

Análise de Dados Longitudinais Medidas Repetidas e Dados Perdidos

Enrico A. Colosimo/UFGM

- 1 Medidas Repetidas / Modelos Multiníveis.
- 2 Dados Perdidos / Imputação

Medidas Repetidas / Modelos Multiníveis

- mesmo indivíduo/unidade amostral é medido várias vezes em diferentes condições.
- Modelo Multinível ou Hierárquico.
- Exemplo (três níveis): criança/moradia/bairro: resposta é a presença de asma. Covariáveis: idade da criança, presença de mofo em casa e nível de poluição da região.
- Dados longitudinais é um caso particular em dois níveis com níveis ordenados.
- Utiliza frequentemente modelos de efeitos aleatórios.

Medidas Repetidas: Trauma Dentário

- 165 pacientes perderam seus dentes em acidentes/traumas.
- Alguns pacientes perderam mais de um dente: Medidas Repetidas: 205 DENTES
- Os dentes são reimplantados no serviço de urgência do Hospital Odilon Behrens.
- Após algo tempo os pacientes são encaminhados para a Clínica de Trauma Dental da Escola de Odontologia da UFMG, (DTC-SD-FUMG) coordenada pela Profa. Juliana Vilela Bastos.
- O banco de dados refere-se ao período de 1994 a 2011. avaliando a presença, tipo e extensão da ERR.

Medidas Repetidas: Trauma Dentário

- Acontece um processo de absorção da raiz dentária que representa um problema após reimplantação do dente.
- Variáveis demográficas e clínicas estão associadas com a prevalência e extensão da absorção dentária.
- Este estudo tem como objetivo investigar a associação destas variáveis com a ocorrência e gravidade da absorção da raiz após um acidente traumático.
- Resposta: A reabsorção é medida através de um índice que varia de 0 a 12 and were grouped into two categories: low (≤ 4) and severe (≥ 4) que representa 75.1/

Dados Perdidos (Dados Longitudinais)

Dados Perdidos acontecem com frequencia em Estudos Longitudinais.

Tipos:

- 1 **"Drop out"**: Perda completa a partir de um certo instante;
- 2 **Intermitente**: Somente uma perda pontual.

Implicações:

- Perda do desbalanciamento;
- Perda de precisão (aumento da variância);
- "Possível" viés de seleção de acordo com o mecanismo de perda dos dados e o método de estimação.

1 "MCAR" Missing Completely at Random:

- Perda completamente aleatória;
- Dados observados podem ser considerados como uma amostra aleatória dos dados completos.

Exemplo: Mudança da cidade na qual estava sendo desenvolvido o estudo (migração)

Impacto:

- Qualquer método de análise gera inferências válidas.
- Perda de precisão.

2 "MAR": Missing at RAndom

- A probabilidade da resposta perdida depende do conjunto de valores observados.

Exemplo: Óbito devido a gravidade da resposta medida.

Impacto:

- As médias amostrais ficam viesadas.
- GEE não produzem estimativas válidas para β .
- EMV produz estimativas válidas se a distribuição normal for corretamente especificada para a resposta.

3 "NMAR": Not Missing at Random

- A probabilidade da resposta perdida depende do conjunto de valores observados e não observados.

Exemplo: gravidade devido a renda/salário.

Impacto:

- As médias amostrais ficam viesadas.
- GEE não produzem estimativas válidas para β .
- EMV não produz estimativas válidas.

1 Imputação Simples

- Substituir o dado perdido por uma única estimativa (média, último valor, etc).
- Subestima a variância dos estimadores.

2 Imputação Múltipla.

- Modelos de Regressão.
- Distância de valores observados (o mais próximo segundo alguma métrica).
- Análise Bayesiana:

$$p(y_{\text{perdido}} / \text{dados})$$

.