"Causal inference for policy evaluation: case studies and statistical complications"
Universitat Pompeu Fabra, Barcelona, 30th March 2012

Causal inference through principal stratification:

basic ideas and an application to the effect of university studies on job opportunities

Leonardo Grilli
grilli@ds.unifi.it
www.ds.unifi.it/grilli

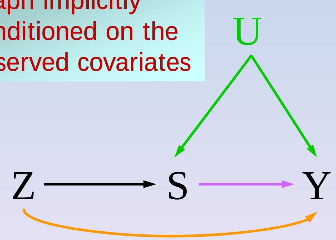
Department of Statistics
University of Florence

Outline

- Causal inference with intermediate variables
- Principal stratification, with emphasis on
  - Bias from conditioning on post-treatment variables
  - Truncation by death
- Case study: Relative effectiveness of two degree programmes with respect to employment

Causal inference with intermediate variables

Graph implicitly conditioned on the observed covariates



U = unobserved variables

Z = treatment

S = intermediate

Y = response

(S, Y: Post-treatment variables)

Two causal estimands of interest:

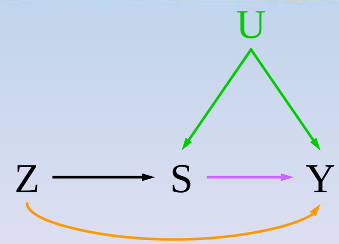
Z → Y   S → Y

Unconfoundedness assumption: conditionally on the observed covariates there are no unobserved confounders,

i.e. no arrow U → Z

Relevant intermediate variables

In causal inference there are cases where we <u>cannot ignore</u> the intermediate variable S:



1) When S is the "real" treatment of interest, e.g. in studies with ***noncompliance***, where Z is merely the treatment assignment
2) When Y is not observed, or even not defined, depending on the value of S, e.g. S is an indicator of ***nonresponse***, or S is a variable whose value determines the existence of Y, e.g. S is the ***survival indicator*** and Y is the quality of life (*truncation by death*)
3) When it is of interest to disentangle the total effect of Z on Y into the ***direct effect*** and the indirect effect through S

Broken balancing

- Under *Unconfoundedness* all pre-treatment variables are balanced among treatment groups
- In general, the balancing property is corrupted for post-treatment variables
- This fact may lead to biases if the post-treatment variable is *relevant*, i.e. if you want to
  - condition on a post-treatment variable, or
  - estimate a causal effect for a post-treatment variable

→ A solution: define potential outcomes for all the post-treatment variables

Potential outcomes

- For each *relevant* post-treatment variable there is one potential version for each level of the treatment
- Every statistical unit is assigned to one and only one level of the treatment → only one of the potential versions is observable
- In case of a binary treatment Z=0,1 then
  - S(1), S(0) ⇒ S=S(Z) is the observed version
  - Y(1), Y(0) ⇒ Y=Y(Z) is the observed version

Principal strata

Simplest case: both Z and S binary → 4 strata

Z=1 drug

S=1 get well

Z	sicker	opposite	responsive	healthier
1	0	0	1	1
0	0	1	0	1

Principal strata are defined by the values of the two potential versions of the intermediate variable S (counterfactual) → they are not influenced by the value taken by Z (like pre-treatment covariates)

Observed values of Z and S do not identify the stratum: if Z=1 and S=1 the unit can belong to two strata: 10 (responsive) or 11 (healthier)

Principal strata are latent classes (→ ***latent class models***)

Principal Causal Effects

Causal effect of Z on Y for a single unit: $Y_i(1)$ vs. $Y_i(0)$

Principal Causal Effect (PCE) of Z on Y:

$f(Y(1))$ vs. $f(Y(0))$ for the units of a principal stratum

Causal effects across principal strata are nonsense!

Conditioning on the *observed* value of the intermediate variable S implies conditioning on different principal strata depending on the value of Z

Frankgakis C.E. & Rubin D.B. (2002) Principal stratification in causal inference, *Biometrics*, 58: 21-29.

Barnard J., Frangakis C.E., Hill J.L. & Rubin D.B. (2003) Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York City, *JASA*, 98: 299-323.

Conditioning on intermediate variables /1

- Conditioning on a post-treatment intermediate variable, often called 'concomitant' variable, is a common practice (it was recommended even by R.A. Fisher), but it gives wrong conclusions
- This type of error can be easily recognized if the problem is cast in the principal stratification framework
- Rubin (JASA 2005) gives the following example
  - Suppose a very large randomized experiment where half of the plots are assigned a new fertilizer and half a standard fertilizer
  - Z = treatment indicator (0=standard fertilizer; 1= new fertilizer)
  - S = number of plants established in each plot
  - Y = yield in each plot

Conditioning on intermediate variables /2

- Each post-treatment variable has potential and observed versions
  - S(1), S(0) whereas S=S(Z) is the observed version
  - Y(1), Y(0) whereas Y=Y(Z) is the observed version
- Suppose we wish to estimate the effect of Z on Y controlling for S (i.e. the effect of the new fertilizer on the yield controlling for the number of plants)
- The standard approach is ANCOVA conditioning on the observed intermediate variable S

$$Y_i = \alpha + \beta Z_i + \gamma S_i + error_i$$

- Hypothetical situation with a treatment effect on the intermediate S, but no treatment effect on the primary outcome Y; example with ***two principal strata***: normal plots and good plots (= more plants)

Fraction of pop.	Potential outcomes				PS	Observed data		
	S(1)	S(0)	Y(1)	Y(0)		Z	S	Y
1 / 4	3	2	10	10	normal plots	0	2	10
1 / 4	3	2	10	10		1	3	10
1 / 4	4	3	12	12	good plots	0	3	12
1 / 4	4	3	12	12		1	4	12

- If we control for the observed values of the intermediate S, the comparison between Y under treatment and Y control is possible only for S=3 → the treatment effect on the primary outcome Y is estimated to be -2 (this is also the estimate of β from the ANCOVA model)
- What's wrong? The comparison is not fair but confounded by the quality of the plots: we are comparing the *yield of the new fertilizer in bad plots* with the *yield of the standard fertilizer in good plots* (i.e. we are estimating an effect using units belonging to different principal strata!)

Conditioning on intermediate variables /4

- Another way to see what's wrong with conditioning on the post-treatment intermediate variable S, is that such conditioning destroys the unconfoundedness of the assignment mechanism
- Recall we assumed a 50/50 randomization:

$$Pr\left(Z_i = 1 \mid Y(0), Y(1)\right) = Pr\left(Z_i = 1\right) = 0.5$$

- When we condition on S such probability depends on Y(1) and thus the assignment mechanism is confounded

$$Pr\left(Z_i = 1 \mid S_i, Y_i(0), Y_i(1)\right) = \begin{cases} 1 & \text{if } S_i = 3 \text{ if } Y_i(1) = 10 \\ 1 & \text{if } S_i = 4 \\ 0 & \text{otherwise} \end{cases}$$

Discrete principal strata & LC models

- In most applications both the treatment variable and the intermediate variable are discrete (often binary) → principal strata are discrete (latent classes) → the statistical model is a special type of LC (Latent Class) model
- The connection with LC models has been recognized in the case of non-compliance (CACE):
  - Jo B., Asparouhov T. & Muthen B. (2008). Intention-to-treat analysis in cluster randomized trials with noncompliance. *Statistics in Medicine*, 27, 5565–5577
  - Skrondal A. & Rabe-Hesketh S. (2004) Generalized Latent Variable Modeling. Chapman & Hall.
- The connection with LC models is discussed in general terms in
  - Grilli L. (2011) Causal inference through principal stratification: a special type of latent class modelling. In Fichet B, Piccolo D, Verde R, Vichi M (Eds) Classification and Multivariate Analysis for Complex Data Structures. Springer-Verlag. pp 265-270. *Download a draft from* `www.ds.unifi.it/grilli`

Continuous principal strata

- Principal strata can also be continuous (e.g. partial compliance in clinical trials: binary treatment & continuous intermediate variable)
- In this framework statistical modelling is challenging since it is necessary to use a more structured model:
  - joint distribution of the intermediate variables
  - functional relationships between potential outcomes and potential values of the intermediate variables

Jin H. and Rubin D.B. (2008), Principal Stratification for Causal Inference With Extended Partial Compliance. *JASA*, 103, 101–111.

Bartolucci F. and Grilli L. (2011) Modelling partial compliance through copulas in a principal stratification framework. *JASA*, 106, 469-479.

Schwartz S.L., Li F., Mealli F. (2011) A Bayesian Semiparametric Approach to Intermediate Variables in Causal Inference. *JASA*, 106, 1331-1344.

Applications of principal stratification /1

NON-COMPLIANCE (estimation of CACE: Complier Average Causal Effect)

- Angrist J., Imbens G.W., Rubin D.B. (1996) Identification of Causal Effects using Instrumental Variables. *JASA,* 91, 444-472.
- Imbens, G.W. and Rubin, D.B. (1997) Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics,* 25, 305–327.
- Hirano K., Imbens G.W., Rubin D.B., Zhou X. (2000) Assessing the effect of an Influenza Vaccine in an Encouragement Design, *Biostatistics,* 1, 69-88.
- Mealli F., and D.B. Rubin (2002) Assumptions when Analyzing Randomized Experiments with Noncompliance and Missing Outcomes. *Health Services and Outcomes Research Methodology,* 3, 225-232.
- Jo B., Asparouhov T. & Muthen B. (2008) Intention-to-treat analysis in cluster randomized trials with noncompliance. *Statistics in Medicine*, 27, 5565–5577.
- Little, R., Long, Q., & Lin, X. (2009) A comparison of methods for estimating the causal effect of a treatment in randomized clinical trials subject to noncompliance. *Biometrics*, 65, 640–649.

Applications of principal stratification /2

ESTIMATION OF DIRECT AND INDIRECT EFFECTS

- Rubin D. B. (2004) Direct and Indirect Causal Effects via Potential Outcomes, *Scand. J. Stat.* 31, pp. 161–170.
- VanderWeele T. (2008) Simple relations between principal stratification and direct and indirect effects, *Stat. Probabil. Lett.* 78, pp. 2957-2962.
- Sjölander A., Humphreys K., Vansteelandt S., Bellocco R., Palmgren J. (2009) Sensitivity Analysis for Principal Stratum Direct Effects, with an Application to a Study of Physical Activity and Coronary Heart Disease. *Biometrics* 65, 514–520.
- Gallop R., Small D. S., Lin J. Y., Elliott M. R., Joffe M., Ten Have T. R. (2009) Mediation analysis with principal stratification. *Statist. Med.*, 28: 1108–1130.
- Mattei A. and Mealli F. (2011) Augmented Designs to Assess Principal Strata Causal Effects. *JRSS B*, 73, 729–752.
- Schwartz S.L., Li F., Mealli F. (2011) A Bayesian Semiparametric Approach to Intermediate Variables in Causal Inference. *JASA*, 106, 1331-1344.

Applications of principal stratification /3

SURROGATES

- Frankgakis C.E. & Rubin D.B. (2002) Principal stratification in causal inference, *Biometrics*, 58: 21-29.
- Gilbert P. B. and Hudgens M. G. (2008) Evaluating Candidate Principal Surrogate Endpoints. *Biometrics*, 64: 1146–1154.
- Li Y., Taylor J. M. G. and Elliott M. R. (2010) A Bayesian Approach to Surrogacy Assessment Using Principal Stratification in Clinical Trials, *Biometrics*, 66, 523–531.

MISSING DATA

- Taylor L. and Zhou X. H. (2009) Multiple Imputation Methods for Treatment Noncompliance and Nonresponse in Randomized Clinical Trials. *Biometrics*, 65: 88–95.
- Jin H., Barnard J., Rubin D.B. (2010) A Modified General Location Model for Noncompliance With Missing Data: Revisiting the New York City School Choice Scholarship Program Using Principal Stratification. *JEBS*, 35, 154–173.

TRUNCATION BY DEATH → Next

Truncation by death

Z=treatment, S=survival, Y=quality of life

→ Y defined only for S=1 (no quality of life for dead persons!)

BUT: non-sense to compare Y under Z=0 and Z=1 <u>among the survivors (i.e. condition on S=1)</u>:

Z=0 and S=1 ⇔ unit ∈ strata 11 or 01

Z=1 and S=1 ⇔ unit ∈ strata 11 or 10

outcome not defined under one value of Z → causal effect undefined

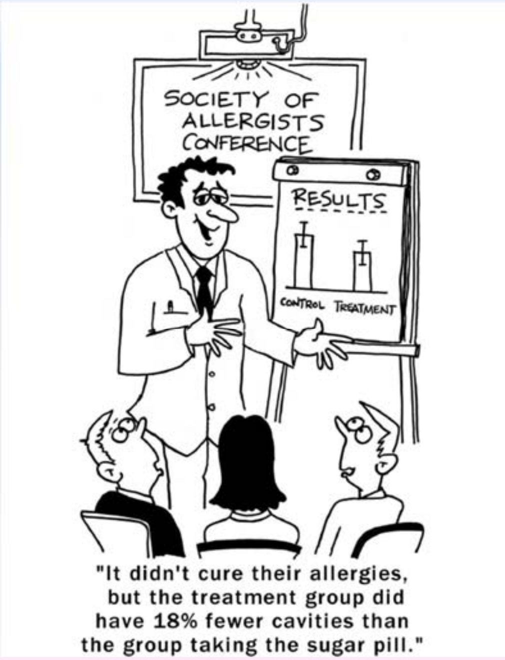
The only conceivable casual effect of Z on Y is the principal effect in the stratum 11, namely {S(0)=1, S(1)=1}

Truncation by death – examples /1

- Evaluating the causal effects of a *special educational intervention* on final test scores
  - ➢ S(z) = Graduation indicator given assignment z
  - Zhang JL, Rubin DB (2003). Estimation of causal effects via principal stratification when some outcomes are truncated by 'death', *JEBS*, 28, 353-368.
- Evaluating the causal *effect of Breast Self-Examination (BSE) teaching courses* on quality of execution of BSE
  - ➢ S(z) = Indicator of BSE practice given assignment z
  - Mattei A, Mealli F (2007) Application of the principal stratification approach to the Faenza randomized experiment on breast self-examination. *Biometrics* 63, 437-446.

Truncation by death – examples /2

- Evaluating the causal effects of *job training programs* on wages
  - ➢ S(z) = Indicator of employment given assignment z
  - Zhang JL, Rubin DB, Mealli F (2009) Likelihood-based analysis of causal effects of job-training programs using principal stratification. *JASA*, 104, 166-176.
  - Frumento P., Mealli F., Pacini B., Rubin D.B. (2012) Evaluating the effect of training on wages in the presence of noncompliance and missing outcome data. To appear in *JASA*. Draft on www.eale.nl/Conference2010/Programme/PaperscontributedsessionsF/add127617_WGTwGDCbRS.pdf
- Evaluating the *effectiveness of degree programs* on employment status of their graduates
  - ➢ S(z) = Graduation indicator given assignment z
  - Grilli L, Mealli F (2008) Nonparametric Bounds on the Causal Effect of University Studies on Job Opportunities Using Principal Stratification. *JEBS*, 33, 111-130.



www.causeweb.org

Case study

Relative effectiveness of two degree programmes with respect to employment

A parametric model
based on *principal stratification*
to deal with *truncation by death*

Scope and motivation /1

AIM: assessing the relative effectiveness of two degree programmes with respect to employment

- 1992 cohort of freshmen of the University of Florence
- Two degree programmes: Economics and Political Science
- Employment: binary indicator for having a permanent job about two years after degree

Scope and motivation /2

Naif approach: compare the employment rates for the graduates

But this is not fair, because the two degree programmes might "select" the individuals in a different way (e.g. one d.p. might be more easy in general or for students with certain features)

(issue is relevant: in our data the graduation rate after 8 years is around 25%)

If the graduates of the two d.p. differ for some unobserved features which are related with the occupational chances then a comparison based only on graduates yields biased results ⇒ need to take into account the graduation process

We exploit the idea of **principal stratification**, since there is a relevant intermediate variable (graduation) between the treatment variable (chosen degree prog.) and the outcome variable (employment)

Data

A. **Administrative database** of the 1992 cohort of freshmen enrolled in *Economics* (1068 students) and *Political Science* (873 students)

B1-B3. Three **census surveys** on the occupational status of the graduates of the University of Florence of years 1998 to 2000

datasets A and B1-B3 are merged

Available covariates: Female, Residence in Florence, Gymnasium (Lyceum), High grade, Late enrolment

covariates are important since the treatment is not randomised!

Treatment variable

Treatment variable Z:

$$Z = \begin{cases} 1 & \text{if enrolled in Economics} \\ 0 & \text{if enrolled in Political Science} \end{cases}$$

- Z is called “treatment” just to conform to the literature on causal inference
- No active vs. placebo → values of Z on an equal footing
- No randomisation → possible confounders (so covariates are important for unconfoundedness)

Intermediate and outcome variables

Intermediate variable S:

$$S = S(z) = \begin{cases} 1 & \text{if graduated when } z \\ 0 & \text{if not graduated when } z \end{cases}$$

S is the observed version of the potential variables S(0), S(1)

Outcome variable Y:

$$Y = Y(z) = \begin{cases} 1 & \text{if employed (after graduation) when } z \\ 0 & \text{if not employed (after graduation) when } z \end{cases}$$

Y is the observed version of the potential outcomes Y(0), Y(1)

For our purposes Y is defined only when S=1 (*truncation by death*)

Principal strata

In our case both Z and S are binary → 4 strata

Z	L=GG	L=GN	L=NG	L=NN
1 (Economics)	G	G	N	N
0 (Political Sc)	G	N	G	N

G=Graduated

N=Not graduated

Principal strata are defined by values of the two potential versions of the intermediate var. S (counterfactual): e.g. **GN** are the students who become **G**raduate if enrolled in Economics and **N**ot graduate if enrolled in Political Sc.

Observed group $O(Z, S^{obs})$	Z_i	S_i^{obs}	Y_i^{obs}	Latent group L_i (principal stratum)
$O(1,1)$	1	1	in {0,1}	GG or GN
$O(1,0)$	1	0	not defined	NG or NN
$O(0,1)$	0	1	in {0,1}	GG or NG
$O(0,0)$	0	0	not defined	GN or NN

mixtures

Sample proportions

Observed group $O(Z, S^{obs})$	Z_i	S_i^{obs}	Y_i^{obs}	Latent group L_i (principal stratum)
$O(1,1)$	1	1	in {0,1}	GG or GN
$O(1,0)$	1	0	not defined	NG or NN
$O(0,1)$	0	1	in {0,1}	GG or NG
$O(0,0)$	0	0	not defined	GN or NN

0.253 sample proportion of graduates (S_i^{obs}=1) among students in Economics (Z_i=1)

0.202 sample proportion of graduates (S_i^{obs}=1) among students in Political Science (Z_i=0)

SAMPLE SIZES	Economics	Political Sciences
Enrolled	1068	873
Inter-viewed	187	99

0.516 sample proportion of individuals with a permanent job (Y_i^{obs}=1) among students in Economics (Z_i=1) who graduated (S_i^{obs}=1)

0.364 sample proportion of individuals with a permanent job (Y_i^{obs} = 1) among students in Political Science (Z_i =0) who graduated (S_i^{obs}=1)

Relevant parameters

Probabilities of the principal strata: $\pi_{GG}, \pi_{GN}, \pi_{NG}, \pi_{NN}$

e.g. probability to be a student who become **G**raduate if enrolled in Economics and **N**ot graduate if enrolled in Political Science

Probabilities of employment: $\gamma_{1,GG}, \gamma_{0,GG}, \gamma_{1,GN}, \gamma_{0,NG}$

e.g. probability to be employed for a student who (i) become **G**raduate if enrolled in Economics and **N**ot graduate if enrolled in Political Science and (ii) actually enrolled in Economics

Causal effect of degree prog. on employment in the GG group: $\gamma_{1,GG} - \gamma_{0,GG}$

Type of analysis

Principal stratification is the conceptual framework for the application of various statistical methods:

- **Non parametric methods** (⇒ bounds)

  Grilli L. & Mealli F. (2008) Nonparametric Bounds on the Causal Effect of University Studies on Job Opportunities Using Principal Stratification. JEBS, 33, pp 111-130.

- **Model-based** (ML or Bayesian) **methods** (⇒ point estimates)

  Grilli L. & Mealli F. (2007) University Studies and Employment. An Application of the Principal Strata Approach to Causal Analysis. In Effectiveness of University Education in Italy: Employability, Competences, Human Capital (L. Fabbris ed.), pp 219-232. Heidelberg: Physica-Verlag.

I will show the model-based ML approach

Likelihood

$$L\left(\boldsymbol{\theta} \mid \mathbf{Z}, \mathbf{S}^{obs}, \mathbf{Y}^{obs}, \mathbf{X}\right) =$$

Enrolled in Economics:

$$\prod_{i\in O(1,1)} \left\{ \pi_{GG:i}\left[\left(\gamma_{1,GG:i}\right)^{Y_i^{obs}}\left(1-\gamma_{1,GG:i}\right)^{1-Y_i^{obs}}\right] + \pi_{GN:i}\left[\left(\gamma_{1,GN:i}\right)^{Y_i^{obs}}\left(1-\gamma_{1,GN:i}\right)^{1-Y_i^{obs}}\right] \right\}$$

$$\times \prod_{i\in O(1,0)} \left\{\pi_{NG:i} + \pi_{NN:i}\right\}$$

Enrolled in Political Sc.:

$$\times \prod_{i\in O(0,1)} \left\{ \pi_{GG:i}\left[\left(\gamma_{0,GG:i}\right)^{Y_i^{obs}}\left(1-\gamma_{0,GG:i}\right)^{1-Y_i^{obs}}\right] + \pi_{NG:i}\left[\left(\gamma_{0,NG:i}\right)^{Y_i^{obs}}\left(1-\gamma_{0,NG:i}\right)^{1-Y_i^{obs}}\right] \right\}$$

$$\times \prod_{i\in O(0,0)} \left\{\pi_{GN:i} + \pi_{NN:i}\right\}$$

Various models can be built by specifying submodels for the π's and the γ's

Model specification

Probabilities of the principal strata: π_{GG}, π_{GN}, π_{NG}, π_{NN}

Principal strata submodel: multinomial logit

Probabilities of employment: $\gamma_{1,GG}$, $\gamma_{0,GG}$, $\gamma_{1,GN}$, $\gamma_{0,NG}$

Outcome submodel: 4 distinct logit models

Principal strata are latent classes

$\Rightarrow$ the model is a *latent class model* with restrictions:

- a given individual can belong to only two of the four classes
- the outcome is not defined for some classes (depending on Z)

Principal strata submodel

$$\pi_{GG:i} = \frac{\exp(\eta^{\pi}_{GG:i})}{1+\exp(\eta^{\pi}_{GG:i})+\exp(\eta^{\pi}_{GN:i})+\exp(\eta^{\pi}_{NG:i})}$$

$$\pi_{GN:i} = \frac{\exp(\eta^{\pi}_{GN:i})}{1+\exp(\eta^{\pi}_{GG:i})+\exp(\eta^{\pi}_{GN:i})+\exp(\eta^{\pi}_{NG:i})}$$

$$\pi_{NG:i} = \frac{\exp(\eta^{\pi}_{NG:i})}{1+\exp(\eta^{\pi}_{GG:i})+\exp(\eta^{\pi}_{GN:i})+\exp(\eta^{\pi}_{NG:i})}$$

$$\pi_{NN:i} = \frac{1}{1+\exp(\eta^{\pi}_{GG:i})+\exp(\eta^{\pi}_{GN:i})+\exp(\eta^{\pi}_{NG:i})}$$

Multinomial logit specification

$$\eta^{\pi}_{GG:i} = \alpha^{\pi}_{GG} + \boldsymbol{\beta}^{\pi}_{GG}{}'\mathbf{x}_i$$

$$\eta^{\pi}_{GN:i} = \alpha^{\pi}_{GN} + \boldsymbol{\beta}^{\pi}_{GN}{}'\mathbf{x}_i$$

$$\eta^{\pi}_{NG:i} = \alpha^{\pi}_{NG} + \boldsymbol{\beta}^{\pi}_{NG}{}'\mathbf{x}_i$$

With 5 covariates there are $3+3\times5=18$ parameters

Outcome submodel

$$\gamma_{1,GG:i} = \frac{1}{1+\exp(-\eta^{\gamma}_{1,GG:i})}$$

$$\gamma_{0,GG:i} = \frac{1}{1+\exp(-\eta^{\gamma}_{0,GG:i})}$$

$$\gamma_{1,GN:i} = \frac{1}{1+\exp(-\eta^{\gamma}_{1,GN:i})}$$

$$\gamma_{0,NG:i} = \frac{1}{1+\exp(-\eta^{\gamma}_{0,NG:i})}$$

Logit link with stratum-specific linear predictor

$$\eta^{\gamma}_{1,GG:i} = \alpha^{\gamma}_{1,GG} + \boldsymbol{\beta}^{\gamma}{}'\mathbf{x}_i$$

$$\eta^{\gamma}_{0,GG:i} = \alpha^{\gamma}_{0,GG} + \boldsymbol{\beta}^{\gamma}{}'\mathbf{x}_i$$

$$\eta^{\gamma}_{1,GN:i} = \alpha^{\gamma}_{1,GN} + \boldsymbol{\beta}^{\gamma}{}'\mathbf{x}_i$$

$$\eta^{\gamma}_{0,NG:i} = \alpha^{\gamma}_{0,NG} + \boldsymbol{\beta}^{\gamma}{}'\mathbf{x}_i$$

With 5 covariates there are $4+5=9$ parameters

Principal Causal Effect in the GG stratum (on the logit scale):

$$\alpha^{\gamma}_{1,GG} - \alpha^{\gamma}_{0,GG}$$

ML inference

- ✓ Maximization algorithm: quasi-Newton with a BFGS update of the Cholesky factor of the approximate Hessian
- ✓ Software: SAS proc NLMIXED

- Principal strata submodel $\Rightarrow$ 18 parameters
- Outcome submodel $\Rightarrow$ 9 parameters

Overall 27 parameters

Some parameters of the Principal strata submodel (a multinomial logit) have highly negative estimates and huge standard errors

$\Rightarrow$ for certain values of the covariates some principal strata are empty so some constraints are needed (the final model has 6 constraints)

Main result: the estimated causal effect (on the logit scale) is 0.666 (s.e. 0.301)

Estimated probabilities (%) for some covariate patterns

Parameter	00000	00100	00110	00101	01100	10100	11100	11111
$\pi_{GG:i}$	1.1	8.0	9.1	10.9	20.3	24.9	52.5	62.2
$\pi_{GN:i}$	6.3	6.0	3.3	0.0	14.0	0.0	0.0	0.0
$\pi_{NG:i}$	3.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\pi_{NN:i}$	89.0	86.0	87.6	89.1	65.7	75.1	47.5	37.8
$\gamma_{1,GG:i}$	77.9	58.2	59.9	60.7	57.3	48.0	47.1	51.5
$\gamma_{0,GG:i}$	64.5	41.7	43.4	44.2	40.8	32.2	31.4	35.3
$\gamma_{1,GN:i}$	61.9	39.0	40.7	41.5	38.1	29.8	29.0	32.8
$\gamma_{0,NG:i}$	20.3	9.1	9.7	10.0	8.9	6.3	6.1	7.1
Causal effect $\gamma_{1,GG:i} - \gamma_{0,GG:i}$	13.5	16.5	16.5	16.4	16.5	15.8	15.7	16.2

The pattern $(x_1, x_2, x_3, x_4, x_5)$ stands for

Gymnasium $= x_1$, *High grade* $= x_2$, *Regular enrolment* $= x_3$, *Female* $= x_4$, *Florence* $= x_5$

Results: principal strata submodel

- the size of GG stratum varies a lot with the covariates, from a minimum of 1.1% (students with weak background) to a maximum of 62.2%
- for most covariate patterns the GN and NG strata (i.e. students able to graduate in only one degree prog.) are very small (but for students with weak background they are larger then the GG stratum)
- the higher graduation rate of Economics is originated by the students with a weak background $\Rightarrow$ orientation policies should be designed especially for this kind of students

Results: outcome submodel

- the level of the probability of being employed varies a lot with the covariates
- in the GG stratum the causal effect on employment (modelled as constant across the covariate patterns) is about 15% (significant at 5%)
- students with a weak background have little chances of being GG, so for them the above causal effect has little relevance

Gracias por su atención!

www.ds.unifi.it/grilli
grilli@ds.unifi.it