Especialização em *Data Science* e Estatística Aplicada

Módulo IV – Análise de Dados Longitudinais

Prof. Dr. Márcio Luis Lanfredi Viola

Goiânia, 2025

IME
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
FEN
FACULDADE DE ENFERMAGEM
UFG
UNIVERSIDADE FEDERAL DE GOIÁS

Conteúdo Programático

- Conceitos Básicos.
- Principais delineamentos para dados longitudinais.
- Modelos lineares gerais para dados longitudinais.
- Modelos lineares generalizados para dados longitudinais.
- Modelos com efeitos aleatórios.
- Aplicações.

Conteúdo - Aula 3

- Modelos lineares gerais para dados longitudinais.
- Modelos com efeitos aleatórios.
- Exemplos.

REVISANDO: PERFIL INDIVIDUAL

O perfil individual de resposta com as m observações referentes à *i*-ésima unidade amostral, **i** = 1, ..., n_j, submetida ao *j*-ésimo tratamento, *j* = 1, ..., *p*, é dado por

$$\mathbf{Y}_{ij} = (Y_{ij1}, Y_{ij2}, \ldots, Y_{ijm})^T$$

em que Y_{ijk} denota o valor da variável resposta para a *i*-ésima unidade amostral submetida ao *j*-ésimo tratamento no *k*-ésimo instante, *k* = 1, ..., *m*.

REVISANDO: PERFIS MÉDIOS

Os perfis médios de respostas nos m níveis da condição de avaliação são dados por

$$E(\boldsymbol{Y}_{ij}) = \boldsymbol{\mu}_j = (\mu_{j1}, \ldots, \mu_{jm})^T, \; j = 1, \ldots, p.$$

Como analisar dados longitudinais?

- **Modelo Marginal:** Modela separadamente a média e a estrutura de covariância;

- **Modelo Condicional ou de Efeitos Aleatórios:** Trata os coeficientes como sendo aleatórios para as covariáveis que mudam no tempo (por exemplo, intercepto e coeficiente do tempo).

    - As diferenças entre os perfis surgem porque os coeficientes de regressão variam entre indivíduos;

    - A correlacão entre as medidas no mesmo indivíduo são induzidas pelos efeitos aleatórios.

Efeitos

- **Efeito transversal:** Variação entre unidades amostrais;

- **Efeito longitudinal:** Variação intraunidades amostrais.

ANOVA para medidas repetidas

- **Análise de Variância (ANOVA):** É uma técnica pela qual a variabilidade total de um conjunto de dados é separada em vários componentes;

- Usualmente, cada um desses componentes de variação está associada a uma fonte específica de variação;

- Em qualquer tipo de experimento, é de interesse conhecer a magnitude das contribuições de cada uma dessas fontes para a variação total.

ANOVA para medidas repetidas

- **Objetivo:** Comparar a resposta média em cada tempo;

- Considere o caso simples em que não há efeito de grupos:

$$Y_{ij} = \mu + \alpha_i + \tau_j + \varepsilon_{ij},$$

em que $\varepsilon_{ij} \sim N(0, \sigma^2)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.

ANOVA para medidas repetidas

Neste nosso caso:

- Os blocos são os indivíduos;
- α_i : o efeito do *i*-ésimo bloco (indivíduo, unidade amostral);
- α_i: pode ser tratado como efeito fixo ou aleatório. Neste último caso, $\alpha_i \sim N(0, \sigma_\alpha^2)$;
- Os tratamentos são os próprios tempos;
- τ_j: O efeito do *j*-ésimo tratamento (tempo).

ANOVA para medidas repetidas

$$SQ_{Total} = \sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij} - \bar{y})^2 \qquad \bar{y} = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{y_{ij}}{nm}$$

$$SQ_{Tratamento} = n\sum_{j=1}^{m}(\bar{y}_j - \bar{y})^2 \qquad \bar{y}_j = \sum_{i=1}^{n}\frac{y_{ij}}{n}$$

$$SQ_{Bloco} = m\sum_{i=1}^{n}(\bar{y}_i - \bar{y})^2 \qquad \bar{y}_i = \sum_{j=1}^{m}\frac{y_{ij}}{m}$$

Sob $H_0 : \tau_1 = \ldots = \tau_m$,

$$F = \frac{QM_{Trat}}{QM_{Res}} \sim F_{(m-1),(n-1)(m-1)}$$

Observação

- Podemos utilizar a ANOVA para testar a igualdade de mais de duas médias;

- O teste F vale se $\text{Cov}(\boldsymbol{Y_i}) = \text{Var}((Y_{i1}, \ldots, Y_{im})^T) = \boldsymbol{\Sigma}$, em que

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix}$$

e $\rho = \frac{Cov(y_{ij}, y_{ij'})}{\sigma^2}$.

Observação

- O Teste de Friedman (teste não paramétrico) é uma alternativa para a ANOVA, quando a suposição de normalidade, igualdade de variâncias ou esfericidade não for válida.

- O Teste de Friedman usa os postos dos dados ao invés de seus valores observados para obter a estatística de teste.

Limitações da ANOVA

- Não se aplica em situações em que ocorre desbalanceamento;
- Trata o tempo como categórico;
- Usualmente, a correlação tende a diminuir à medida que aumentamos a distância temporal, ou seja, a estrutura esférica não é adequada;
- Difícil de ser utilizado na presença de covariáveis contínuas.
- Variável resposta possuindo distribuição Normal.

Modelos Mistos

Em geral, podemos analisar dados provenientes de estudos com medidas repetidas, por meio de modelos mistos da forma

$$\boldsymbol{Y}_i = g(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) + \boldsymbol{\varepsilon}_i, \; i = 1, ..., n,$$

em que:

- $\boldsymbol{Y}_\mathrm{i} = (y_{i1}, ..., y_{imi})^T$, com dimensão $m_i \times 1$, é o perfil de respostas da *i*-ésima unidade experimental.

Modelos Mistos

- $\boldsymbol{\beta}$ é um vetor de parâmetros (efeitos fixos ou parâmetros de localização) desconhecidos, com dimensão $p \times 1$;

- $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ip})$ é uma matriz de especificação dos efeitos fixos com dimensão $(m_i \times p)$, conhecida e de posto completo, em que $\boldsymbol{x}_{ij} = (x_{ij1}, \ldots, x_{ijm_i})^T$ representa o vetor com os m_i valores da j-ésima variável independente, $j = 1, \ldots, p$, para a i-ésima unidade amostral.

Modelos Mistos

- $\boldsymbol{b}_i$ é um vetor, com dimensão $q \times 1$, de variáveis latentes, comumente denominadas efeitos aleatórios, que refletem o comportamento individual da i-ésima unidade amostral;

- $\boldsymbol{Z}_i$ é uma matriz de especificação dos efeitos aleatórios, com dimensão $m_i \times q$, conhecida e de posto completo;

- g é uma função com derivadas contínuas até à segunda ordem;

- $\boldsymbol{\varepsilon}_i$ é um vetor de erros aleatórios com dimensão $m_i \times 1$.

Modelos Mistos

- Em muitos casos, é razoável supor que $\boldsymbol{b}_i \sim N_q(\mathbf{0}, \boldsymbol{G})$ e $\varepsilon_i \sim N_{mi}(\mathbf{0}, \boldsymbol{R}_i)$, em que $\boldsymbol{G}$, com dimensão $q \times q$, e $\boldsymbol{R}_i$, com dimensão $m_i \times m_i$, são matrizes simétricas definidas positivas e, além disso, que $\boldsymbol{b}_i$ e ε_i são variáveis aleatórias independentes;

- Quando g é uma função linear dos efeitos fixos $\boldsymbol{\beta}$ e dos efeitos aleatórios $\boldsymbol{b}_i$, o modelo definido anteriormente se reduz ao modelo linear misto

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i + \varepsilon_i, \; i = 1, ..., n.$$

Modelos Mistos

Sob esse modelo, o vetor de respostas associado à *i*-ésima unidade amostral tem distribuição normal multivariada com vetor de médias e matriz de covariâncias dados, respectivamente, por

$$\mathrm{E}(\boldsymbol{Y}_i) = \boldsymbol{X}_i \boldsymbol{\beta}$$

e

$$\mathrm{Var}(\boldsymbol{Y}_i) = \boldsymbol{\Omega}_i = \boldsymbol{Z}_i \boldsymbol{G} \boldsymbol{Z}_i^{\mathrm{T}} + \boldsymbol{R}_i \, .$$

Modelos Mistos

- A primeira componente da decomposição de Var($\boldsymbol{Y}_i$) modela a dispersão dos perfis individuais de resposta e a segunda componente está relacionada com a dispersão da resposta em torno dos perfis individuais;

- Se eliminarmos os componentes aleatórios $\boldsymbol{b}_i$ e, consequentemente, as matrizes $\boldsymbol{Z}_i$ do modelo linear misto, obteremos os chamados modelos marginais, nos quais a estrutura de covariância é modelada diretamente por meio das matrizes $\boldsymbol{R}_i$.

Modelos Mistos

- O modelo linear misto pode ser interpretado como um modelo linear em dois estágios;

- No primeiro estágio, consideramos fixos os efeitos aleatórios $\boldsymbol{b}_i$, de forma que $\boldsymbol{Y}_i|\boldsymbol{b}_i \sim \mathrm{N}(\boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i, \boldsymbol{R}_i)$;

- No segundo estágio, supomos que os vetores $\boldsymbol{b}_i$ são independentes com distribuição $\mathrm{N}_q(\mathbf{0}, \boldsymbol{G})$ e, consequentemente, o modelo para a distribuição marginal de $\boldsymbol{Y}_i$ é $\boldsymbol{Y}_i \sim \mathrm{N}(\boldsymbol{X}_i \boldsymbol{\beta}, \boldsymbol{Z}_i \boldsymbol{G} \boldsymbol{Z}_i^{\top} + \boldsymbol{R}_i)$.

Observação

Quando $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}_{mi}$, o modelo é chamado de modelo de independência condicional homocedástico, indicando que as m_i observações associadas à i-ésima unidade amostral são condicionalmente independentes dado $\boldsymbol{b}_i$.

Estrutura de covariância

- Grande parte do esforço empregado na modelagem de dados com medidas repetidas se concentra na estrutura de covariância;

- Em geral, o modelo para a matriz de covariâncias $\boldsymbol{\Omega}_i$ deve depender da maneira pela qual as observações foram obtidas e do conhecimento sobre o mecanismo gerador das observações.

Estrutura de covariância

- Diggle (1988) e Diggle et al. (2002) comentam que a matriz de covariâncias deve ser suficientemente flexível para incluir no mínimo três fontes diferentes de variação aleatória:

  (i) a variação devida a efeitos aleatórios, quando as unidades de investigação formam uma amostra aleatória da população de interesse;

  (ii) a variação que pode ser explicada por correlação serial, em que se esperam observações próximas mais fortemente correlacionadas que observações mais distantes;

  (iii) a variação devida a erros de medida.

Estrutura de covariância

- No contexto dos modelos mistos, a covariância entre as observações obtidas em uma mesma unidade amostral poderá ser modelada diretamente por meio da matriz $\boldsymbol{R}_i$, que representa a estrutura de covariância das observações intraunidades amostrais ou indiretamente, por meio de uma combinação da matriz **G** associada aos efeitos aleatórios $\boldsymbol{b}_i$, que representam a variabilidade entre as unidades amostrais com a matriz $\boldsymbol{R}_i$;

- Diggle et al. (2002) apresentam sugestões sobre as diferentes maneiras em que modelos para matrizes $\boldsymbol{G}$ e $\boldsymbol{R}_i$ podem incluir três possíveis fontes de variação na estrutura de covariância: aquelas correspondente aos efeitos aleatórios, à correlação serial e aos erros de medida.

Estrutura de covariância

- A seguir apresentamos alguns exemplos de estruturas para a matriz $\boldsymbol{\Omega}_i$ como para suas componentes $\boldsymbol{R}_i$ e $\mathbf{G}$;

- Para isto, usamos a notação $\boldsymbol{V}_i = \text{Var}(\boldsymbol{Y}_i)$ e consideramos, para fins de ilustração, $m_i = 4$.

Estrutura Uniforme

- $\theta = (\sigma^2, \tau)^T$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2+\tau & \tau & \tau & \tau \\ \tau & \sigma^2+\tau & \tau & \tau \\ \tau & \tau & \sigma^2+\tau & \tau \\ \tau & \tau & \tau & \sigma^2+\tau \end{bmatrix}$$

Estrutura AR(1)

- $\theta = (\sigma^2, \phi)^T$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

Estrutura ARMA(1,1)

- $\theta = (\sigma^2, \gamma, \phi)^T$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \gamma & \gamma\phi & \gamma\phi^2 \\ \gamma & 1 & \gamma & \gamma\phi \\ \gamma\phi & \gamma & 1 & \gamma \\ \gamma\phi^2 & \gamma\phi & \gamma & 1 \end{bmatrix}$$

Estrutura uniforme heterogênea

- $\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \rho)^T$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$$

Não Estruturada

- $\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_{12}, \sigma_{13}, \sigma_{14}, \sigma_{23}, \sigma_{24}, \sigma_{34})^T$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

Observações

- Um modelo para a estrutura de covariância deve ser escolhido com base no modelo para a média;

- Utilizar os resíduos para remover o efeito de covariáveis;

- Uma forma alternativa é utilizar variável resposta padronizada por tempo, que somente é válido para estudos balanceados;

- A matriz de correlação empírica para os resíduos indica uma possível estrutura de covariância.

Observações

- **IMPORTANTE:** Existe uma interdependência entre a resposta média e a estrutura de covariância;
- Variâncias e covariâncias são inflacionadas quando a estrutura para a média é incorretamente especificada;
- A covariância entre pares de resíduos depende do modelo para a média;
- O modelo para a covariância deve, então, ser escolhido com base no modelo para a média;
- A decisão deve ser feita com base nos resíduos.

Observações

- Para modelar a estrutura de covariância, pode-se utilizar o modelo "maximal" para a média e testar diferentes estruturas para a covariância;

- Existem algumas estruturas disponíveis para desenhos balanceados com poucos tempos de observação.

Exemplo 1: Níveis de chumbo no sangue

- Tratamento de crianças expostas ao chumbo;

- Estudo clínico aleatorizado para placebo e um tratamento em criança com níveis de chumbo no sangue entre 20-44 microgramas/dL;

- Quatro medidas repetidas de níveis de chumbo na linha de base semana 0, semana 1, semana 4 e semana 6 (desenho balanceado mas não igualmente espaçado);

- 100 crianças aleatoriamente alocadas entre tratamento e placebo.

Tabela 1: Estatísticas descritivas do nível de chumbo, por semana.

Semana	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0	19.70	22.05	25.60	26.41	29.60	41.10
1	2.80	12.38	20.60	19.09	25.02	40.80
4	3.00	15.25	19.70	19.79	24.60	40.40
6	4.10	18.25	21.25	22.20	25.60	63.90

Figura 1: Gráfico Box Plot do nível de chumbo, por semana.

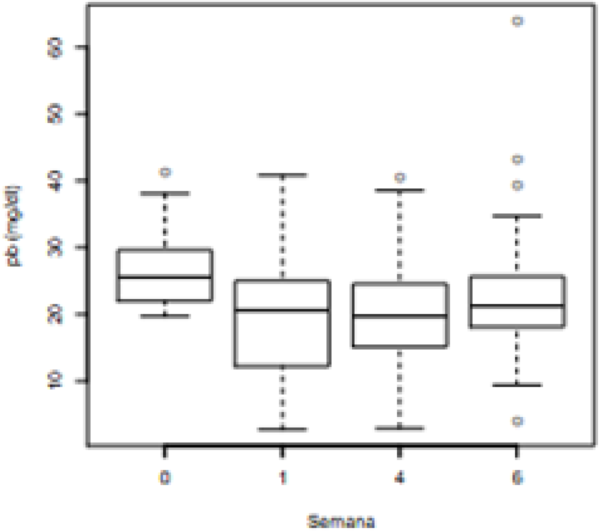


Tabela 2: Estatísticas descritivas do nível de chumbo, por grupo.

Grupo	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
A	2.8	12.38	19.15	19.08	24.6	63.9
P	13.5	20.67	23.90	24.66	27.9	43.3

Figura 2: Gráfico Box Plot do nível de chumbo, por grupo.

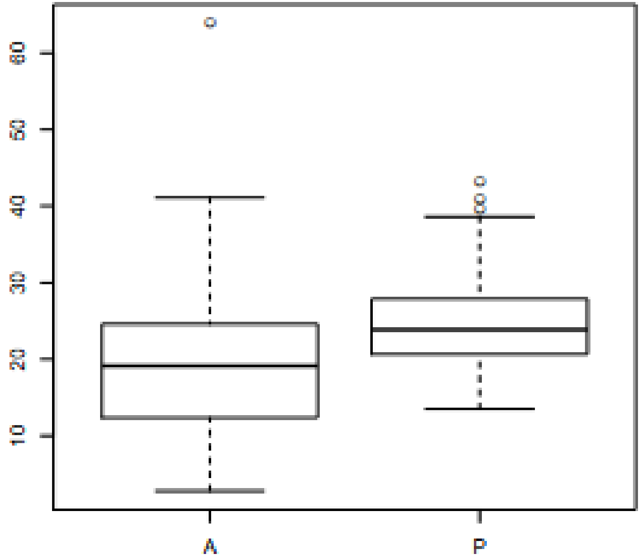


Tabela 3: Estatísticas descritivas do nível de chumbo, por grupo e semana.

	Semana	Min	1º Quartil	Mediana	Média	3º Quartil	Max
Grupo A	0	19.7	22.13	26.20	26.54	29.55	41.1
	1	2.8	7.23	12.25	13.52	17.50	39.0
	4	3.0	9.13	15.35	15.51	19.73	40.4
	6	4.1	15.40	18.85	20.76	23.75	63.9
Grupo P	0	19.7	21.88	25.25	26.27	29.73	38.1
	1	14.9	20.92	24.10	24.66	27.82	40.8
	4	15.3	19.83	22.45	24.07	27.45	38.6
	6	13.5	19.95	22.35	23.65	27.50	43.3

Figura 3: Gráfico de perfis.

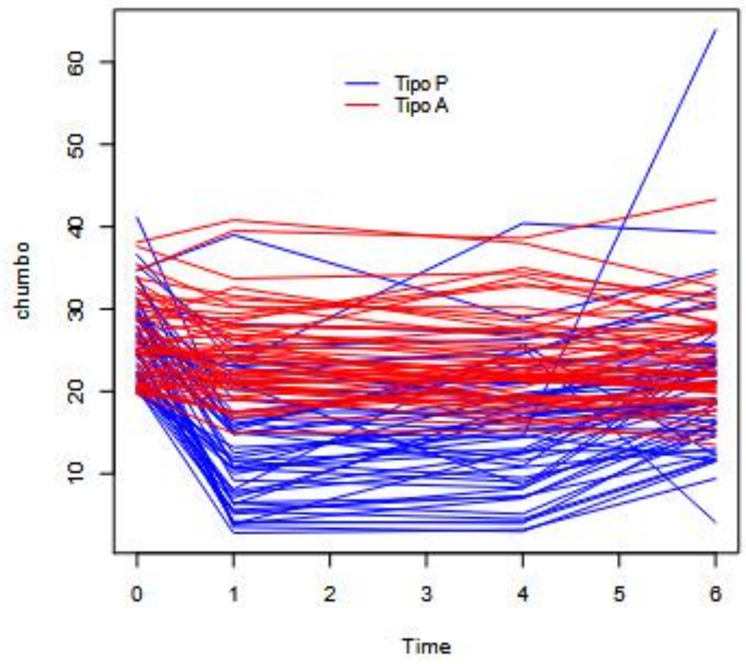
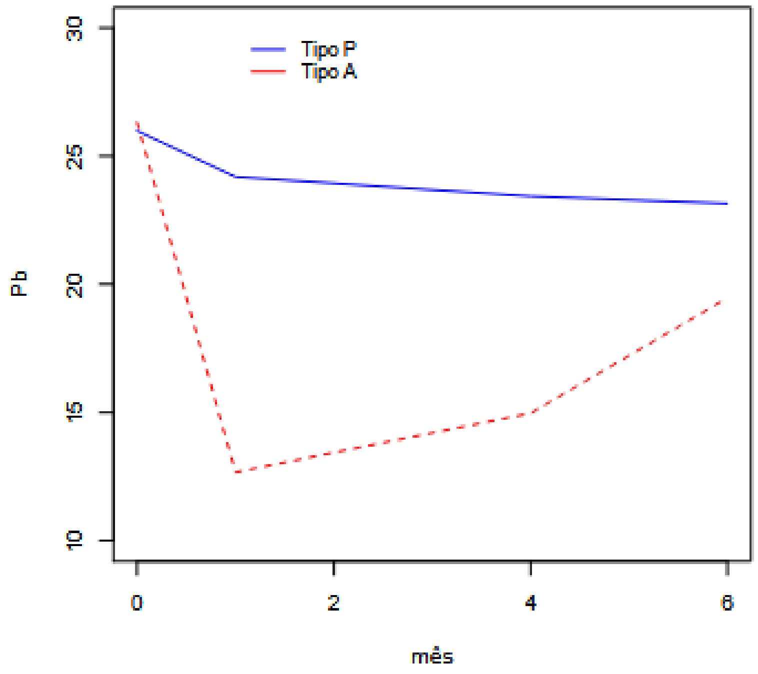


Figura 4: Gráfico de perfis médios.



Exemplo 1: Níveis de chumbo no sangue

- Observa-se que, para o grupo de crianças que tomaram placebo, os níveis de chumbo praticamente permanecem inalterados ao longo do estudo;

- Já para o grupo de crianças submetidas ao agente, os níveis de chumbo se reduzem bastante da semana 0 para a semana 1, e a partir daí têm um ligeiro aumento até a semana 4 e a semana 6;

- Excetuando-se os níveis obtidos na *baseline* (semana 0), que foram praticamente idênticos para ambos os grupos, os níveis de chumbo do grupo tratamento são inferiores aos níveis do grupo placebo.

Tabela 4: Matriz de correlações entre os níveis de chumbo por tempo, usando os dados brutos.

	dados.Week0	dados.Week1	dados.Week4	dados.Week6
dados.Week0	1.000	0.419	0.468	0.562
dados.Week1	0.419	1.000	0.845	0.557
dados.Week4	0.468	0.845	1.000	0.583
dados.Week6	0.562	0.557	0.583	1.000

Observação: Usualmente não é adequado obter as correlações utilizando os dados brutos.

Tabela 5: Matriz de correlações entre os níveis de chumbo padronizados por tempo.

	Week0	Week1	Week4	Week6
Week0	1.00			
Week1	0.62	1.00		
Week4	0.61	0.80	1.00	
Week6	0.63	0.63	0.66	1.00

Figura 5: Gráficos de dispersão dos níveis de chumbo padronizados por tempo.

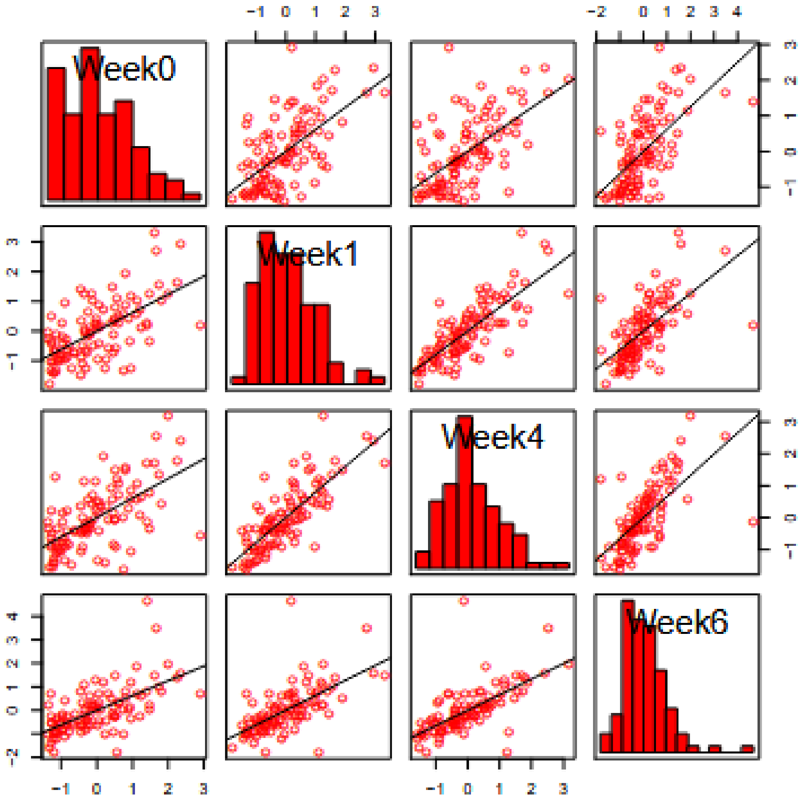


Tabela 6: Matriz de correlações entre os níveis de chumbo padronizados por tempo, para o grupo placebo.

	Week0	Week1	Week4	Week6
Week0	1.00			
Week1	0.83	1.00		
Week4	0.84	0.86	1.00	
Week6	0.76	0.76	0.87	1.00

Tabela 7: Matriz de correlações entre os níveis de chumbo padronizados por tempo, para o grupo tratamento.

	Week0	Week1	Week4	Week6
Week0	1.00			
Week1	0.40	1.00		
Week4	0.38	0.73	1.00	
Week6	0.50	0.50	0.45	1.00

Exemplo 1: Níveis de chumbo no sangue

- Utilizando o modelo com todos os termos (tempo + grupo + interação) e ajuste de Mínimos Quadrados Ordinários.

Tabela 8: Matriz de correlações para os resíduos.

	Week0	Week1	Week4	Week6
Week0	1.00			
Week1	0.57	1.00		
Week4	0.57	0.78	1.00	
Week6	0.58	0.58	0.58	1.00

Exemplo 1: Níveis de chumbo no sangue

- Utilizando o modelo com todos os termos (tempo + grupo + interação) e ajuste de Mínimos Quadrados Ordinários.

Tabela 8: Matriz de covariâncias para os resíduos.

	Week0	Week1	Week4	Week6
Week0	25			
Week1	19	44		
Week4	20	35	47	
Week6	22	29	30	58

Exemplo 1: Níveis de chumbo no sangue

- Ajustou-se o modelo "maximal" que consiste nos seguintes termos: (1) intercepto; (2) todos efeitos principais (1 + 3 parâmetros) e (3) interação tempo e grupo (3 parâmetros);

- Iniciou-se com a forma não estruturada de covariância e testou-se possíveis formas estruturadas de covariância.

Exemplo 1: Níveis de chumbo no sangue

- O modelo final com estrutura de variância heterocedástica e não estruturada pareceu ser o mais indicado para a análise destes dados;

- O gráfico de resíduos versus ajustados pode ser útil para validarmos o modelo ajustado.

Tabela 9: Estimativas dos coeficientes do modelo.

Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	26.272	0.7102939	36.98751	0.0000
factor(tempo)1	-1.612	0.7919197	-2.03556	0.0425
factor(tempo)4	-2.202	0.8149099	-2.70214	0.0072
factor(tempo)6	-2.626	0.8885239	-2.95546	0.0033
factor(Groupo)A	0.268	1.0045072	0.26680	0.7898
factor(tempo)1:factor(Groupo)A	-11.406	1.1199436	-10.18444	0.0000
factor(tempo)4:factor(Groupo)A	-8.824	1.1524566	-7.65669	0.0000
factor(tempo)6:factor(Groupo)A	-3.152	1.2565625	-2.50843	0.0125

Observações

- Os termos de ordem inferior devem sempre aparecer no modelo na presença daqueles de ordem superior, ou seja, ao incluirmos a interação entre grupo e tempo, os efeitos principais de grupo e tempo devem aparecer na estrutura de regressão;

- Uma exceção pode ser aberta, ao assumirmos, em estudos experimentais, igualdade de média na linha de base. Desta forma, excluímos o efeito principal de grupo;

- A vantagem deste ajuste é a parcimônia, ao reduzirmos um parâmetro do modelo

Tabela 10: Estimativas dos coeficientes do modelo supondo igualdade de médias na linha de base.

```
Coefficients:
                                    Value Std.Error   t-value p-value
(Intercept)                     26.406000 0.4998905  52.82357  0.0000
factor(tempo)1                  -1.644501 0.7824043  -2.10186  0.0362
factor(tempo)4                  -2.231356 0.8073816  -2.76369  0.0060
factor(tempo)6                  -2.642065 0.8864611  -2.98046  0.0031
factor(tempo)1:factor(Grpo)A   -11.340999 1.0931205 -10.37488  0.0000
factor(tempo)4:factor(Grpo)A    -8.765289 1.1312579  -7.74827  0.0000
factor(tempo)6:factor(Groupo)A-3.119870 1.2507769  -2.49435  0.0130
```

- Tempo = 0 (linha de base): $E(Y_{ij}|Grupo = P) = E(Y_{ij}|Grupo = A) = \beta_1$
  Estimativa: $26,4 \pm 1,96 * 0.5 \quad (25,5; 27,3)$

- Tempo = 1: $E(Y_{ij}|Grupo = A) = E(Y_{ij}|Grupo = P) = (\beta_1 + \beta_2 + \beta_5) - (\beta_1 + \beta_2) = \beta_5$
  Estimativa: $= -11,3 \pm 1,96 * 1.09 \quad (-13,5; -9,1)$
- Houve uma redução da resposta do tempo =1 para o 6 no grupo =P (efeito placebo)?
  Grupo = P: $E(Y_{ij}|Tempo = 1) = E(Y_{ij}|Tempo = 6) = \beta_2 - \beta_4$. Estimativa:
  $-1,64 + 2.64 = 1.0$ e a $\widehat{Var}(\hat{\beta}_2 - \hat{\beta}_4) = \widehat{Var}(\hat{\beta}_2) + \widehat{Var}(\hat{\beta}_4) - 2\widehat{Cov}(\hat{\beta}_2, \hat{\beta}_4) = 0.6^2$,
  IC $(-0,17; 2,17)$, não existe evidência de efeito placebo.

Especialização em *Data Science* e Estatística Aplicada

Módulo IV – Análise de Dados Longitudinais

Prof. Dr. Márcio Luis Lanfredi Viola

lanfredi@ufscar.br

IME INSTITUTO DE MATEMÁTICA E ESTATÍSTICA FEN FACULDADE DE ENFERMAGEM UFG UNIVERSIDADE FEDERAL DE GOIÁS