

Dados longitudinais: introdução a *anova de medidas repetidas (ANOVA-MR)*

Pedro S.R. Martins

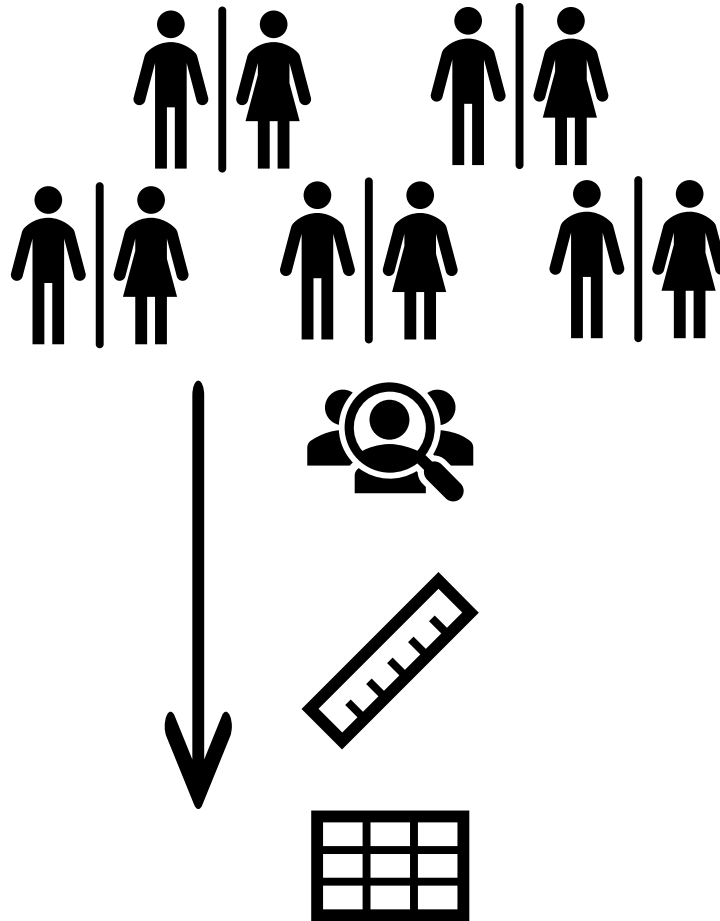
Lavis 2023



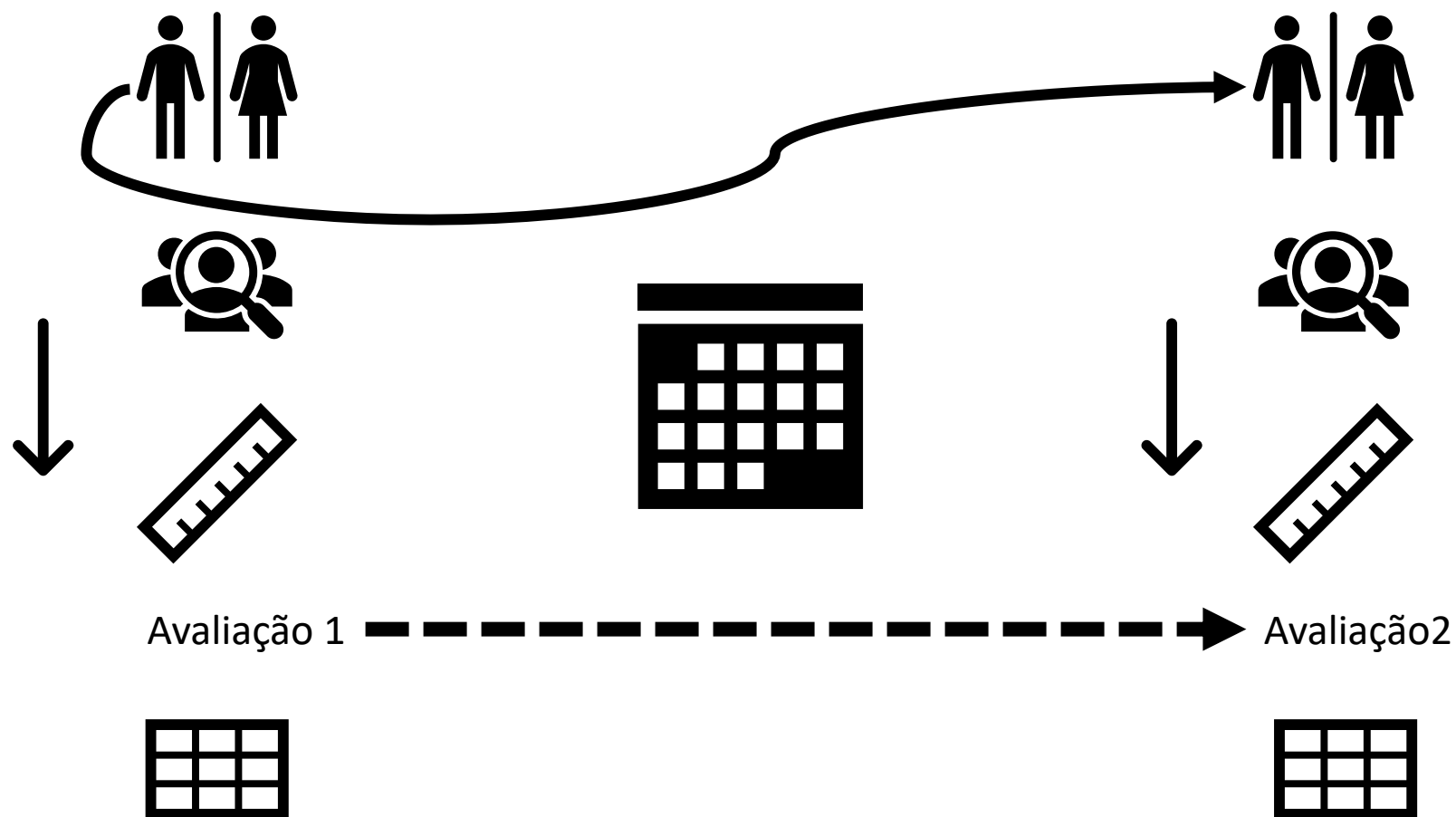
Roteiro

- Dados longitudinais
- Comparações com medidas repetidas

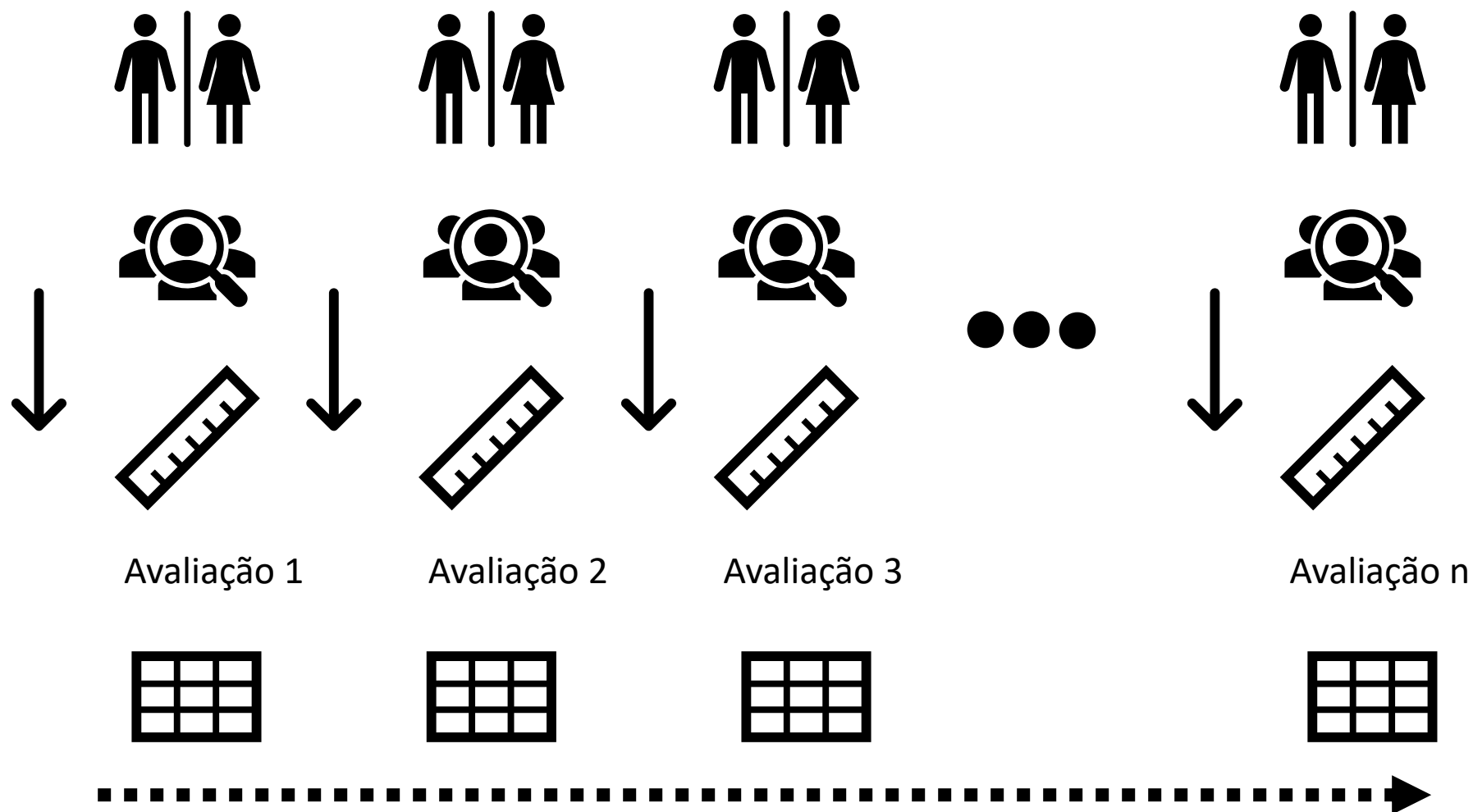
Dados Transversais



Dados longitudinais



Dados longitudinais

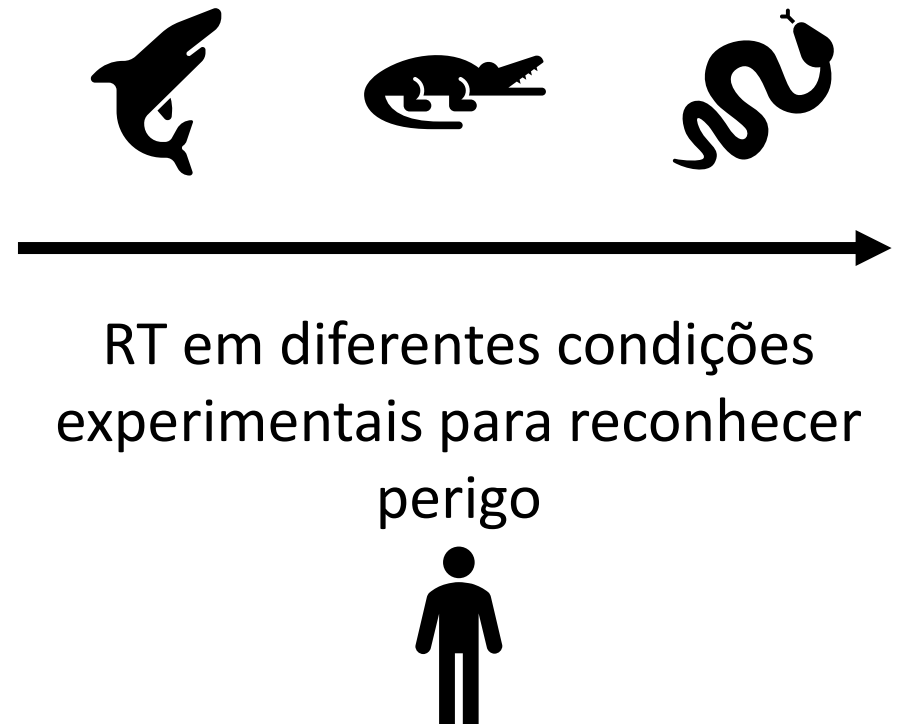


Dados longitudinais

- Dados longitudinais são aqueles em que as mesmas pessoas são avaliadas mais de uma vez
- **Dados em painel** (*panel data*), termo mais usado em econometria e faz referência a dados longitudinais
- **Cuidado:** “medidas repetidas” não é sinônimo de dado longitudinal



A pessoa pode ser avaliada uma vez só em diferentes condições experimentais



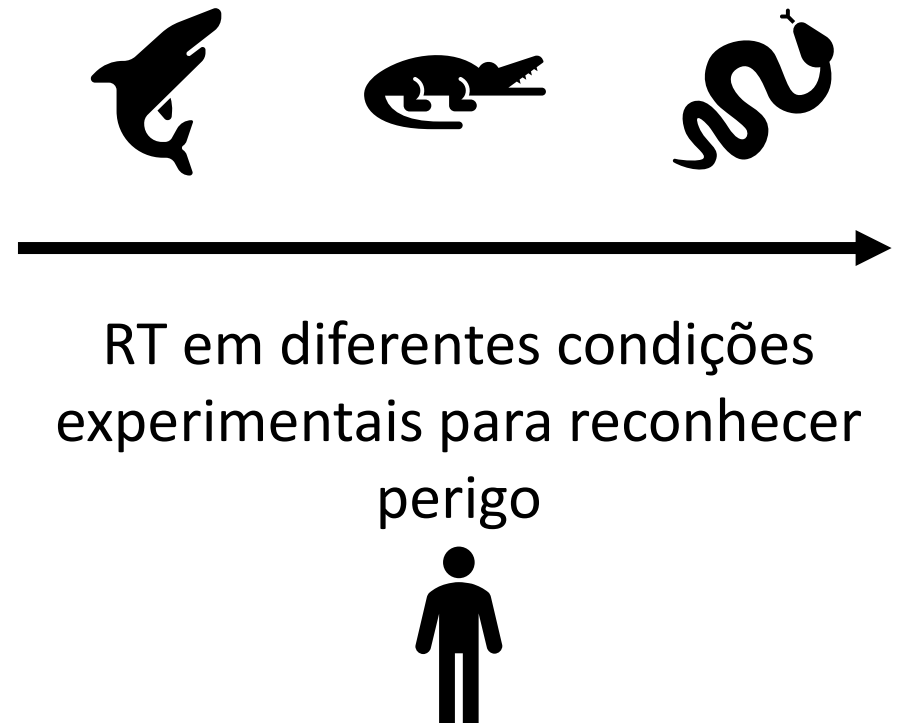
Dados longitudinais

- Dados longitudinais são aqueles em que as mesmas pessoas são avaliadas mais de uma vez
- **Dados em painel** (*panel data*), termo mais usado em econometria e faz referência a dados longitudinais
- **Cuidado:** “medidas repetidas” não é sinônimo de dado longitudinal



A pessoa pode ser avaliada uma vez só em diferentes condições experimentais

→ Outros nomes: delineamento intra-participante ou intrassujeito



Dados longitudinais

Vantagens

- Pode necessitar de menos indivíduos para achar um efeito
- Cada sujeito serve como seu próprio controle, *within subject design*
- Oferece informações sobre padrões de mudança
- Possibilidades de avaliar a confiabilidade das medidas (e.g., McCrae e Möttus, 2019, McCrae, 2015)*

Desvantagens

- Diferenças entre efeitos marginais e condicionais
- Dados desbalanceados (e.g., diferenças entre N nos tempos)
- Dificuldades para analisar e interpretar dados longitudinais
- \$

Estrutura do banco de dados

Dados em formato *wide* (largo)

Id	Medida_av1	Medida_av2	Medida_av3	Medida_avn
1	2	3	6	20
2	5	6	3	1
3	7	5	6	2
4	10	5	6	5

Estrutura do banco de dados

Dados em formato *long* (longo)

Id	Medida_av1	Medida_av2	Medida_av3	Medida_avn
1	2	3	6	20
2	5	6	3	1
3	7	5	6	2
4	10	5	6	5

Id	Tempo	Medida
1	av1	2
1	av2	3
1	av3	6
1	avn	20
2	av1	5
2	av2	6
2	av3	3
2	avn	1
3	av1	7
3	av2	5
3	av3	6
3	avn	2
4	av1	10
4	av2	5
4	av3	6
4	avn	5

Estrutura do banco de dados

- O formato do banco pode variar de acordo com o programa e com a análise que será realizada. **Não existe um formato melhor do que outro.**
- O mais importante é saber que os dois existem e entender como cada programa trabalha os dados

Id	Medida_av1	Medida_av2	Medida_av3	Medida_avn
1	2	3	6	20
2	5	6	3	1
3	7	5	6	2
4	10	5	6	5

Id	Tempo	Medida
1	av1	2
1	av2	3
1	av3	6
1	avn	20
2	av1	5
2	av2	6
2	av3	3
2	avn	1
3	av1	7
3	av2	5
3	av3	6
3	avn	2
4	av1	10
4	av2	5
4	av3	6
4	avn	5



Recurso extra:

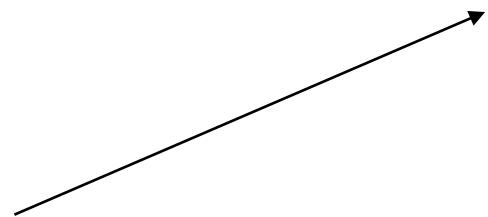
<https://github.com/Pedro-SR-Martins/GrupoEstudosEstatistica2023/tree/main/manipula%C3%A7%C3%A3o%20banco%20de%20dados>

Modelos de regressão

Todas as análises mencionadas anteriormente são extrapolações dos modelos de regressão

- $y = a \cdot x + b$

- $Y = \beta_0 + \beta_{x1} + \varepsilon$

- 
- Em dados longitudinais, o fato de ser a mesma pessoa variáveis vezes aparece no termo do erro.
 - Na regressão, os erros não devem ser correlacionados
 - Em dados longitudinais, esse pressuposto é quebrado

Dados longitudinais

- Duas tradições para o uso de dados longitudinais usando os modelos lineares
 - *Alterações na equação para inserir a variabilidade intra participantes*
 - *Imposição de pressupostos adicionais para calcular a variabilidade intra participantes*

Comparações dados longitudinais

Dados transversais

2 grupos

Distribuição normal – teste t

Não paramétrico – Mann-Whitney

3 + grupos

Distribuição normal – ANOVA

Não paramétrico – Kruskal-Wallis

Dados longitudinais

2 tempos

Distribuição normal – teste t pareado

Não paramétrico – Wilcoxon

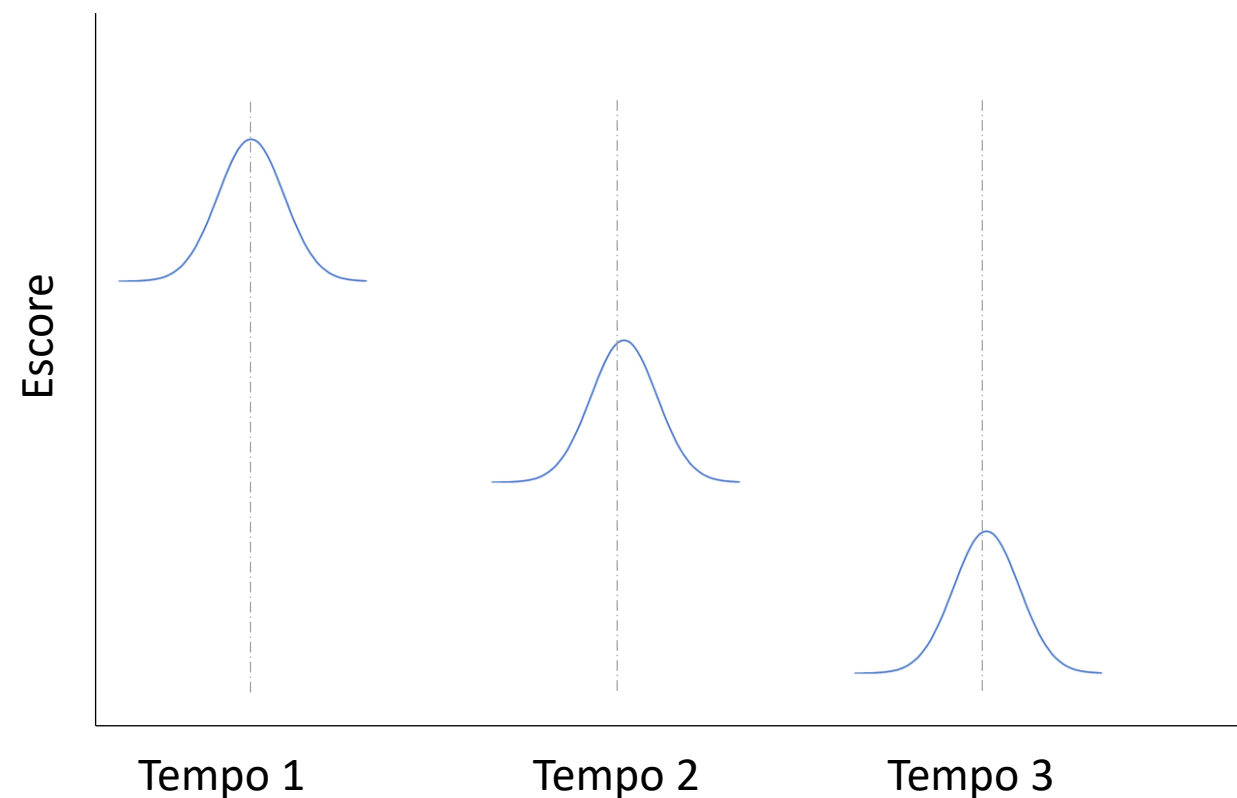
3 + tempos

Distribuição normal – ANOVA de medidas repetidas

Não paramétrico – ANOVA de Friedman

ANOVA de medidas repetidas

Id	Medida_av1	Medida_av2	Medida_av3	Medida_avn
1	2	3	6	20
2	5	6	3	1
3	7	5	6	2
4	10	5	6	5



Pressupostos

- *VD quantitativa e intervalar*
- *VD com Distribuição normal multivariada*
- *Igualdade das variâncias entre os momentos de avaliação*
- *Covariância constante entre os tempos de avaliação (esfericidade)*

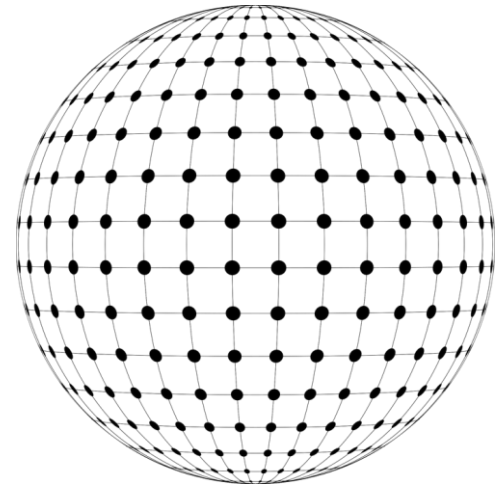
Esfericidade

$$\begin{array}{c} \text{T1} \quad \text{T2} \quad \text{T3} \\ \text{T1} \left[\begin{array}{ccc} 1 & \rho & \rho \\ \text{T2} & \rho & 1 & \rho \\ \text{T3} & \rho & \rho & 1 \end{array} \right] \end{array}$$

Efeitos lineares

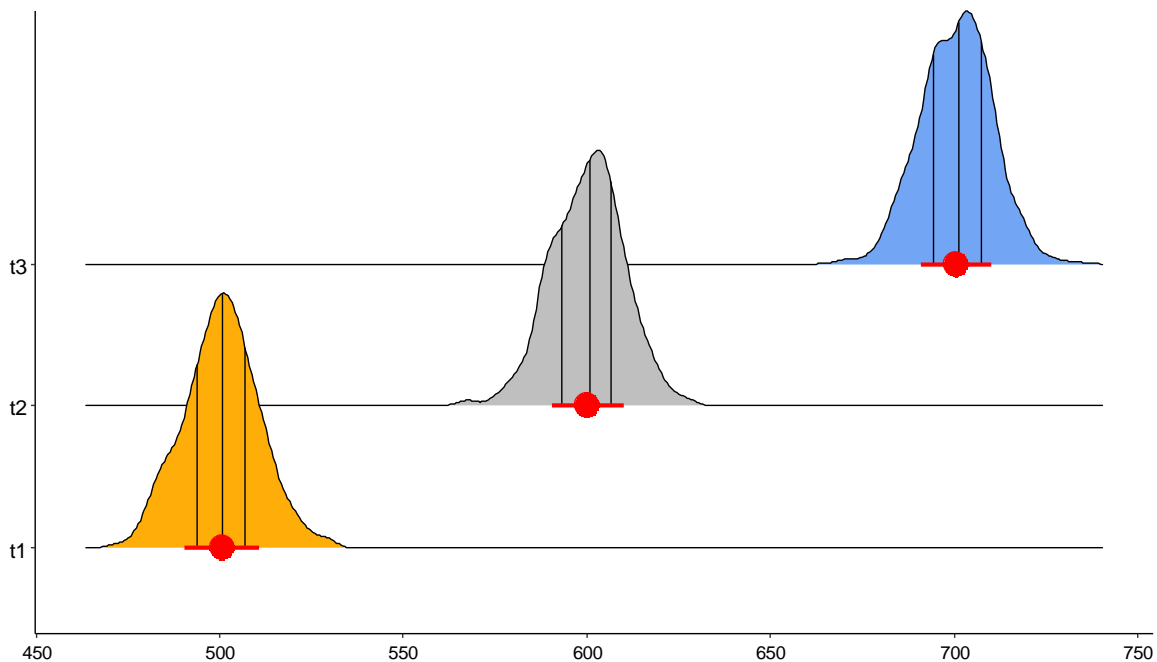
Correlação é a mesma entre todos os tempos

Variância similar ao longo do tempo

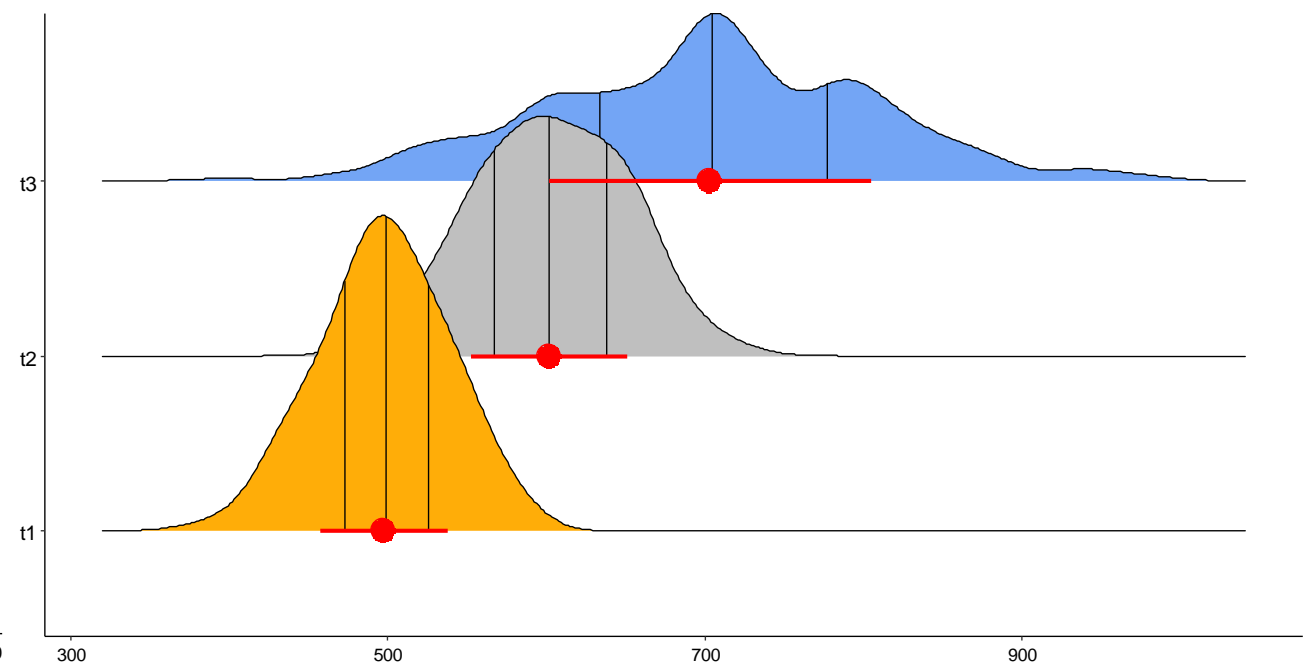


Esfericidade

Esférico :)



Não esférico :(



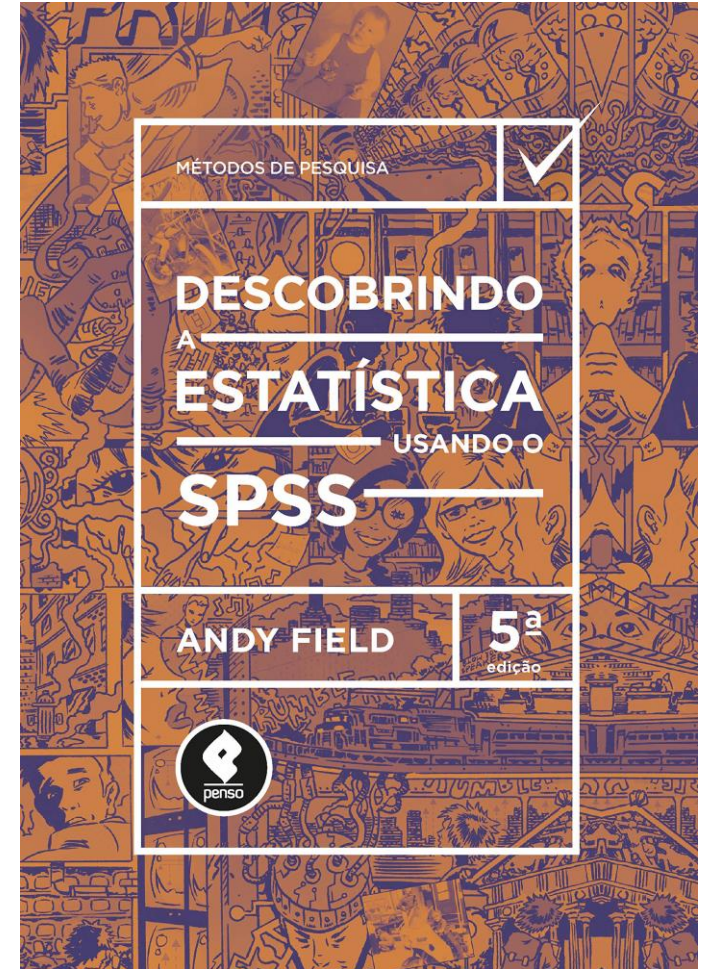
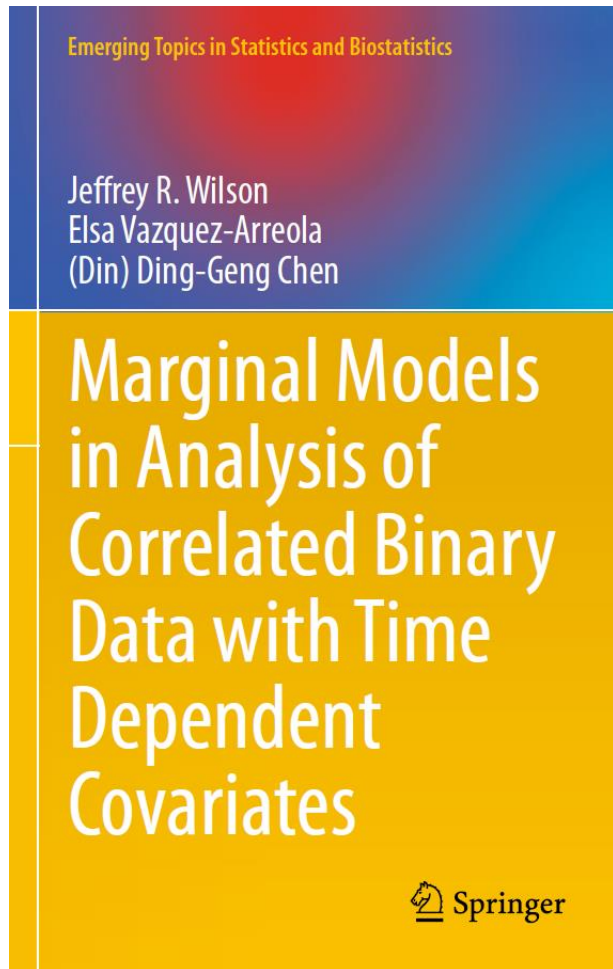
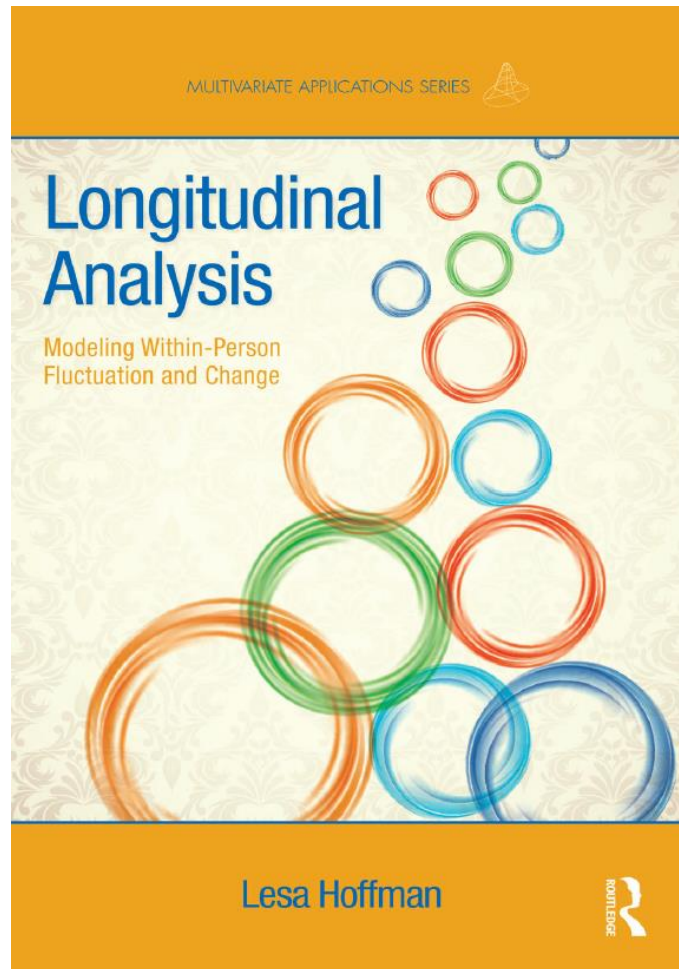
Esfericidade

- Em caso de violação do pressuposto, existem correções para a estatística F (não para os dados)
- Greenhouse–Geisser correction ($\hat{\epsilon}$)
- Greenhouse–Geisser: $\hat{\epsilon} < 0.75$
- Huynh–Feldt $\hat{\epsilon} > 0.75$

Post-hoc

Esfericidade OK	Esfericidade não OK
Tukey	Games-Howell
Bonferroni	Bonferroni

Indicações de livros



Recursos adicionais

- Podcast sobre invariância longitudinal:
<https://quantitudepod.org/s4e24-longitudinal-invar/>
- Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2016, May 23). Testing Measurement Invariance in Longitudinal Data With Ordered-Categorical Measures. Psychological Methods. Advance online publication.
<http://dx.doi.org/10.1037/met0000075>
- [https://github.com/Pedro-SR-Martins/GrupoEstudosEstatistica2023/blob/main/Inv
ariancia/longitudinal%20invariance%20example.R](https://github.com/Pedro-SR-Martins/GrupoEstudosEstatistica2023/blob/main/Inv%20ariancia/longitudinal%20invariance%20example.R)

