

Especialização em *Data Science* e Estatística Aplicada

Módulo IV - Métodos Não Paramétricos

Profa. Dra. Teresa Cristina Martins Dias

Goiânia, 2025



Conteúdo Programático

- Introdução aos métodos não paramétricos: apresentação e discussão do uso.
- Revisão de conceitos básicos: população e amostra, tipos de variáveis, escala de medidas e medidas descritivas, testes de hipóteses e p-valor.
- Testes de hipóteses não paramétricos baseados na distribuição binomial.
- Testes baseados na distribuição qui-quadrado.
- Teste para duas e mais amostras pareadas.
- Teste para duas e mais amostras independentes.
- Testes de dispersão.
- Aplicações em dados na área da saúde na linguagem **R**.

Aula 1 - Introdução

Para realizar uma análise estatística ou fazer inferência estatística, aplicamos técnicas, métodos ou testes estatísticos, sendo que nestes existem várias suposições para que os resultados obtidos sejam válidos, incluindo análise descritiva adequada ao tipo de variável.

Duas situações:

a) Estamos interessados em verificar se existe diferença entre médias de duas amostras, pareadas ou independentes, ou ainda se uma amostra foi retirada de uma população com determinada média

(verificar se a diferença entre as médias de duas amostras é estatisticamente significativa ou se a diferença é “casual”).

=> testes t-Student para comparação:

- para duas amostras independentes:
- para duas amostras dependentes (ou pareadas)
- para uma amostra (comparar com a média da população).

Suposições:

- independência entre as observações
- normalidade (ou n suficientemente)
- igualdade de variâncias (no caso de amostras independentes)
- variável quantitativa.

b) Estamos interessados em comparar várias amostras (três ou mais) para verificar se as médias são diferentes

(verificar se, pelo menos, um dos tratamentos produz resultados estatisticamente diferentes dos demais)

Suposições:

- variável quantitativa
- normalidade (ou n suficientemente grande)
- independência entre as observações
- igualdade de variâncias (amostras independentes)

=> ANOVA: análise de variância

... e se ...

- a(s) variável(is) for(em) qualitativa(s)?
- a(s) amostra(s) for(em) pequena(s)?
- a(s) amostras(s) não tem(tiverem) distribuição normal ou (aproximadamente normal)?



Métodos Não Paramétricos

- A estatística não paramétrica pode ser definida como uma coleção de métodos e testes para inferência estatística, nos quais não fazemos (muitas!) suposições sobre a (forma da) população da qual os dados foram coletados.
- Suposições
nos casos em que existem, estas são mais brandas.
- Em que situações pode ser aplicada?
os métodos e testes são usados em substituição aos testes paramétricos usuais em diversas situações (não normalidade dos dados, variáveis qualitativas, *etc.*).

Uso de testes não paramétricos

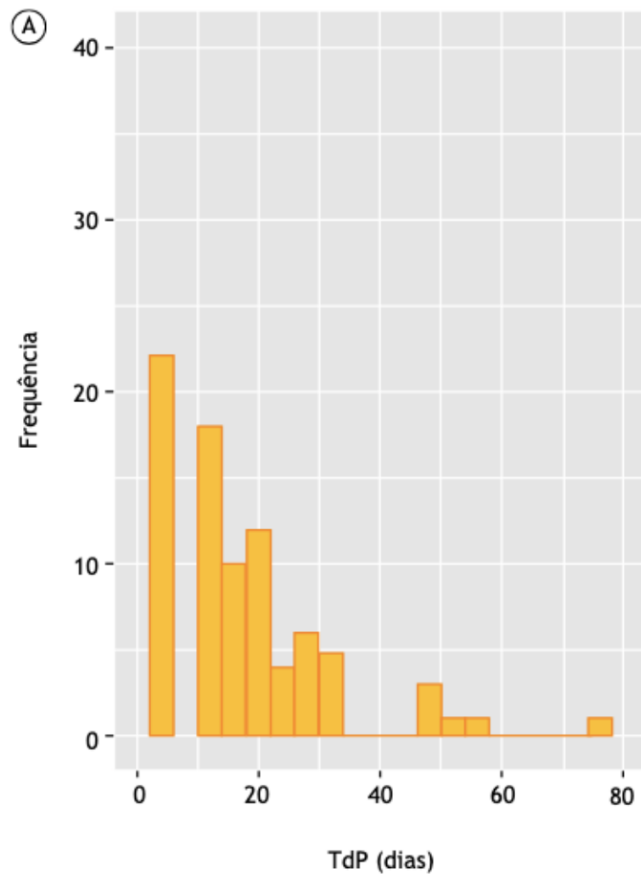
| Vantagens | Desvantagens |
|---|--|
| Dispensa normalidade e, são mais eficientes do que os paramétricos no caso de não normalidade | Desperdício de informação (geralmente não são levadas em consideração a magnitude dos dados) |
| Pode ser usado para variáveis qualitativas e, úteis no caso de ser difícil estabelecer uma escala quantitativa de valores para os dados | Se as suposições necessárias são atendidas, muitos dos testes paramétricos são conhecidos por serem os mais poderosos. Nos estudos de comparações (para alguns testes) os não paramétricos são frequentemente quase tão poderosos quanto os paramétricos, especialmente para amostras pequenas |
| Pode ser aplicado para amostras pequenas | |
| Testes mais simples | A utilização das tabelas dos testes é mais complicada |

Discussão:

A chefe de uma UTI gostaria de avaliar se pacientes obesos admitidos por exacerbação da doença pulmonar obstrutiva crônica (DPOC) têm um tempo de permanência (TdP) hospitalar mais longo do que pacientes não obesos.

Depois de selecionar **200** pacientes, ela descobriu que a distribuição do TdP é fortemente assimétrica a direita. Se ela fizesse um teste de hipótese, seria apropriado usar um teste t para comparar o TdP de pacientes obesos e não obesos com exacerbação de DPOC?

(Politi, M. T.; Ferreira, J. C.; Patino, C. M. "Testes estatísticos não paramétricos: mocinho ou bandido?". "Jornal Brasileiro de Pneumologia 47 (2021): e20210292).



Por fim...

Nesta situação, como a distribuição do tempo de permanência (TdP) hospitalar é fortemente assimétrica à direita, a relação entre obesidade e TdP em pacientes hospitalizados por exacerbações de DPOC deve ser analisada com um **teste não paramétrico** (em vez de aplicar o teste t-Student).

Alguns testes não paramétricos

| | |
|-------------------------------------|---|
| Uma amostra | para proporção, mediana, quantis, média (Wilcoxon) |
| Duas amostras dependentes | McNemar, do sinal, Wilcoxon e medidas de associação |
| Três ou mais amostras dependentes | Friedman |
| Duas amostras independentes | Mann-Whitney |
| Três ou mais amostras independentes | Kruskal-Wallis |