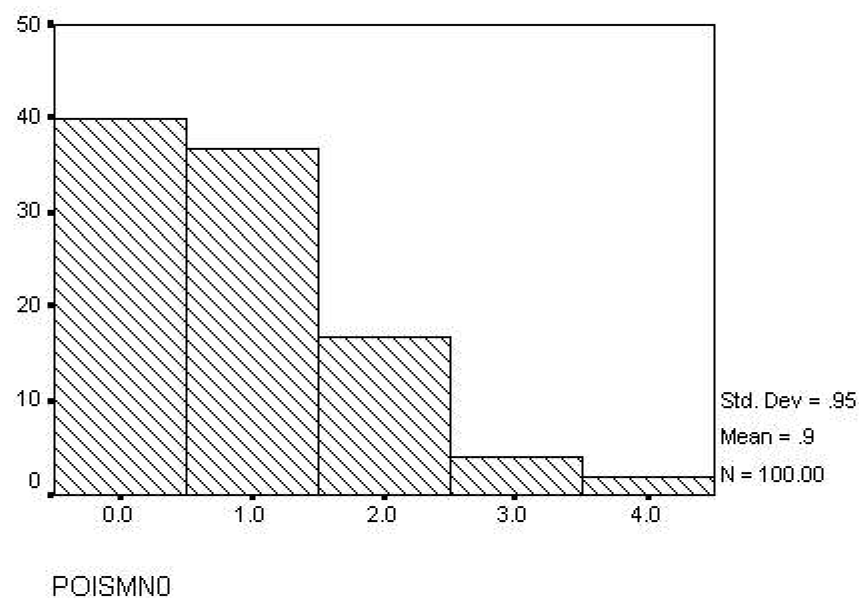


Figure 3

Poisson Mean=1



THE EXAMPLE
Outcome = do know of local drug fighting organization (KNODFORG)
1 = yes
0 = no or don't know or missing