

Estimands

Peter J Diggle

Lancaster University



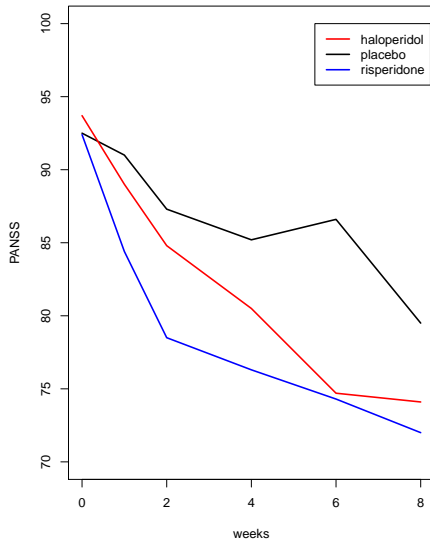
Lancaster
Medical School

Lancaster
University



- 1 a randomised trial of drug treatments for schizophrenia
- 2 what's the question?
- 3 a model-based answer
- 4 estimation or prediction?
- 5 marginal or conditional effects?

PANSS data



- **what benefit could a patient expect to receive from either treatment?**
- **what treatment should we recommend for general use?**
- **...?**

Measurement sub-model

$$Y_{ij} = \{\alpha_{k(i)} + \beta_{k(i)}t_{ij} + \gamma_{k(i)}t_{ij}^2\} + S_i(t_{ij}) + Z_{ij}$$

Dropout sub-model

$$\log[P_{ij}/(1 - P_{ij})] = \tau + \phi Y_{i,j-1}$$

What's the difference between treatments 2 and 1 with respect to the change in symptom severity over the 8-week follow-up period?

Estimation

An **estimand** is a function of the model parameters

Example:

$$(\beta_2 - \beta_1) \times 8 + (\gamma_2 - \gamma_1) \times 64$$

Prediction

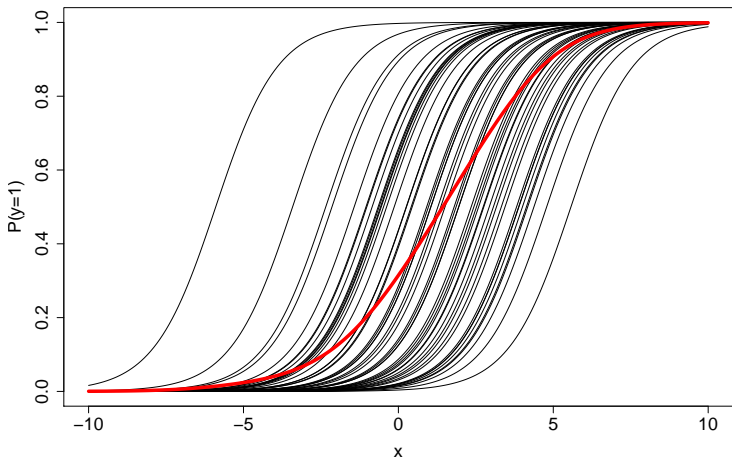
A **predictand** is a function of parameters and unobserved random variables

Example:

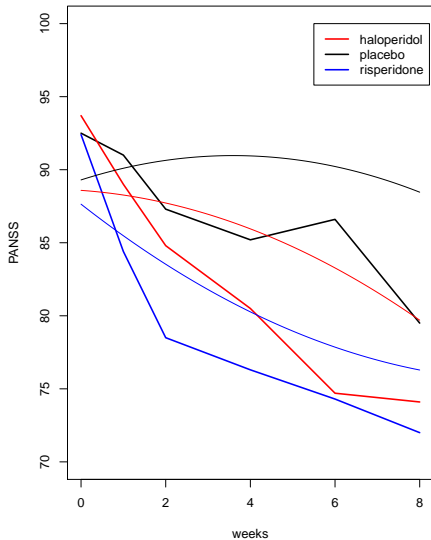
$$(\beta_2 - \beta_1) \times 8 + (\gamma_2 - \gamma_1) \times 64 + \{S^{(2)}(8) - S^{(2)}(0)\} - \{S^{(1)}(8) - S^{(1)}(0)\}$$

Marginal or conditional effects?

Logistic regression example



The PANSS data re-visited



- **I don't understand what LOCF is estimating**
- **I find it hard to define the precise question of interest without using mathematics**
- **Writing down a model makes two things explicit:**
 - **the relevant estimand or predictand**
 - **the underlying assumptions**
- **The combined effects of non-linearity, missing values and serial correlation are subtle ... and can be counter-intuitive**