FI=HLML05

MORE HLM SUBMODELS:

RANDOM COEFFICIENTS REGRESSION (RCR) MODEL

AND

INTERCEPTS AND SLOPES AS OUTCOMES (IASAO)

AND

A COUPLE OF MINOR POINTS AT THE END (True Scores, Deviance)

by
Ralph B. Taylor
breck@rbtaylor.net

The RCR model represents an additional elaboration over the ANCOVA model. Now we <u>will</u> allow L-1 slopes to vary.

In the RCR model we simply allow this variation to happen, and see how big it is.

With IASAO, for the first time ever, we will try and <u>predict</u> these differences in L-1 slope with a L-2 predictor.

RCR

Our L-1 model is the same as when we are working with 1 group centered predictor (B&R Eq. 2.2):

YIJ = B0J + B1J(X-X.j) + RIJ

Note that X has been group centered so we are looking at "frog pond" effects.

But now at level two, in addition to modeling the intercepts, which we have already done in the ANOVA

B0J = G00 + U0J [B&R Eq. 2.17a]

we will now try to see how much variability we have associated with the L-1 slope across the J units:

B1J = G10 + U1J [B&R Eq. 2.17b]

the new terms we have here include:

G10 "the average regression slope across the Level-2 units" (p. 20). This has been precision weighted

U1J "the unique increment to the slope associated with L-2 unit J" (p. 20). This will have <u>a different</u> value for each individual neighborhood; when we start going into residual files we will be able to see what this is like for each neighborhood; for example, in what neighborhoods does age have a stronger effect on fear; in what neighborhoods does it have a weaker effect?

VARIANCE OF U1J = T11 (unconditional)

Notice at this point that since we have both random effects for intercepts, and random effects for slopes at L-2, we can see how these effects relate to each other; we can examine the unconditional covariance of U0J and U1J.

COV(U0J, U1J) = T01

The full equation we get with this model is as follows (B&R Eq. 2.19):

YIJ = G00 + G10(XIJ-X.J) + U00J + UIJ((XIJ-X.J) + RIJ

IASAO

Now at long last we come to the full model where we have both slopes and intercepts as the outcomes.

These equations are shown in Eq. 2.2 and 2.4 in B&R

Our Level-2 model becomes:

B0J = G00 + G01(WJ) + U0J

B1J = G10 + G11(WJ) + U1J

G01 We have already seen before; it is the impact of the L-2 predictor on the group mean true score estimate

G11 This is the last term to introduce. It explains the slope of a L-2 predictor **on a L-1 slope**

MINOR POINTS

“True Scores”

With the group means (B0J): I have been saying we end up with "true scores" of each group on the Y outcome.

To be completely correct, these are **estimates** of true scores of each group on Y, derived using empirical Bayes techniques and precision weighting and all the stuff discussed in Chapter 3.

A note on deviance: the reason the value does not make any sense is because it only makes sense <u>when we are comparing two different models</u>.

See B&R p. 56.

Deviance Statistic

D is a measure of "lack of fit" and is "-2 times the value of the log likelihood function ... the higher the deviance, the poorer the fit"

To get clearer on likelihood and log likelihood, here are some quotes from Darlington (1990):

> Consistency between the data and a model is measured by the likelihood or probability of observing the particular data if the model is correct (442) .... the model’s likelihood or consistency with the data equals the product of the values of Fit(i) for all i subjects (443)

If PS(i) = probability of success for person i , and 1 - PS(i) = probability of failure for person i

> Fit(i) = Y(i) X PS(i) + (1- Y(i)) X (1 - PS(i)) [p. 443]

> Values of likelihood can be extremely small in large samples ... therefore we usually report the natural logarithms of likelihood values rather than the raw values. But these logarithms are always negative, since a likelihood is always below 1; thus the usual value reported is -ln(likelihood). We shall denote this as NLL for negative log likelihood. NLL is always positivce and measures <u>lack of fit</u> between data and model; the smaller the value the better the model fits the data. Since the sum of the logarithms is the logarithm of the product we have: NLL = -(sum(ln(fit(i))) (443)

> Other measures of overall fit are more closely related to chi square tests on an overall model .... If the regressors in one model are a subset of those in another, then the difference in fit between the two models can be tested with the test (449):

> chi square = 2 X (larger NLL - smaller NLL) [p. 449]

Deviance can be used to test a <u>composite</u> hypothesis; rather than testing one parameter, you compare two different models, <u>where one is an elaboration of another</u>.

Then you compare the two D values, and test the difference using a chi square test with m df where m = number of parameters different for a model.

But the two models you are comparing "must be identical with respect to the specification of the fixed effects."

So what you are seeing is the differences in adding <u>random</u> effects.

On p. 75 and 76 B&R give an example. In a full HLM model they allow slopes and intercepts to covary; they also allow the level-1 slopes to vary.

When happens if they restrict these conditions to 0?

They see how much D changes. Since they restricted 2 parameters, they test it against chi square with 2 df and find it NS.

In other words, the more restrictive model, not allowing covariance of slopes and intercepts, and restricting the variance of L-1 slopes to 0, results in a model that fits the data no more poorly than the unrestricted model. (Restricting the variance of the slope to 0 means forcing the same slope across schools.)

Of course, they had a clue to this to begin with. See Table 4.5, p. 72. Notice that the variance associated with the SES achievement slopes is NS.

Kreft and DeLeeuw see deviance in a slightly different light. They do not talk about how it can only be limited to comparing different models with the same fixed effects but added varying effects. Instead they suggest that <u>you should use deviance as an omnibus test</u>, as a preliminary to looking at whether particular coefficients are significant. See the discussion on pages: 54, 64, 65.

New Reference
Darlington, R. (1990). Regression and linear models. New York: McGraw Hill.