STATISTICAL MODELS FOR STUDENT EVALUATION IN THE UNIVERSITY EDUCATIONAL PROCESSES

Bologna, 6th February 2007

Dipartimento di Scienze Statistiche "Paolo Fortunati"

Multilevel models as a tool for student guidance

Leonardo Grilli & Carla Rampichini

grilli@ds.unifi.it carla@ds.unifi.it

Department of Statistics "Giuseppe Parenti"

University of Florence

Outline

- Effectiveness of educational institutions
- Multilevel models for assessing effectiveness
- Estimation and prediction
- Limitations
- References

Effectiveness of educational institutions

Student results (outcomes)

- High school graduates planning to enrol in University need information on:
  - chances of obtaining a degree
  - time to graduate
  - occupational opportunities after degree
- Student results depend both on her characteristics and university quality

Effectiveness of educational institutions (schools, universities)

- Remark: the outcome of an educational institution cannot be defined in absolute terms, but only *with respect to the effects on the students*
- Problem: the effects on the students are affected by the features of the students themselves (if two institutions of similar quality have students with rather different motivation, ability... the outcome of the two institutions is likely to be quite different)

How to make a fair assessment?

Value added

- The analysis of the educational process is difficult ➔ the quality of educational institutions is usually measured via an **input/output** approach:
  - the process is a sort of *black-box*
  - the output (*outcome*) is evaluated in the light of the input → the effectiveness is just the **value added** by the school:

VALUE-ADDED = ACTUAL OUTCOME
minus
EXPECTED OUTCOME GIVEN THE INPUT

Internal/external effectiveness

The educational process leads to multiple outcomes → many measures of effectiveness

- Internal effectiveness:
  - Dropout (1=Yes, 0=No)
  - Duration of studies (time to the degree)
  - Number of credits after a given period
- External effectiveness:
  - Occupational status after degree (1=Yes, 0=No)
  - Duration of unemployment (time to first job)
  - Wage or job satisfaction

7

Need for value added analysis

- Empirical research has found that the differences in student outcomes across schools are due
  - mainly to differences in student prior achievement and socio-economic background
  - for a minor part to differences in school factors such as teachers ability, organization...
- Thus comparing the unadjusted outcomes is markedly unfair and a value added approach is needed

8

Type A and B effectiveness

A potential student (or its family) and the government are interested in different types of effectiveness:

- **Type A - Potential student**: interested in comparing the results she can obtain by enrolling in different institutions, *irrespective of the way such results are yielded*
- **Type B - Government**: interested in assessing the "production process" in order to evaluate the ability of the institutions to exploit the available resources

The two types of effectiveness are called **A** and **B** after Raudenbush & Willms (1995)

9

Type A and B effectiveness

- **Type A**: performance of the institution *adjusted* for the features of the students
- **Type B**: performance of the institution *adjusted* for the features of
  - the students
  - the context in which it operates (e.g. resources, local labour market, socio-economic composition of enrolled students)

In practice the adjustment required for the assessment of **Type B effectiveness** is not easy (many variables whose measurement is problematic)

10

Student guidance

- Student guidance requires evaluation of **Type A** effectiveness
- The educational process leads to multiple outcomes → many measures of Type A effectiveness
- Any student gives different weights to the outcomes according to her preferences → the evaluation system should avoid summarizing the various kinds of effectiveness into a single overall indicator

11

Multilevel models for assessing effectiveness

Statistical issues

- The statistical models for assessing the relative effectiveness of educational institutions must face two main issues:
  - Adjustment: the measures must be adjusted at least for the features of the students (necessary for a fair comparison)
  - Quantification of uncertainty: this is necessary in order to make assessments strongly supported by empirical evidence (avoiding judgements that may be originated by the sampling variability or other sources of error)

The raw rankings (so called 'League Tables') ignore both issues (Goldstein & Spiegelhalter, 1996)

13

Statistical issues

Adjustment &

Quantification of uncertainty

Regression models

But *standard* models are not adequate!

- Standard models make unsuitable assumptions on the variance-covariance structure (independence among observations, while the results of the students of the same school usually are positively correlated) ➔ poor quantification of uncertainty
- Standard models are unable to represent some key features, e.g. varying slopes

14

Multilevel models

- Multilevel (mixed, random effects) models overcome the main limitations of standard models and are well suited for assessing the relative effectiveness of schools
  - The effectiveness of a school is explicitly represented by the random effects

15

Random intercept model

Features of the student | Features of the school/context

Outcome of the student | School random effect

$$Y_{ij} = \alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + u_j + e_{ij}$$

i = student

j = school

$$= (\alpha + u_j) + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + e_{ij}$$

Intercept of j-th school

16

Standard assumptions

1. IID errors at each level
2. Errors have a Normal distribution with zero means
3. Errors at any level are uncorrelated with the covariates (level 1 and 2 exogeneity)
4. Errors at different levels are uncorrelated

17

Value added interpretation

Actual outcome | Expected outcome given student and school/context features

$$Y_{ij} - \left(\alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j\right) = u_j + e_{ij}$$

i = student

j = school

The difference between actual and expected outcome is decomposed in two parts:

- School-level component (random effect) u_j
- Student-level component e_{ij}

The random effect u_j is the school value added, or effectiveness. It is a *residual* term ➔ its meaning depends on which covariates are in the model

18

Type A effects from random intercept model

$$Y_{ij} = \alpha + \beta x_{ij} + \gamma \bar{x}_j + u_j + e_{ij}$$

$$A_{ij} = A_j = \underbrace{\gamma \bar{x}_j}_{\text{context}} + \underbrace{u_j}_{\text{practice}}$$

Uniform TYPE A effect of school j

- Y_{ij} n. of credits by the end of 1st year (student i, school j)
- x_{ij} student grade (final high-school grade "maturità")
- $\bar{x}_j$ average grade of the students enrolled in the j-th school
- β within-school slope relating X to Y
- γ contextual effect

All school factors beyond $\bar{x}_j$ are included in u_j (note: $\bar{x}_j$ cannot be omitted if $\gamma \neq 0$, since then β would be biased)

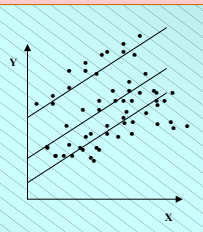
19

Random intercept model

- Uniform effects ➔ same school effect for all the students
- Constant slopes ➔ parallel regression lines
- Ranking of the schools on the basis of the Type A effect

Intercept j-th school: $\alpha + \gamma \bar{x}_j + u_j$

Slope j-th school: β



$$A_j = \gamma \bar{x}_j + u_j$$

20

Random slope model

Features of the student — Features of the school/context — School random effects

Outcome of the student

$$Y_{ij} = \alpha + \boldsymbol{\beta} \mathbf{x}_{ij} + \boldsymbol{\gamma} \mathbf{w}_j + u_{0j} + \mathbf{u}_{1j} \mathbf{z}_{ij} + e_{ij}$$

i = student
j = school

$\mathbf{z}$ is a subset of $\mathbf{x}$

$$= \underbrace{(\alpha + u_{0j})}_{\text{Intercept of } j\text{-th school}} + \underbrace{(\boldsymbol{\beta}_1 + \mathbf{u}_{1j})}_{\text{Slopes of } j\text{-th school on covariates } \mathbf{z}} \mathbf{z}_{ij} + \ldots$$

21

Type A effects from random slope model

$$Y_{ij} = \alpha + \beta x_{ij} + \gamma \bar{x}_j + u_{0j} + u_{1j} x_{ij} + e_{ij}$$

$$A_{ij} = \underbrace{\gamma \bar{x}_j}_{\text{context}} + \underbrace{u_{0j} + u_{1j} x_{ij}}_{\text{practice}}$$

TYPE A effect of school j on student i

- Y_{ij} n. of credits by the end of 1st year (student i, school j)
- x_{ij} student grade (final high-school grade "maturità")
- $\bar{x}_j$ average grade of the students enrolled in the j-th school
- β within-school slope relating X to Y
- γ contextual effect

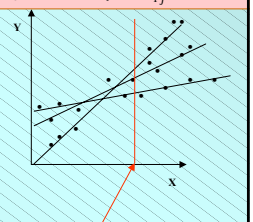
22

Random slopes model

- Varying effects ➔ different school effects, depending on student characteristics
- random slopes ➔crossing regression lines
- No unique ranking of the schools ➔ different rankings conditionally on student characteristics

Intercept j-th school: $\alpha + \gamma \bar{x}_j + u_{0j}$

Slope j-th school: $\beta + u_{1j}$



1. Define student profiles
2. Build rankings by profile

23

Outcomes and models

The nature of the outcome determines the kind of multilevel (mixed) model

Outcome	*Mixed model*	
Continuous (e.g. wage)	linear	Generalized Linear Mixed Models
Binary (e.g. dropout)	logit, probit	
Count (e.g. credits)	poisson	
Time (e.g. time to degree)	duration	

24

Estimation and prediction

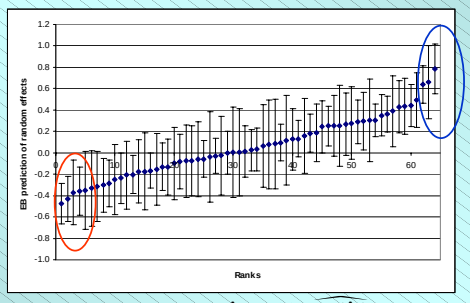
Student guidance from the model

- Evaluation of Type A effects allows:
  - Ranking the schools
  - Predicting the outcome for a given student
- Need estimation of model parameters and prediction of random effects

Estimation and prediction

- Model parameters
  - Slopes: α, β and γ
  - Stand. dev. and covariances: σ_e σ_{u_0} σ_{u_1} $\sigma_{u_0u_1}$
- Model errors
  - Level 1 errors e_{ij}
  - Level 2 errors (random effects): u_{0j} u_{1j}
- Two approaches:
  - ML for the parameters + empirical Bayes for the random effects (two distinct steps)
  - Bayesian (all at once)

Uncertainty of predicted random effects



Intervals for pair-wise comparisons $\hat{u}_j \pm 1.39 \times \widehat{SE(\hat{u}_j)}$ e.g. Goldstein (2003)

The interval width depends on the cluster size

Only extreme values differ significantly!

Prediction of the outcome

- After estimation of the parameters and prediction of the random effects it is easy to predict the outcome for a given student in a given school:

$$\hat{Y}_{ij} = \hat{\alpha} + \hat{\beta} x_{ij} + \hat{\gamma} \overline{x}_j + \hat{u}_j$$

student school

- The university could build a system where the student plug-in his characteristics and obtain the predicted outcome for every school

Limitations of value added approach

- Need more information to understand why some schools are more or less effective
- Studies of school effects are quasi-experiments → causal conclusions are questionable
- An effective adjustment for the input requires several good-quality covariates
- Measurement error in the covariates (especially prior achievement) may bias the slope estimates
- Difficult to fully account for all the uncertainty
- Difficult to communicate the results to a non specialized audience

References

- Raudenbush SW & Willms JD (1995) The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*, 307-335.
- Goldstein H & Spiegelhalter DJ (1996) League tables and their limitations: statistical issues in comparisons of institutional performances. *JRSS A*, *159*, 385-443.
- Draper D & Gittoes M (2004) Statistical analysis of performance indicators in UK higher education. *JRSS A*, *167*, 449-474.
- Battauz M, Bellio R & Gori E (2005) A multilevel measurement error model for value-added assessment in education. Atti del Convegno S.Co. 2005 (Bressanone, 15-17 settembre), 91-96.
- Bird SM, Cox D, Farewell VT, Goldstein H, Holt T & Smith PC (2005) Performance indicators: good, bad, and ugly, *JRSS A*, 168, 1-27.
- Fabbris L (Ed.) (2007) *Effectiveness of University Education in Italy: Employability, Competences, Human Capital*, Heidelberg: Springer-Verlag.
- Chiandotto B, Grilli L & Rampichini C (Eds.) (2005) *Valutazione dei processi formativi di terzo livello: contributi metodologici*, Collana Valmon n. 12, Università degli studi di Firenze. **http://valmon.ds.unifi.it**

31