Multiple imputation by substantive model compatible methods:

Case study of 90-day colon cancer mortality

Helen Fowler and Edmund Njeru Njagi

Aim:

Investigate the roles of various prognostic factors in socio-economic differences in 90-day colon cancer mortality

Analysis model:

Logistic regression model

- Outcome: 90-day mortality, main exposure: deprivation group

- Explanatory variables: stage, age (non-linear), comorbidity score, receipt of surgical treatment

- Interactions:
    - Stage and i) age, ii) treatment, iii) deprivation and iv) comorbidity score
    - Deprivation and i) comorbidity and ii) treatment
    - Comorbidity and treatment

Missing data:

- Stage (30%)
- Grade (20%)

MI by full conditional specification (FCS):

- Joint distribution approximated – all conditionals

- Model* for stage given all other variables

- Model* for grade given all other variables

- *Imputation model

- Analysis model page 2 = substantive model

Incompatibility in FCS:

- No joint model for which the conditionals equal the imputation and substantive models

- Problem if following not handled well:
  - Interactions
  - Non-linear effects

- Results in biased parameter estimates

Impute stage accounting for stage*age interaction:

Wrong move 1: Passive imputation

- Do nothing about the interaction at imputation

- Test for the interaction in substantive model

- Results in attenuated parameters

Impute stage accounting for stage*age interaction:

Wrong move 2: ``Just another variable'' approach

- Compute new variable $Z = \text{stage*age}$

- Impute Z as if different variable from stage and age

- Only valid under MCAR

- May even invalidate Rubin's combination rules

Substantive model compatible FCS (SMC – FCS)

- Values drawn from a candidate distribution

- Draw accepted if satisfy certain bound

- Bound constructed based on the substantive model

- Imputation compatible with substantive model

Availability in software

- *smcfcs* in STATA

- *smcfcs* in R

Read more about SMC – FCS

- Bartlett JW, Morris TP. Multiple imputation of covariates by substantive-model compatible fully conditional specification. The Stata Journal 2015; 15: 437-56.

- Bartlett JW, Seaman SR, White IR, Carpenter JR. Multiple imputation of covariates by fully conditional specification: Accommodating the substantive model. Stat Methods in Med Research 2015; 24: 462-87.