

Análise de Dados Longitudinais

Introdução à Disciplina

Enrico A. Colosimo-UFMG
www.est.ufmg.br/~enricoc

Medidas Repetidas/Dados Longitudinais

Medidas Repetidas

Medidas Repetidas são obtidas quando uma resposta é medida repetidamente em um grupo de unidades.

- Unidades: indivíduos, pacientes, animais, plantas, conglomerados (família, empresa, cidade, etc), etc.
- Caso especial: Dados Longitudinais: são medidas ao longo do tempo em uma mesma unidade/indivíduo.

Tipos de Estudos

- Transversal: uma única resposta é medida em cada unidade em um certo instante de tempo.
- Longitudinal: unidades são, geralmente, medidas repetidamente ao longo do tempo.

Obs.: Medidas repetidas podem estar associadas à estudos transversais.

Estudo Transversal vs Longitudinal

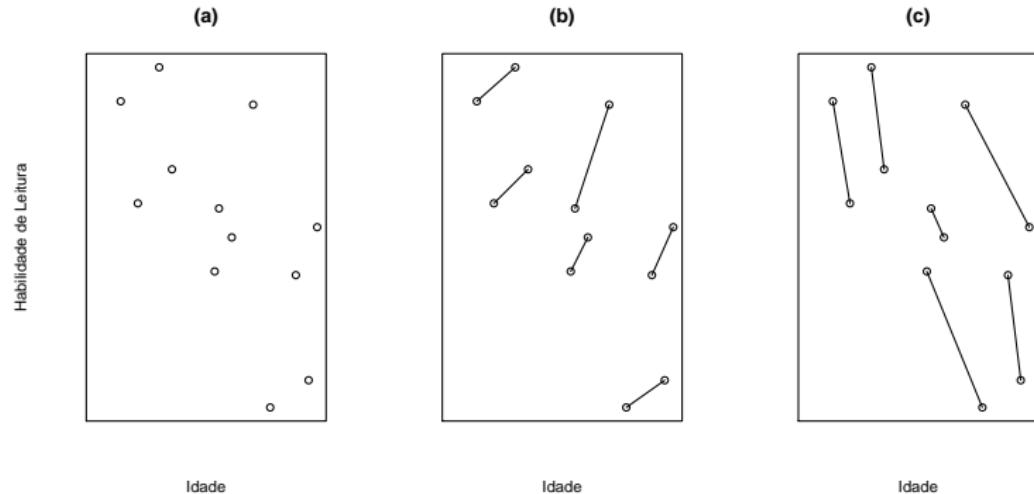


Figura: Habilidade de Leitura: transversal (a) vs longitudinal (b) e (c).
Fonte: Diggle et al. (2002).

Vantagens do Desenho Longitudinal

- Permite avaliar o comportamento da resposta ao longo do tempo.
- Permite avaliar a variação intra-indivíduo.
- No exemplo, o efeito de idade pode ser confundido com um possível efeito de coorte.
- Se quisermos comparar o depois com o antes, cada indivíduo atua como seu próprio controle evitando o efeito de variação entre indivíduo (fatores de confusão).

Medidas Repetidas/Dados Longitudinais

Característica Principal

Medidas Repetidas na mesma unidade são
CORRELACIONADAS.

Exemplos: Estudos Longitudinais

- ① Indivíduos foram divididos aleatoriamente em dois grupos (A e B) e a pressão sistólica foi medida em 5 tempos distintos.
(**Dados Longitudinais**)
- ② Indivíduos foram divididos aleatoriamente em dois grupos (A e B) e foi registrado o tempo até que a pressão sistólica atingiu um certo patamar.
(Análise de Sobrevivência)
- ③ Uma série histórica (200 valores) de medidas de pressão sistólica foi registrada para o Sr. João.
(Séries Temporais)

Análise de Dados Longitudinais - Medidas Repetidas

1 Características:

- Dados de mesma natureza;
- grande número de pequenas séries;
- os tempos de medição são fixos (**balanceado ou não balanceado**);
- covariáveis: fixas ou dependentes do tempo;
- áreas de aplicação: saúde, economia, engenharia, etc.

2 Vantagens:

- **avaliar mudança no tempo**;
- eficiência no custo das observações;
- homogeneidade nas comparações.

3 Dificuldades:

- **observações correlacionadas**;
- usualmente mais demorados;
- fonte de vício: dados perdidos (viés de seleção).

Exemplos Reais

Estudo "cross-over"

- O Dr. Emílio Suzuki quer comparar o efeito de dois colírios (A, B) redutores da pressão ocular com relação ao fluxo sanguíneo.
- Para tal ele submeteu cada paciente aos dois colírios por um período de dois meses com um descanso de igual tamanho.
- A ordem da aplicação dos colírios foi aleatória
- Duas medidas de pressão (colírio A e colírio B) foram tomadas ao fim do estudo para cada paciente.
- Eventualmente, a medida de linha de base pode ser útil na análise estatística.

Exemplos Reais

Estudo Longitudinal Desbalanceado: Avaliação longitudinal do crescimento de lactentes nascidos de mães infectadas com o HIV-1.

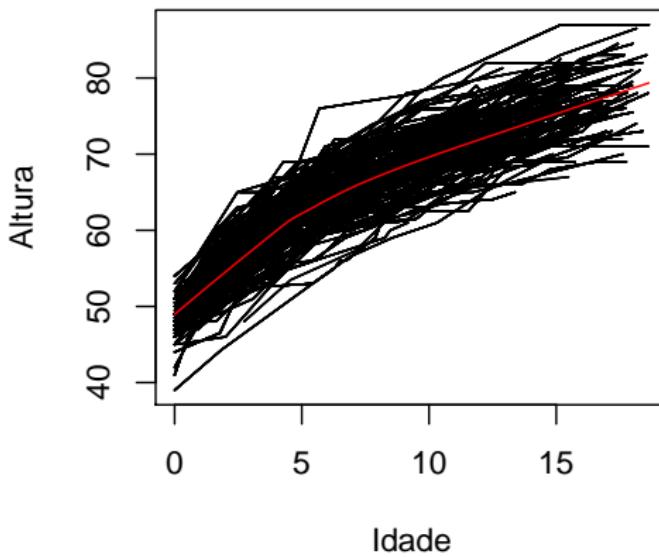
- Comparar longitudinalmente altura de lactentes infectados e não-infectados nascidos de mães infectadas pelo HIV.
- Uma coorte aberta acompanhada no ambulatório de AIDS pediátrica do Hospital das Clínicas da Universidade Federal de Minas Gerais.
- Período: 1995 a 2003.
- Inclusão: primeiros três meses de vida.
- Grupos: (1) não-infectados: 97; (2) infectados: 42.
- Controlado por sexo.

Estrutura Longitudinal

- Visitas regulares ao pediatra.
- Planejado para acompanhamento de 18 meses.
- Tempo: idade da criança.
- Tempo mediano de acompanhamento foi 15 meses (7 a 18).
- Número total de medidas: Não-infectados: 907; Infectados: 411.
- Número médio de visitas por criança: 9,5.
- Delineamento não-balanceado.

Perfis individuais e médio

Perfis das Crianças



Perfis médio por grupo

Gráfico para os Grupos

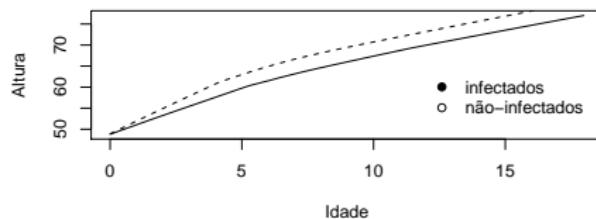
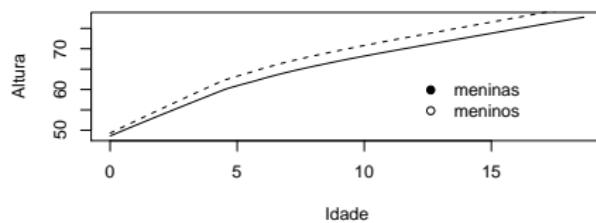


Gráfico para Meninos e Meninas



Exemplos Reais

Marcadores Psicofisiológicos de Proteção e Vulnerabilidade ao Estresse Psicosocial

Os objetivos gerais deste estudo são:

- Investigar as reações cardíacas a uma situação de estresse social.
- Investigar a capacidade de regulação dessas respostas em função da afetiva individual (fatores internos) e da indução prévia de um estado de afeto positivo ou negativo (fator externo).

Exemplos: Estresse Psicosocial

- Participaram do experimento 72 estudantes universitários da Universidade de Granada (Espanha) de ambos os sexos, com idade entre 18 a 30 anos.
- Foram utilizadas 40 fotos agradáveis (famílias e bebês) e 40 fotos desagradáveis (pessoas com mutilações) selecionadas do catálogo International Affective Picture System - IAPS.
- O objetivo das fotos é induzir um estado de humor positivo ou negativo, respectivamente.
- Resposta: período cardíaco médio avaliado em 12 momentos.

Perfis individuais e médio

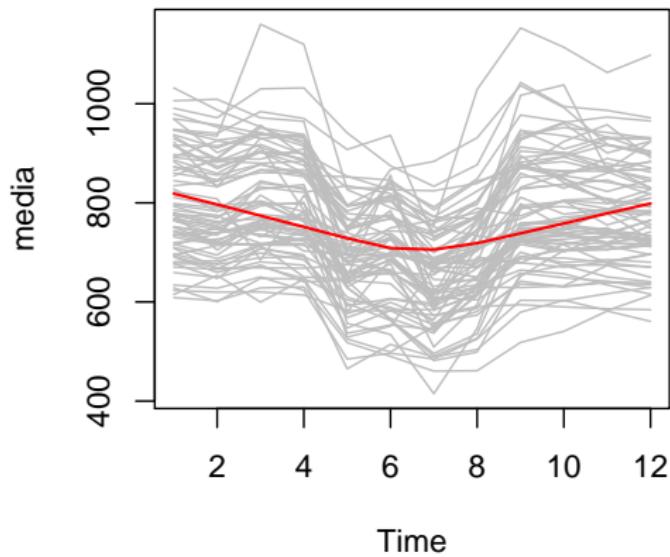


Figura: Perfis dos estudantes e uma curva alisada

Exemplo - Medidas Repetidas

- 1 Trauma Odontológico - Profa. Juliana Bastos
 - O indivíduo perde o(s) dente(s) por acidente.
 - O dente é reimplantado em um serviço de urgência (Odilon Behrens).
 - Em seguida ele é encaminhado ao serviço de trauma da Faculdade de Odontologia da UFMG para tratamento de canal.
 - No período entre o reimplante e o canal, existe um processo de reabsorção inflamatória.
 - Este processo de reabsorção é medido por um índice.
 - A resposta de interesse é a avaliação deste índice, em especial se ele ultrapassou ou não o valor 4 na chegada ao tratamento de canal.
- 2 O objetivo é identificar fatores que aceleram ou desaceleram o crescimento do índice.
- 3 Fatores: período extra-oral, meio de armazenamento, idade, etc.
- 4 Alguns pacientes contribuem com mais de um dente.
- 5 A princípio o estudo é transversal.

Comparação de duas Médias

Retomar a Comparação dos Colírios A e B

- Pacientes com pressão intra-ocular (PIO) elevada irão participar do estudo.
- A pressão será medida após dois meses de uso do colírio.
- O objetivo é comparar a redução média de PIO dos dois colírios.

Então, queremos o seguinte:

$$\delta = \mu_A - \mu_B.$$

O interesse é então testar a hipótese:

$$H_0 : \delta = 0$$

Comparação de duas Médias

- Existem duas formas de conduzir o estudo:
 - 50 pacientes são submetidos ao colírio A e ao colírio B (medidas repetidas). Considera-se uma período de descanso de dois meses entre a aplicação dos colírios. É indicado aleatorizar a ordem de aplicação de A e B.
 - 100 pacientes são selecionados e 50 são sorteados para receber o colírio A e os demais recebem o B.
- Ambos estudos são experimentais
 - Pareado: Estudo Cross-over
 - Amostras Independentes: Estudo Clínico Aleatorizado.
- Qual forma você utilizaria?

Amostra Pareada ou Independente?

1 Vantagens de Parear as Amostras

- Controlar por possíveis fatores de confusão.
- Menos pacientes/unidades na amostra.
- Teste mais preciso com menos suposições.
- Controla pelo efeito de coorte.

2 Vantagens de Amostras Independentes

- Dados são obtidos de forma mais rápida.

Amostras Pareada ou Independente?

Quando devemos parear?

SEMPRE (que for possível).

- Caso típico: antes e depois.
- Situações impossíveis: fumantes e não-fumantes, etc.

Teste-t pareado

- Comparar duas medidas pareadas.
-

$$\delta = \mu_A - \mu_B.$$

Uma estimativa natural para δ é a diferença das médias. Ou seja

$$\widehat{\delta} = \widehat{\mu}_A - \widehat{\mu}_B.$$

A variância de $\widehat{\delta}$ é

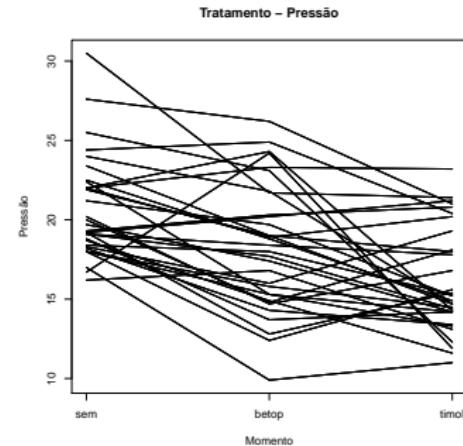
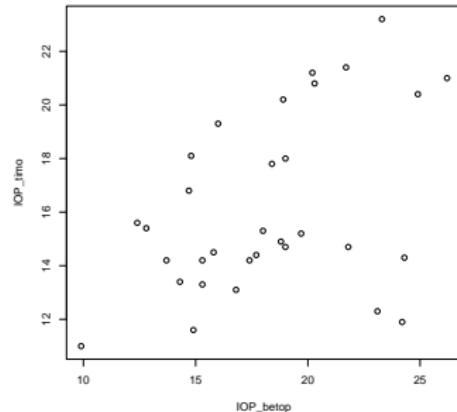
$$Var(\widehat{\delta}) = \frac{1}{n}(\sigma_A^2 + \sigma_B^2 - 2\sigma_{AB})$$

Teste-t pareado

Usualmente dados longitudinais têm correlação positiva. Ou seja $\sigma_{AB} > 0$.

Isto significa que a estatística a ser utilizada tem menor variância do que aquela com dados independentes.

Exemplo: Colírio A: Timoptol (timo) e Colírio B: Betoptic (cor=0,43).



Teste-t pareado

Considere as diferenças:

$$d_i = y_{i1} - y_{i2} \quad i = 1, \dots, n.$$

A estatística é:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

que sob H_0 tem uma distribuição t com $n-1$ graus de liberdade.

Teste para uma única amostra.

Suposição: d_i vem de uma distribuição normal.

Análise de Dados Longitudinais

1 Características:

- As respostas de diferentes unidades são independentes;
- As respostas para a mesma unidade são correlacionadas. De uma forma geral, as respostas próximas no tempo devem ser mais correlacionadas.

2 Medida Temporal

- Idade;
- Calendário medido a partir de um certo evento. Evento: aplicação de um tratamento, diagnóstico de doença, etc.

3 Objetivos do Estudo:

- avaliar o comportamento temporal;
- avaliar o efeito de covariáveis sobre a resposta;
- predição.

4 Modelos de Regressão

- Modelos marginais (modelar a média e a estrutura de covariância);
- Modelo de efeitos aleatórios.
- Modelo de transição.

Características da Correlação dos Dados

- As correlações usualmente são positiva;
- as correlações usualmente diminuem a medida que aumenta a separação no tempo;
- as correlações entre medidas repetidas raramente aproximam do zero.

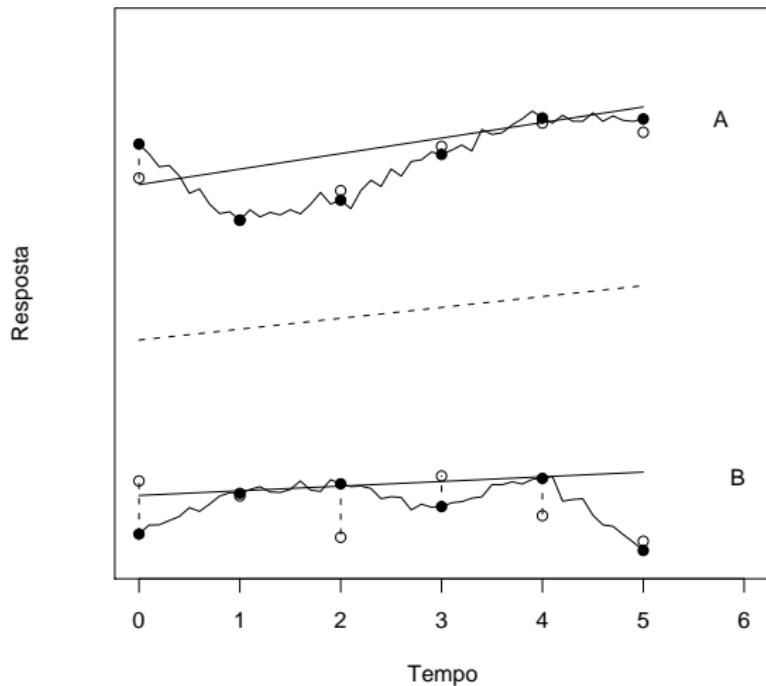
Fontes de Variabilidade em Estudos Longitudinais

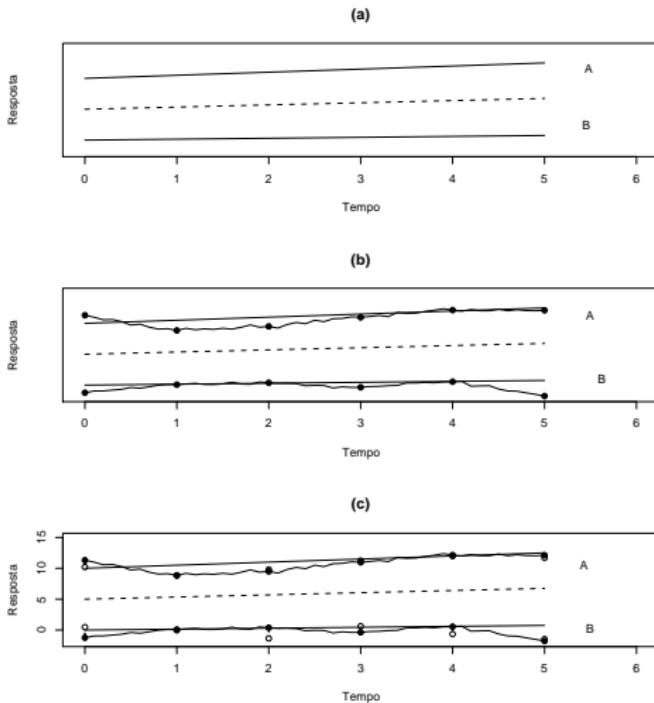
- Variação entre-unidades;
- Variação intra-unidade;
- Erro de medição.

Fontes de Variabilidade em Estudos Longitudinais

Estas três fontes de variação podem ser visualizadas de forma gráfica.

- pontos pretos são respostas livre de erro de medição;
- pontos brancos são as respostas observadas;
- A e B são diferentes indivíduos.





Comparação de mais de duas Médias

Comparação dos Colírios A e B e C

- Pacientes com pressão intra-ocular elevada irão participar do estudo.
- A pressão será medida após dois meses de uso do colírio.
- O objetivo é comparar a redução média dos três colírios.

Então, queremos testar a seguinte hipótese:

$$H_0 : \mu_A = \mu_B = \mu_C.$$

ANOVA é válido?