

# Avaliação parcial da Área 1

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DEPARTAMENTO DE ESTATÍSTICA  
MAT02035 - MODELOS PARA DADOS CORRELACIONADOS

Rodrigo Citton P. dos Reis  
citton.padilha@ufrgs.br

**Questão 1 (1 ponto)** Apresente as principais características de um estudo longitudinal.

**Questão 2 (1 ponto)** Para um estudo longitudinal, descreva as diferenças entre os delineamentos balanceado e desbalanceado.

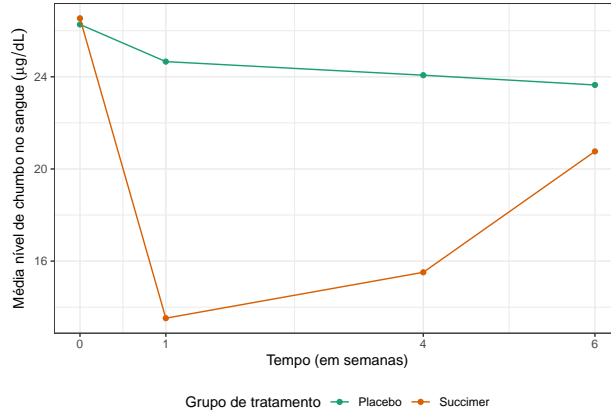
**Questão 3 (2 pontos)** No estudo TLC (*estudo sobre tratamento de crianças expostas ao chumbo*), considere somente as duas primeiras medidas do estudo: linha de base (semana 0) e semana 1. Suponha que é de interesse determinar se existe uma mudança na resposta média ao longo do tempo. Uma estimativa da mudança na resposta média é dada por  $\hat{\delta} = \hat{\mu}_2 - \hat{\mu}_1$ , em que  $\hat{\mu}_j = \frac{1}{N} \sum_{i=1}^N Y_{ij}$ ,  $j = 1, 2$ . A expressão da variância de  $\hat{\delta}$  é dada por

$$\text{Var}(\hat{\delta}) = \text{Var} \left\{ \frac{1}{N} \sum_{i=1}^N (Y_{i2} - Y_{i1}) \right\} = \frac{1}{N} (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}).$$

Note que a expressão acima inclui o termo  $-2\sigma_{12}$ . Este termo é o responsável por levar em consideração a correlação entre as duas primeiras medidas repetidas. Para os dados do estudo, no grupo **succimer** ( $N = 50$ ), temos  $\hat{\delta} = 13.5 - 26.5 = -13$ ,  $\hat{\sigma}_1^2 = 25.2$ ,  $\hat{\sigma}_2^2 = 58.9$  e  $\hat{\sigma}_{12} = 15.5$ .

1. Com os dados apresentados acima, calcule  $\widehat{\text{Var}}(\hat{\delta})$ . Logo em seguida, calcule  $\widehat{\text{Var}}(\hat{\delta})$ , mas agora assumindo que todas as observações são independentes.
2. Utilize os resultados do item anterior e discuta as consequências de se ignorar a correlação entre os dados de medidas repetidas.

**Questão 4 (3 pontos)** Mais uma vez, considere o estudo TLC, um estudo aleatorizado, controlado por placebo, de um agente quelante administrado por via oral, o *succimer*, em crianças com níveis confirmados de chumbo no sangue de 20 a 44  $\mu\text{g/dL}$ . As crianças do estudo tinham idades entre 12 e 33 meses e viviam em moradias deterioradas no centro da cidade. Dados sobre os níveis de chumbo no sangue na linha de base (semana 0), e nas semanas 1, 4 e 6 do primeiro período de tratamento foram coletados. Na figura abaixo é apresentado o gráfico de médias de repostas por grupos de tratamento em cada ocasião de medição.



1. Sugira um modelo de regressão para a resposta média. Justifique a sua escolha. Como você classifica a abordagem do modelo escolhido: perfis de resposta ou curvas paramétricas?
2. Com base no modelo que você propôs no item anterior, apresente a matriz de delineamento  $X_i$  para um indivíduo do grupo *succimer* e para um indivíduo do grupo placebo.
3. Com base no modelo que você propôs, apresente uma hipótese estatística para testar se o padrão de mudanças ao longo do tempo difere entre os grupos. Além disso, descreva o teste (Wald, Razão de Verossimilhanças, ou outro) e a distribuição de referência.

**Questão 5 (3 pontos)** Considere o estudo de crescimento dentário, as medidas da distância (mm) do centro da glândula pituitária à fissura pteromaxilar foram obtidas em 11 meninas e 16 meninos nas idades de 8, 10, 12 e 14 (Potthoff e Roy, 1964). Um modelo saturado (considerando efeitos de tempo, grupo [gênero], e interações  $tempo \times grupo$ ) foi assumido para a resposta média, e os seguintes modelos para a covariância foram ajustados: (**mod.ne**) covariância não estruturada; (**mod.sc**) simetria composta; (**mod.ar**) autorregressiva.

1. Para este estudo acima, descreva  $N$ , o número de indivíduos,  $n$ , o número de medidas repetidas, e  $t_{ij}$ .
2. Para cada um dos modelos de covariância, escreva a matriz  $Cov(Y_i)$ , em que  $Y_i$  representa o vetor de respostas.
3. Com base nas comparações abaixo ( $gl$  representa os graus de liberdade, AIC é o Critério de Informação de Akaike, logVero é a log-verossimilhança maximizada, RV é a razão de verossimilhança), escolha um modelo de covariância se ajuste adequadamente aos dados. Justifique a sua resposta.

	Modelo	$gl$	AIC	logVero	Teste	RV	Valor $p$
mod.ne	1	18	450.0348	-207.0174			
mod.ar	2	10	454.5472	-217.2736	1 vs 2	20.51237	0.0086

	Modelo	$gl$	AIC	logVero	Teste	RV	Valor $p$
mod.ne	1	18	450.0348	-207.0174			
mod.sc	2	10	443.4085	-211.7043	1 vs 2	9.373732	0.3118

	Modelo	$gl$	AIC	logVero
mod.sc	1	10	443.4085	-211.7043
mod.ar	2	10	454.5472	-217.2736