Multilevel Models

6. Multilevel Logit Models (continued)

Germán Rodríguez

Princeton University

April 11, 2018

Contraceptive Use in Bangladesh

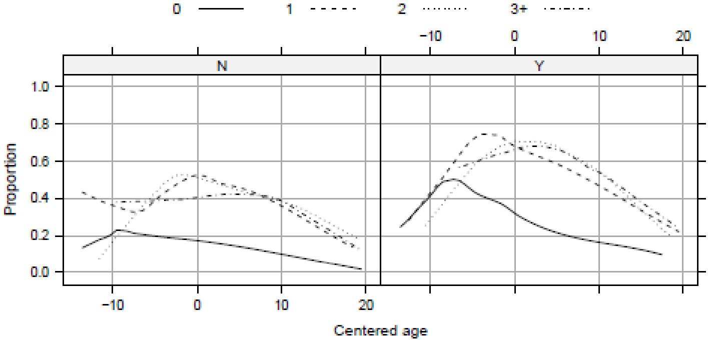
Our second dataset concerns contraceptive use in Bangladesh from Huq and Cleland (1990) and makes an appearance in the Stata manual, Bates's `lme4` book, and other papers.

The data pertain to 1934 women grouped in 60 districts. The outcome is a binary indicator of current contraceptive use. The predictors of interest include age and number of children, as well as an indicator of urban residence. We also have a district identifier.

Most districts have both urban and rural parts. We will entertain random-intercept models where the level of contraceptive use varies by district, and random-slope models where the urban-rural differential varies by district, in both cases net of observed covariates.

The first issue, however, is how to specify the fixed part of the model.

Plotting binary data is harder than continuous data, but still necessary. A useful tool is to use scatterplot smoothers such as splines or loess. The figure below shows contraceptive use by age for rural and urban women grouped by number of children.



Contraceptive use is clearly a non-linear function of age, but many analyses use just a linear term. Kudos to Bates!

Model Selection

Perhaps the first model to fit is a random-intercept model with a linear term on age, indicators of 1, 2 and 3 or more children, and an indicator for urban residence, the model used in the Stata manual and in several previous analyses.

Following Bates we'll introduce a quadratic term on age. This addition improves the fit by a remarkable 44.12 points in the chi-squared scale, which is not surprising in light of the graph.

The figure also suggests that there are very small differences between 1, 2 and 3+ children, so we'll follow Bates and use a single indicator for any children, losing 0.37 χ^2 points while saving 2 d.f.

A final improvement is to add an interaction between the linear term on age and the indicator for children. This allows the curve for mothers to have a different peak and improves the fit by 8.00 chi-squared points at the expense of one d.f.

The Random-Intercept Model

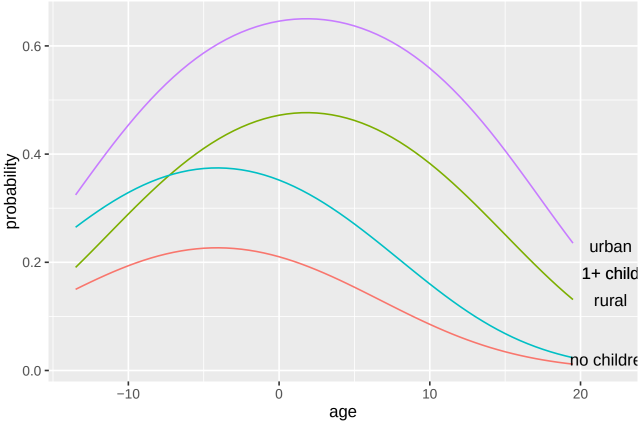
Here's the Stata output from the final random-intercept model

```
Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =     146.77
Log likelihood  = -1182.4584                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       c_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       urban |   .7134563   .1213548     5.88   0.000     .4756053    .9513074
         age |  -.0472872    .021841    -2.17   0.030    -.0900949   -.0044795
        age2 |  -.0057577   .0008414    -6.84   0.000    -.0074068   -.0041086
       child |   1.210876   .2075937     5.83   0.000     .8039994    1.617752
    ageXchild |   .0683467   .0254687     2.68   0.007     .0184289    .1182645
       _cons |  -1.323606   .2154606    -6.14   0.000    -1.745901   -.9013106
-------------+----------------------------------------------------------------
    /lnsig2u |   -1.48611   .3397138                     -2.151937   -.8202836
-------------+----------------------------------------------------------------
     sigma_u |   .4756585   .0807939                      .3409673    .6635561
         rho |   .0643468   .0204529                      .0341322    .1180392
------------------------------------------------------------------------------
LR test of rho=0: chibar2(01) = 44.46                  Prob >= chibar2 = 0.000
```

Results using R's `glmr()` are very similar. Try your hand at interpreting these results before peaking at the next slide.

Use by Age, Children and Residence

Seems clear that contraceptive use increases with age and then declines, is higher for women with children than those without, peaks at a later age for women with children, and is generally higher in urban areas. These effects are best shown in a graph



Variation in Use Across Districts

There is also evidence of substantial variation in contraceptive use across districts:

- The estimated standard deviation of the intercept, 0.476, means that the odds of using contraception are 60% higher in a district one standard deviation above the mean than in an average district, everything else being equal.
- The intraclass correlation between the latent propensity to use contraception of two women in the same district is 0.06. Equivalently, we can say that districts account for only 6% of the variation in propensity to use contraception net of the observed covariates.

In case you are curious the manifest correlation at the median, calculated using `xtrho`, is equivalent to an odds ratio of 1.23.

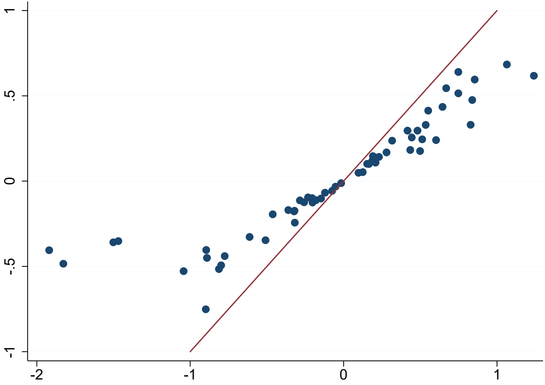
Estimation of the Random Effects

We can identify districts where women are more or less likely to use contraception by predicting the random effects. There are two ways to proceed:

1. Calculate maximum likelihood estimates by treating the estimated linear predictor from the multilevel model as an offset and then running a separate logit model in each district. The estimate is not defined when all women in a district have the same outcome, which happens in three districts.
2. Compute empirical Bayes estimates using the mean or mode of the posterior distribution of the random effects, which requires using numerical integration.

Comparison of ML and EB estimates

The graph below compares EB and ML estimates and shows the usual shrinkage towards zero.



The shrinkage is particularly noticeable in four districts, all with fewer than 15 women and effects quite far from zero.

Predicted Probabilities

Subject-specific probabilities are easily computed from first principles by setting the observed covariates and the random effects to selected values. The predicted probabilities for women of average age with children in urban and rural areas of the average district are 0.6458 and 0.4718, an odds-ratio of 2.04.

Population-average probabilities can also be computed, although they require integration over the distribution of the random effect, which can be done "by hand" or using `gllamm`. Using 12-point quadrature we obtain population-average probabilities of 0.6389 and 0.4732, or an odds-ratio of 1.97

As usual the population average effect is smaller than the subject-specific, but the difference here is modest because the intra-class correlation is low.

The Random-Slope Model

The next step is to see whether the urban-rural differential in contraceptive use varies by district, which we'll do by treating the urban effect as a random slope.

This model is analogous to allowing an interaction between urban residence and district, but instead of estimating a separate urban-rural difference for each district, we assume that they are drawn from a normal distribution. Estimation is possible because most districts have urban and rural areas; in fact we find only 15 districts with no rural women and 3 with no urban women.

The Random-Slope Model

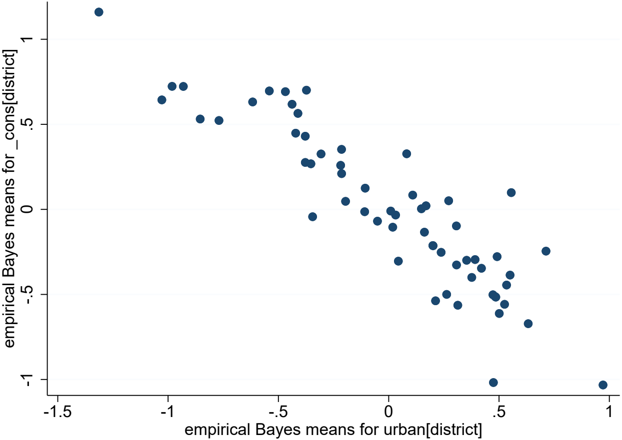
Here's the output from Stata using 7 points per effect

```
 Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(5)      =     135.42
Log likelihood = -1176.4767                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------------
           c_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
           urban |   .7906825   .1648731     4.80   0.000     .4675372    1.113828
             age |  -.0461515   .0219589    -2.10   0.036    -.0891903   -.0031128
           agesq |  -.0056484   .0008514    -6.63   0.000    -.0073171   -.0039797
           child |   1.211711   .2091521     5.79   0.000     .8017802    1.621641
       ageXchild |    .066423   .0256306     2.59   0.010     .0161879     .116658
           _cons |  -1.344866   .2244245    -5.99   0.000     -1.78473   -.9050024
-----------------+------------------------------------------------------------------
district         |
       var(urban)|   .5453645   .2931897                      .1901421    1.564212
       var(_cons)|   .3859845   .1280172                      .2014915    .7394059
-----------------+------------------------------------------------------------------
district         |
 cov(urban,_cons)|   -.363198   .1660099    -2.19   0.029    -.6885714   -.0378246
------------------------------------------------------------------------------------
LR test vs. logistic model: chi2(3) = 56.42               Prob > chi2 = 0.0000
```

The negative covariance should reinforce the importance of specifying `covariance(unstructured)`.

Empirical Bayes Estimates

Let us look at estimates of district effects on rural levels and urban-rural differentials in contraceptive use. We could compute ML estimates as we did for the random intercept model, but I will focus on EB estimates



Empirical Bayes Estimates (continued)

We see a clear negative correlation as noted earlier. Districts where contraceptive use in rural areas is higher than expected after considering the age and motherhood status of women, tend to have a smaller urban-rural differential in contraceptive use.

An alternative parametrization estimates separate urban and rural levels and omits the constant in the fixed and random parts. This formulation leads to exactly equivalent estimates of the fixed part but the two random effects turn out to be nearly independent. Details are left as an exercise.

Immunization in Guatemala

The final example is our analysis of childhood immunization in Guatemala. This is a three-level dataset with 2159 children of 1595 mothers who live in 161 communities, analyzed in our *Demography* paper and RG2, and used as a detailed illustration of 3-level models in MALMUS §16.2-16.8, pages 873–897.

The sample consists of children age 1-4 who have received at least one immunization, and the outcome of interest is whether they have received the full set appropriate for their age. Predictors include

1. age of child at child level,
2. mother's ethnicity and education and father's education at the family level, and
3. urban indicator and percent indigenous in 1981 at the community level.

We will return to this dataset when we compare those results with Bayesian methods.