Modelling partial compliance through copulas in a principal stratification framework

Bartolucci F. and Grilli L. (2011) To appear in JASA.

Leonardo Grilli

University of Florence

Some history

- **EFRON & FELDMAN (1991):** analysis of a randomized trial with partial non-compliance

- **FRANGAKIS & RUBIN (2002):** principal stratification
  - general framework to deal with non-compliance
  - but earliest applications are for all-or-none compliance ($\rightarrow$ discrete strata)

- **JIN & RUBIN (2008):** new analysis of Efron & Feldman data using *continuous principal strata*

- **BARTOLUCCI & GRILLI (2011):** new analysis of Efron & Feldman data following the approach of Jin & Rubin but *with a different modeling strategy*

Efron-Feldman data /1

- **EF**: Efron B. & Feldman D. (1991) Compliance as an Explanatory Variable in Clinical Trials, *JASA* 86, 9-17.

- Subset of data from LRC-CPPT, a *placebo-controlled double-blinded randomized clinical trial* designed to study the effectiveness of *cholestyramine* for lowering cholesterol levels

- data on 335 men: 164 → active pills of the drug
   171 → placebo pills

- Visits at two-month intervals (average length of 7.3 years)

- Variables:
  - binary indicator for treatment assignment (1 = active pills)
  - compliance (proportion of assigned packets not returned, averaged over all visits – rounded to the second decimal)
  - continuous outcome variable: average decrease in the cholesterol level during the study
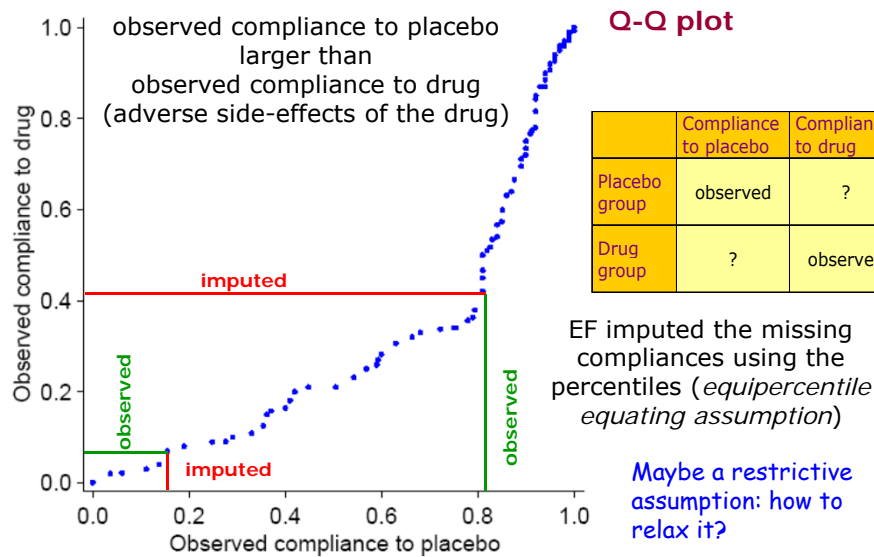
3

Efron-Feldman data /2
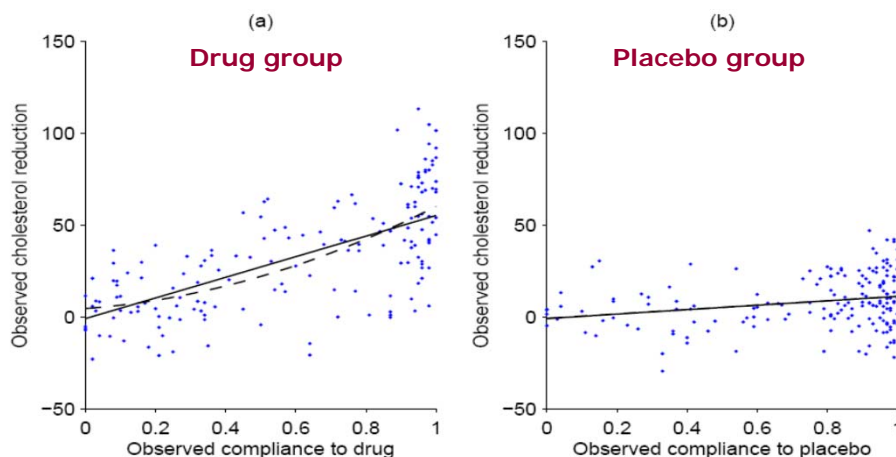
The effect of the observed compliance to drug is likely the combination of a ***"genuine" effect*** (chemical action of the drug) and a ***"collateral" effect*** (correlation between the degree of compliance and some unobserved characteristics of the patients which affect the cholesterol level, such as the propensity to eat healthy food or to exercise)

4

Efron-Feldman data /3

observed compliance to placebo larger than observed compliance to drug (adverse side-effects of the drug)

Q-Q plot

	Compliance to placebo	Compliance to drug
Placebo group	observed	?
Drug group	?	observed

EF imputed the missing compliances using the percentiles (*equipercentile equating assumption*)

Maybe a restrictive assumption: how to relax it?

5

Preliminary analysis

□ Treatment indicator Z_i (1=drug, 0=placebo)

□ Potential outcomes
- Compliance: d_i placebo, D_i drug
- Outcome: $Y_i^{(0)}$ placebo, $Y_i^{(1)}$ drug

□ Preliminary analysis: separate regressions for the two treatment arms

$$Y_i^{(0)}|d_i, Z_i = 0$$
$$\sim \mathrm{N}[-0.869 + 12.081 d_i, \exp(5.294)],$$
$$Y_i^{(1)}|D_i, Z_i = 1$$
$$\sim \mathrm{N}[-0.869 + 56.106 D_i, \exp(5.294 + 1.366 D_i)]$$

SELECTED MODELS: no quadratic terms, but hetero-schedasticity for $Z_i=1$

6

3

Modelling strategy

□ Principal stratification approach of **JR**: Jin and Rubin (2008) Principal Stratification for Causal Inference With Extended Partial Compliance, *JASA* 103, 101-111.

□ Principal stratification

  ▪ (d_i, D_i) principal strata (***continuous***)

  ▪ $E(Y_i^{(1)} - Y_i^{(0)} \mid d_i, D_i)$ principal causal effect (PCE)

□ Regression models for the outcomes

  ▪ $Y_i^{(0)}$ on d_i and D_i

  ▪ $Y_i^{(1)}$ on d_i and D_i

> • Like JR we adopt principal stratification and focus on PCE
>
> • In the regressions we allow for nonlinearities (like JR) but also for interactions and heteroschedasticity
>
> • Key point: we relax one of the assumptions of JR

7

Assumptions /1

1. SUTVA

2. Ignorable treatment assignment

3. Strong access monotonicity (drug compliance is null for patients assigned to placebo and viceversa)

> • We rely on assumptions 1 to 3 like JR and EF
>
> • These assumptions are untestable but reasonable in carefully designed randomized trials such as the *cholestyramine* study

8

Assumptions /2

- An advantage of our approach is that we do not invoke assumptions about the relationship between drug and placebo compliances

- In contrast
  - EF assumed *equipercentile equating* of the compliances (a deterministic relationship)
  - JR assumed *negative side-effect monotonicity*, i.e. the drug compliance is no larger than the placebo compliance (a stochastic relationship with $D_i \leq d_i$)

  > - Negative side-effect monotonicity seems plausible in placebo-controlled experiments where the drug has some negative side effects
  > - However, it is violated if there exists a subset of patients having positive side effects from reported cholesterol reductions after periodic blood tests

Joint distribution of compliances

- The compliances d_i and D_i are never jointly observed …

- … but empirical evidence on their correlation is induced by the equations for the outcomes where both compliances enter as regressors

- For the joint distribution of the compliances, JR assumed that d_i has distribution $Beta(\alpha_1, \alpha_2)$ and (conditional on d_i) the ratio D_i/d_i has distribution $Beta(\alpha_3, \alpha_4)$, which implies $D_i \leq d_i$ (*Negative side-effect monotonicity*)

  > - The specification of JR allows the compliances to be correlated, although in a particular way that should not be viewed as the only possible way
  > - *How sensitive is the inference on the causal effect to the model on the compliances?*

Copulas

- To model the joint distribution of the compliances (d_i,D_i) we use a copula, which is a flexible way to define a joint distribution from the marginals

- Definition: let X and Y be two random variables with distribution functions $F_X(x)$ and $F_Y(y)$, then a single-parameter **copula** (with parameter ψ) is a function $C_\psi(\cdot\ ,\ \cdot)$ such that $C_\psi(F_X(x),F_Y(y))$ is a joint distribution function

- The copula has the merit of allowing us to study the association between the compliances <u>without specifying a model for their marginal distributions</u>, which are estimated by their *empirical distribution functions*

11

Plackett copula

- We use a Plackett copula which has a single association parameter ψ

  - $0 < \psi < 1$ → negative association
  - $\psi = 1$ → independence
  - $\psi > 1$ → positive association

- Advantages of using a copula instead of a parametric density:

  - no constraints on the marginal distributions
  - association captured by a single parameter (to be estimated or used in a sensitivity analysis)

12

6

ML estimation via EM

1. Compute the univariate empirical distribution functions of d_i and D_i

2. For a set of values of the association param. ψ

    i. Compute the joint distribution function of (d_i, D_i) using the copula

    ii. Compute the likelihood

    iii. Maximize the likelihood via EM

    we wrote a MATLAB routine

3. Plot the profile likelihood for ψ

This allows us to
- see how the different values of ψ are supported by the data
- check for local maxima

Model selection /1

❑ We begin with a general form with quadratic terms, interactions and heteroschedasticity

INITIAL SPECIFICATION:

1. *Regression models for the means*
    - $E(Y_i^{(0)} \mid d_i, D_i) = \ldots$
    - $E(Y_i^{(1)} \mid d_i, D_i) = \ldots$

    9 parameters

2. *Regression models for the log-variances*
    - $\text{logVar}(Y_i^{(0)} \mid d_i, D_i) = \ldots$
    - $\text{logVar}(Y_i^{(1)} \mid d_i, D_i) = \ldots$

    6 parameters

3. *Plackett copula for the joint distribution of* (d_i, D_i) with association parameter ψ

    1 parameter

We test several restrictions using the LR test

Model selection /2

FINAL SPECIFICATION:

1. *Regression models for the means*

   - $E(Y_i^{(0)} \mid d_i, D_i) = \alpha + \beta \, d_i$

   - $E(Y_i^{(1)} \mid d_i, D_i) = \alpha + \beta \, d_i + \gamma \, D_i + \delta \, (d_i \times D_i)$

   } 4 parameters

2. *Regression models for the log-variances*

   - $logVar(Y_i^{(0)} \mid d_i, D_i) = \phi$

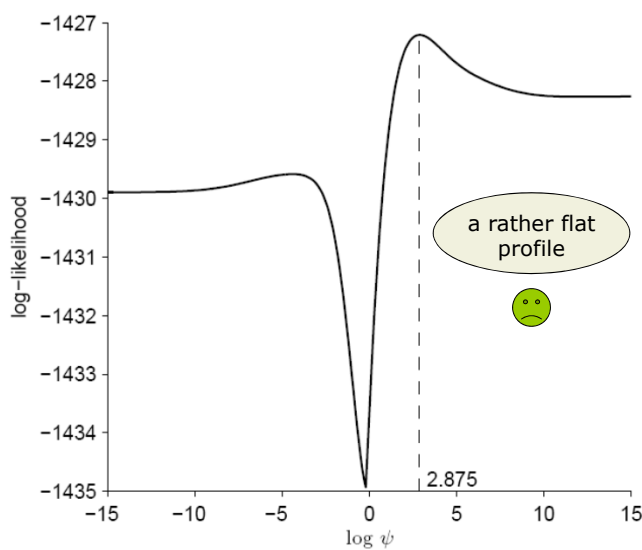   - $logVar(Y_i^{(1)} \mid d_i, D_i) = \phi + \eta \, D_i$

   } 2 parameters

3. *Plackett copula for the joint distribution of* (d_i, D_i) *with association parameter* ψ

   } 1 parameter

Profile log-likelihood of $\log \psi$

a rather flat profile

Point estimate of ψ is 17.727

Independence between d_i and D_i (i.e. $\psi = 1$) is rejected (p-value<0.001)

Pearson correlation between d_i and D_i is 0.689

Type of inference

- The PCE depends on the regression coefficients, which depend on the value of the Plackett association parameter ψ which is
  - estimated with low precision because of scarce empirical support (flat profile log-likelihood)
  - identified thanks to the regression equations (which cannot be tested separately since they depend on both compliances)
- Thus it is not advisable to base the inference exclusively on the model with ψ at its point estimate
- First, I will illustrate the inference drawn with ψ at its ML estimate
- Next, I will show the results of a sensitivity analysis on the estimated PCE by letting ψ vary on a suitable interval
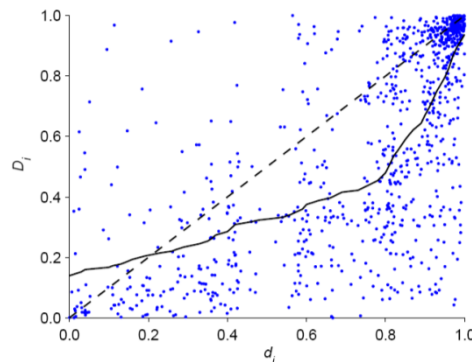
17

Association of compliances [ψ at MLE]

Random draws from the bivariate distribution of the compliances
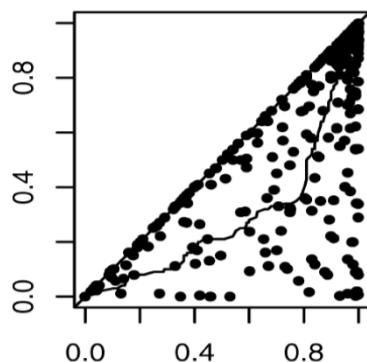
Jin and Rubin (2008)   Bartolucci and Grilli (2011)

We relax the *negative side-effect monotonicity* (i.e. $D_i \leq d_i$)

→ 21.6% of the points go beyond the bisectrix, corresponding to individuals with $D_i > d_i$ (they would take more drug than placebo)

18

Models for the means and PCE [ψ at MLE]

□ $E(Y_i^{(0)}|d_i, D_i) = -0.269 + 11.243\,d_i$

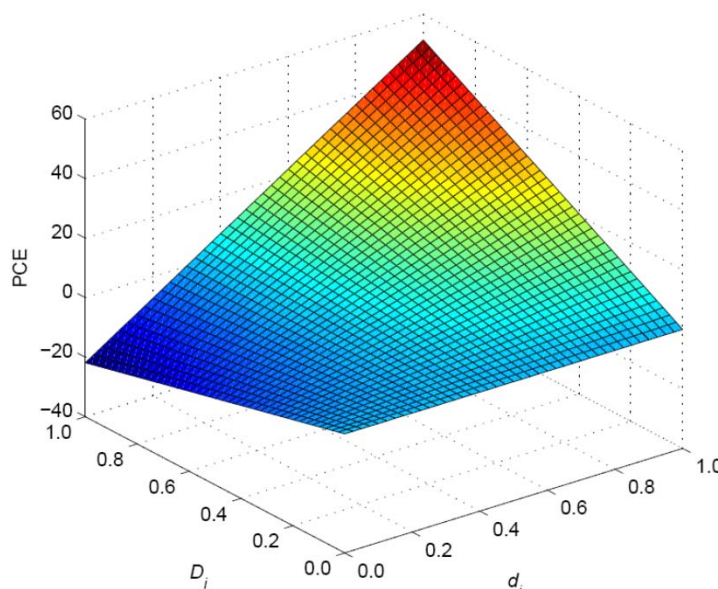□ $E(Y_i^{(1)}|d_i, D_i) = -0.269 + 11.243\,d_i - 21.878\,D_i + 73.359\,(d_i \times D_i)$

The estimated **Principal Causal Effect** is

$$PCE(d_i, D_i) = (-21.878 + 73.359\,d_i) \times D_i$$

□ The PCE depends on the dose of the drug D_i and the slope is

- positive, except when $d_i < 0.298$ (but this is rare: 12.3% of the subjects in the placebo arm)
- steeper at higher levels of the placebo compliance d_i

19

Estimated PCE surface [ψ at MLE]

Good agreement with the estimates of Jin and Rubin

At the median point (d_i=0.89, D_i=0.70) our ML estimate of PCE is 30, compared with 24 of JR (Bayesian posterior median)

20

Confidence intervals for PCE [ψ at MLE]

◻ Confidence intervals for the PCE are easily obtained via the *nonparametric bootstrap* (1000 samples)

◻ There is a confidence interval for each couple of values of the compliances, for example:

- at the medians (d_i=0.89, D_i=0.70) → interval (22.5, 39.2) → significant effect

- at first quartiles (d_i=0.59, D_i=0.27) → interval (-3.1, 9.9) → non-significant effect

- at first quartile for placebo compliance and at third quartile for drug compliance (d_i=0.59, D_i=0.95) → interval (-10.8, 34.7) → non-significant effect

Remark: the third case refers to an unlikely couple of compliance levels (rarely D_i is much larger than d_i) → very large interval

Sensitivity analysis

◻ We perform a **sensitivity analysis** to assess how the PCE depends on ψ (we let it vary in its profile likelihood interval):

- at the median point (d_i=0.89, D_i=0.70): PCE $\in$(27.4, 34.8)

- at the Q1-Q3 point (d_i=0.59, D_i=0.95): PCE $\in$(14.0, 29.5)

Remark: this case (corresponding to an unlikely couple of compliance levels, i.e. far from the bulk of the data) has an high sensitivity (in addition to a large sampling variance – recall the bootstrap CI)

◻ Principal Causal Effects are reliably estimated at drug and placebo compliance levels near the sample medians, while inference at unlikely compliance levels appears to be unduly affected by model assumptions

Model checks

- The estimates of the PCE may critically depend on some modelling choices such as
    - the **normality** of the conditional distributions of the potential outcomes
    - the **type of copula**
- We check the normality assumption using a Box-Cox transformation of the outcomes
    - the LR test does not reject the hypothesis of an identity transformation (i.e. no evidence against normality)
    - the estimates of the PCE are quite stable (except for unlikely couples of compliance levels – again!)
- We check the role of the copula by replacing the Plackett copula with a Gaussian copula:
    - the estimates of the PCE are very stable (even for unlikely couples of compliance levels)

23

Warnings

- <u>Be aware of hidden extrapolation</u>: the model yields estimates of the PCE at any couple of compliance levels, but for unlikely couples (i.e. far from the bulk of the data) the estimates are highly sensitive to model assumptions (beyond having a large sampling variance)

- <u>We are not estimating a dose-response function</u>: for a fixed value of d_i, we can draw PCE(d_i,D_i) as a function of D_i: however, this is not interpretable as a dose-response function since the dose of the drug is not randomly assigned but chosen by the patients → the effect of the dose of the drug is mixed with the effect of the unobserved features associated with the degree of compliance (the interpretation in terms of a dose–response function would require further problematic assumptions – see JR's section 4)

24

Final remarks

- Our ML estimates of the causal effects for the Efron-Feldman data are similar to the Bayesian estimates of Jin and Rubin (2008) …

- … but we offer an *alternative modelling strategy* yielding a different interpretation due to
  - *interaction between the compliances* in the principal causal effect
  - *flexible specification of the joint distribution of the compliances through a copula* (→ doubts on the negative side-effect assumption)

- Merits of our approach:
  - The joint distribution of the compliances is modelled in a flexible way (relaxing assumptions)
  - Sensitivity analysis is straightforward
  - ML estimation via EM is computationally simple

25

Thanks for your attention!

grilli@ds.unifi.it