STATISTICAL HORIZONS

Multilevel Modeling

Tenko Raykov, Ph.D.

Upcoming Seminar:

April 7-8, 2017, Philadelphia, Pennsylvania

Multilevel Modeling

Part 1 – *Introduction,* *Basic and Intermediate Modeling Issues*

Tenko Raykov
Michigan State University

(www.msu.edu/~raykov)

Plan:

0. Resources for this course and what it is about.

1. Why do we need multilevel modeling (MLM), and how come aggregation and disaggregation do not do the job?

2. The beginnings of MLM – Why what we already know about regression analysis is so useful.

3. The intra-class correlation coefficient – The basics of a multilevel model.

4. Proportions of variance at third level and at intermediate level in three-level settings, and how to evaluate them.

5. The random intercept model and model adequacy assessment.

6. Robust (ordinary least squares-based) modeling of lower-level variable relationships in presence of clustering effect.

Appendices:

- terminology in multilevel modeling (glossary, notation);
- reshaping data (if needed) for modeling with Stata.

0. Resources for this course and what it is about

Literature:

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling with Stata* (3rd Edition). College Station, TX: Stata Press.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear and nonlinear modeling* (2nd Edition). Thousand Oaks, CA: Sage.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent linear and mixed models*. Boca Raton, FL: Chapman & Hall.

Snijders, T. A. B., & Bosker, R. (2012). *Multilevel models. An introduction to basic and advanced modeling* (2nd Edition). Thousand Oaks, CA: Sage.

Software:

- **Stata.**

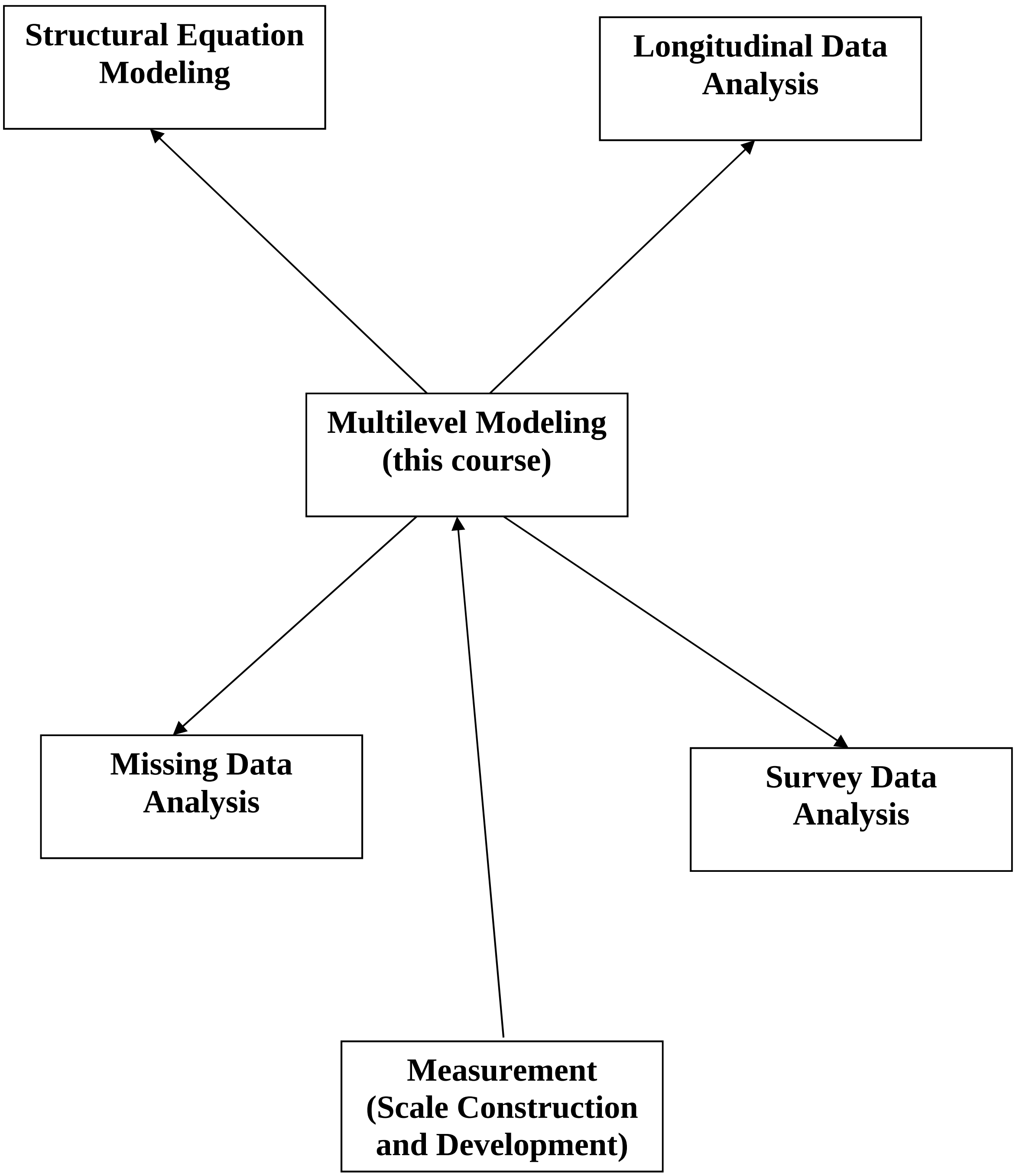
Further excellent software and voluminous literature on multilevel modeling/hierarchical linear and nonlinear modeling are available as well (see also papers sent along with data files to participants).

Goals of course:

It is application oriented but with coherent discussion of theoretical issues at an introductory to intermediate level, with a few more advanced issues.

***Disclaimer*: Data sets are used only for method illustration purposes.**

***Note*. This course provides connections to the following main applied statistics areas (methodologies; see alternative/possible short courses).**



What this course is about:

It is about the applied statistical modeling methodology usually referred to as:

- **Hierarchical (linear and nonlinear) modeling,**
- **Random effects modeling,**
- **Random coefficient modeling,**
- **Mixed effects modeling,**
- **Mixed modeling,**
- **Generalized linear mixed modeling,**
- **Variance component models.**

These are largely synonyms for *multilevel modeling*.

Closely related are these fields (effectively synonymous here):

- **Longitudinal modeling,**
- **Panel data modeling,**
- **Repeated measure analysis,**
- **Cross-sectional time series analysis.**

The common theme underlying the above and this course is:

> ***Regression modeling when data are clustered (nested, hierarchical) in some way.***

Standard applied statistical modeling – e.g., regression analysis (OLS) – does *not* handle this case (as it makes the assumption of independence).

Clustered data sets are *very rich* since they contain information about processes at *different* levels (e.g., personal *and* institutional characteristics) as well as, by implication, about their interactions.

Multilevel modeling allows one to (1) *disentangle* them, (2) study them *simultaneously*, and (3) examine their *interactions*, all this (4) permitting answering *complex questions* about studied phenomena.

1. Why do we need multilevel modeling, and how come aggregation and disaggregation do not do the job?

Data from empirical studies in the social, behavioral, and biomedical sciences as well as in business, marketing, and economics very often exhibit distinct *hierarchical (multi-level) structure*. The reason is that studied individuals are (already) *grouped* into larger units—e.g., firms, companies, industries; neighborhoods, cities, states, countries; centers, physicians, hospitals; families, dyads, twins pairs, etc. This nesting or clustering may have an effect on the subjects' scores on measures used, in particular on outcome scores, which thus need *not* be <u>un</u>correlated.

The units of analysis – on which the dependent variable is measured – are usually referred to as *level*-1 units. The higher-order units in which they are nested (firms, companies, cities managers) are called *level*-2 units, also often referred to as 'groups' for simplicity (as we will do not infrequently below).

In some studies, depending on the research question, it is possible that other measurements are considered level-1 units (e.g., subsection 1.2 below). In general, if the level-2 units are in turn nested themselves within even higher-order units, the latter are called *level*-3 units, and so on (e.g., Section 4).

Consider the following illustrations.

1.1. Examples of nested data and the hallmark of MLM

1. **(*Business, Industrial/Organizational Psychology*) Employees are nested (clustered) within companies: former are level-1 units, the latter are level-2 units.**

2. **(*Economics/Sociology*) Individuals nested (clustered) within cities: former - level-1 units, latter - level-2 units; also,**

interviewees (level-1 units) are nested w/in interviewers (level-2 units).

3. (*Management*) Workers are nested within managers; employees are nested within teams.

The list of examples can go on, and they all share this *common feature*: the units in a lower level of the hierarchy are *nested* (grouped, clustered, similar, 'correlated', inter-related) within the units of its higher level(s).

Why is the issue of nesting, or clustering, of subjects relevant?

Because nesting implies a possibly serious *lack of independence* of individual scores on the dependent variable(s) of concern (abbreviated to DV and denoted *Y* in this course;
we refer to the DV as a *response* or *outcome* variable in it).

The reason is that the <u>*Y* scores of subjects within the level-2 units are in general *correlated* (more similar), unlike the *Y* scores of subjects from different level-2</u> (or highest level) <u>units</u>.

This nesting (clustering, hierarchy, similarity, correlation) is the *hallmark* of multilevel modeling. (I will use the abbreviation "MLM" for multilevel/hierarchical modeling, the subject of this course, rather than the abbreviation "HLM" that is reserved for a popular software program for MLM.)

To examine in more detail this feature of MLM, let's look at the scores on say a job satisfaction (JS) measure in a given industry. Here, to properly study their properties, we need to keep in mind that *employees are nested within firms (companies)*.

Suppose that we (i) randomly sample 30 firms/companies from the industry in question and then (ii) randomly sample 50 say employees from within each firm (who have been administered the JS measure).

Then, due to the fact that *employees in the same firm <u>share</u> the same working and related conditions* (as well as possibly same manager/s), and a host of other *experiences* of various kinds, their *Y* scores will tend to be *correlated* ('*similar*'; see next).

Let's look at the following illustration (to clarify this important point).

Employee (ID1)	Firm ID (ID2)	JS Score (*Y*)
1	Company 1	45
2	Company 1	46
3	Company 1	44
4	Company 1	42
5	Company 2	79
6	Company 2	78
7	Company 2	77
8	Company 2	75
9	Company 3	92
10	Company 3	91
11	Company 3	93
12	Company 3	94
… etc.		

Here, even though firms/companies and employees within them are randomly chosen, the individual worker's JS *scores <u>within</u> company seem to be relatively <u>similar</u>* ('correlated'). Specifically, *(any) two JS scores <u>within</u> firm are more similar than (any) two across (<u>between</u>) firms*.

This within-firm similarity is unlike any such for scores across companies (which scores in general might be so too, though, but often to a lower extent). This property, often referred to as *within-cluster or within-group correlation*, is a fundamental feature that will go like a red thread throughout this course.

All standard statistical methods (e.g., the general and generalized linear models, classical SEM) do *not* account for this lack of independence in the response variable scores across subjects. In particular, the *standard* (traditional, conventional) methods—like regression analysis, ANOVA, multivariate statistics, SEM, analysis of categorical data, etc.—do assume that these subjects' scores are *independent.* For this reason, they are called *single-level methods*.

Hence, to the degree to which this assumption is violated, the results of an application of those classical/standard methods on hierarchical data of the kind we have been discussing so far, will yield *less trustworthy if not even misleading* results.

A rather frequent consequence of a serious violation of the above independence assumption is the phenomenon of *spurious significance*, if this violation is neglected (see also Rabe-Hesketh & Skrondal, 2012, for alternative examples).

This phenomenon follows from:

- **(i) ignoring the nesting (clustering) of subjects or the lowest units in the analyzed data,**
- **(ii) proceeding then with the above mentioned conventional methods (single-level models/methods), which tends to yield *too many rejections of pertinent hypotheses,***
- **(iii) usually *spuriously small standard errors*, which also lead to *too short confidence intervals*, and *too small p-values* associated with statistical tests.**