
Testes não-paramétricos

Mostre os dados usados nos cálculos e as expressões usadas nos testes (ou os comandos de pacotes/linguagens usados). Não deixe de comentar os resultados.

1. No exercício 1 foram dadas as idades das melhores atrizes e dos melhores atores premiados com Oscar. Uma das perguntas foi:
“Existe uma tendência das melhores atrizes serem mais jovens que os melhores atores?”
Na época, isto foi respondido com base em medidas de centro, como média e mediana. Agora pode ser respondido com testes de hipótese, mais especificamente com testes não paramétricos.
 - a) Selecione uma amostra (ao acaso) de tamanho apropriado das idades dos melhores atores e verifique a hipótese nula que a mediana da população dos atores é igual à 33.5, a mediana da idade das atrizes, contra a hipótese alternativa que é maior. Use o teste dos sinais e o valor-*p*. Comente o resultado.
 - b) Faça o mesmo, com os mesmos dados, mas usando o valor-*p* do teste dos postos sinalizados de Wilcoxon. Comente eventuais diferenças em relação ao resultado anterior.
 - c) Selecione duas amostras ao acaso, uma das idades das atrizes e outra das idades dos atores, e use o teste da soma dos postos para responder à pergunta do exercício 1 citada acima. Escreve as hipóteses nula e alternativa que está verificando, apresente e comente o resultado.
2. A planilha **Ex-08-dados** contém dados da UNICEF sobre mortalidade infantil: número de mortes por 1000 crianças de 1 a 4 anos. Nesta questão, considere os dados de 2018.
 - a) Selecione ao acaso 10 países da América do Sul e verifique a hipótese que a mediana do índice de mortalidade infantil apresentado é maior que 2 em 1000. Informe qual o teste feito, como foi feito e comente o resultado.
 - b) Faça o mesmo considerando países da Europa.
 - c) Com as amostras de tamanho 10 selecionadas, faça um teste para verificar a hipótese que a mortalidade é menor na Europa que na América do Sul. Informe os dados, o teste e o resultado.
 - d) Faça o mesmo que a letra (c), mas para América do Sul e África.
3. Verifique se houve redução no índice de mortalidade de 2010 para 2018, das seguintes formas:
 - a) Teste pareado: selecione ao acaso alguns países e colete seus dados nos anos de 2010 e 2018.
 - b) Teste não-pareado: selecione ao acaso vários dados de 2010 e outros de 2018. Certifique-se que a amostra seja aleatória e independente (sem escolher os países por conveniência).
4. Enuncie e verifique alguma outra hipótese que achar interessante.
5. Escolhe pelo menos dois dos testes de comparação de populações feitos acima e mostre o gráfico *boxplot* dos dados. Comente se o resultado visual do gráfico confirma o resultado do teste de hipóteses feito.
[Dicas na próxima página]

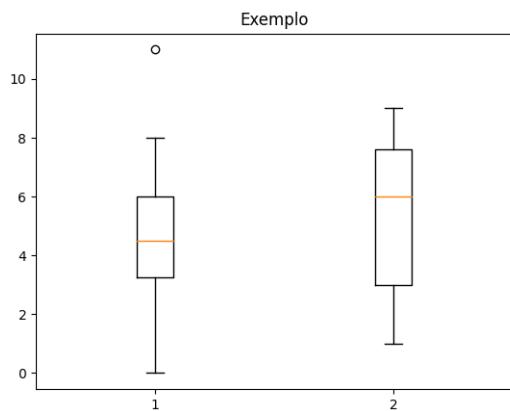
Em Python:

```
import matplotlib.pyplot as plt

dados1 = [6,3,3,4,5,11,6,2,8,4,5,4,6,0]
dados2 = [3,6,5,7.5,1,7.6,8,3,9]

plt.boxplot([dados1, dados2])
plt.title("Exemplo")
plt.show()
```

Obs.: haverá um botão para salvar o gráfico



On-line, free:

<http://shiny.chemgrid.org/boxplotr/>

Data Upload: digite, copie e cole de uma tabela, ou importe dados
Data Visualization: ver, editar e fazer download do gráfico

Obs.: os dados abaixo foram copiados e colados de uma planilha na caixa de entrada

About Data upload Data visualization

Paste data below:

d1	d2
1	6.00
2	3.00
3	6.00
3	5.00
4	7.50
5	1.00
6	11.00
7	7.60
8	6.00
9	8.00
10	3.00
11	9.00
12	4.00
13	2.00
14	5.00
15	4.00
16	0.00

Separator:

Comma

Tab

Semicolon

Clear data

A box plot generated by the shiny app. The x-axis has two categories labeled d1 and d2. The y-axis ranges from 0 to 12. Dataset d1 has a median of approximately 5, a box from 3 to 6, whiskers from 0 to 8, and one outlier at 10. Dataset d2 has a median of approximately 6, a box from 4 to 7, whiskers from 1 to 9, and one outlier at 10.