Teste para duas amostras pareadas

Em alguns estudos, o interesse está em comparar ou o efeito de dois grupos ou efeito de dois tratamentos: verificar a diferença em entre duas situações!

Alguns exemplos:

a) aplicação de herbicida (em plantas) para comparar o efeito do produto, registrando a quantidade de ervas daninhas, nas plantas sem e com herbicida. Existe diferença entre os tratamentos (**controle e tratamento** - herbicida)?

b) antes do sistema de cotas ser implantado, alunos de uma turma foram entrevistados sobre a implantação, se manifestando **a favor ou contra**. Após ser implantado, os mesmos alunos foram entrevistados e novamente foram interrogados sobre serem favoráveis ou não à implantação. **Houve mudança significativa de opinião**?

c) Vinte pessoas foram selecionadas para um "experimento": para cada uma destas foi registrado o **tempo** de reação a determinado evento ao dirigir um carro (frear, desviar, *etc*.), **antes e após** ingerir bebida alcoólica. O álcool afeta o tempo de reação das pessoas?

d) Homens e mulheres tem a mesma altura? Para verificar esta afirmação, **uma amostra da população feminina** foi observada e registradas as alturas das mulheres e **uma amostra da população masculina** foi retirada e suas alturas registradas. O que podemos concluir?

Quais as diferenças nestes exemplos?

Diferenças:

- temos dois tratamentos diferentes aplicados um em cada amostra (grupo) ou temos dois tratamentos aplicado na mesma amostra?
- qual o tipo de variável e escala de medidas?
- os grupos são independentes ou são pareados?
- ? ? ?

Características:

- caso a) duas amostras independentes (porém, com elementos parecidos)

- casos b) e c) a mesma amostra é utilizada duas vezes: o elemento é usado como seu próprio controle, submetendo-o aos dois tratamentos, em ocasiões diferentes. Este tipo é chamado de pré e pós tratamentos (antes e depois). Outro tipo de pareamento é o pareamento natural, como exemplo, gêmeos, órgãos (ouvidos, braços, pés, etc.), ou ainda quando selecionamos (ou tentamos selecionar) indivíduos muito semelhantes!

- caso d) amostras retiradas de populações independentes

Outros pontos importantes:
- caso a) variável quantitativa ou ainda, o registro pode ser uma resposta ordinal (menor/maior, pior/melhor, menos/mais)
- caso b) variável qualitativa
- casos c) e d): variável quantitativa

=> situações diferentes => testes diferentes!

Teste do sinal para dados ordinais

- as observações são emparelhadas, seguindo algum critério.

- útil nos casos em que a mensuração quantitativa não é possível de ser estabelecida, sendo possível estabelecer uma **relação** para cada membro do par. Pode ser aplicado para dados quantitativos

- **ideia do teste**: uma amostra é submetida a um tratamento e os resultados da variável são registrados. A mesma amostra é submetida a outro tratamento e os resultados da variável são registrados. Comparamos o efeito dos dois tratamentos, para cada par.

- podemos observar as seguintes situações: ou (a) o efeito aumentou; ou (b) o efeito diminuiu ou ainda, (c) o efeito permaneceu constante

- a cada par é atribuído um sinal (+) ou (-) (daí vem o nome do teste!), dependendo do resultado observado (uso apenas do sinal!)

- eficiente para pequenas amostras (ideal quando $n \leq 50$ e não há empates de postos)

- **teste binomial** no caso em que $p = 0,5$

- **teste usado para comparar evoluções avaliadas em uma escala ordinal**.

- **Dados**:
  - consiste de *n* observações de uma v.a. bivariada (X_1, Y_1), (X_2, Y_2), …, (X_n, Y_n).
  - os valores registrados, para cada par, são comparados e classificados da seguinte forma:

"+" , se $X_i < Y_i$; "-" , se $X_i > Y_i$ e "0" , se $X_i = Y_i$

- **Suposições**:
  - as v.as. bivariadas (X_i, Y_i) são mutuamente independentes
  - a escala de medidas é, pelo menos, ordinal dentro de cada par
  - os pares (X_i, Y_i) são internamente consistentes: se para um par $P(+) > P(-)$ então esta probabilidade vale para todos os pares; o mesmo ocorre para $P(+) < P(-)$ e para $P(+) = P(-)$.

- **Hipóteses**: (observando que, para $X < Y \Rightarrow$ "+"):

  a) H_0: $E(X_i) = E(Y_i)$ vs H_1: $E(X_i) \neq E(Y_i)$ (mesmo parâmetro de locação!)

  b) H_0: $E(X_i) \geq E(Y_i)$ vs H_1: $E(X_i) < E(Y_i)$ (os valores de X tendem a ser maiores do que de Y =>
  H_1: "+" é mais provável do que um "-")

  c) H_0: $E(X_i) \leq E(Y_i)$ vs H_1: $E(X_i) > E(Y_i)$, para todo i (os valores de X tendem a ser
  menores do que de Y).

- **Teste estatístico**:
  número de pares em que $X_i < Y_i$ => T = número de "+" .

- **Regra de decisão**:
  considerando apenas os ***n'*** pares em que $X_i \neq Y_i$, um nível α de significância e as hipóteses temos (usando a distribuição binomial com parâmetros n' e p = 0,5):

  a) rejeite H_0 se $T \leq t$ ou $T \geq n' - t$

  b) rejeite H_0 se $T \geq t$ (valores grande de T indicam H_1)

  c) rejeite H_0 se $T \leq t$ (valores pequenos de T indicam H_1).

Aplicação: **“Tower Building”** (Johnson e Courtney, Child Development, Vol. 3)

Usando blocos, os autores avaliaram os tempos (em segundos) que crianças pequenas levaram para construir uma torre (a mais alta possível). A fim de fornecer alguma medida de desenvolvimento mental, eles registraram os tempos para uma atividade inicial de construção de torres e, um mês depois, observando as mesmas crianças, registraram os tempos para uma segunda experiência de construção de torres:

- qual é a população de interesse?
- qual é a questão inerente de interesse aqui?
- quais as são amostras?
- qual é a variável aleatória avaliada?
- qual é o parâmetro populacional de interesse?
- qual é a estatística amostral correspondente a esse parâmetro?
- qual é uma declaração ou afirmação populacional apropriada?
- quais são as hipóteses estatísticas correspondentes?
- qual é a estatística teste usada para testar essas hipóteses?
- qual a decisão tomada e a conclusão, com base nos dados disponíveis?

- **definição de sucesso**: X > Y (evolução do experimento)

- **dados**: cada uma das crianças construiu duas torres (uma em cada momento) e, os tempos, em segundos, das construções foram registrados:

Crianças	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1ª tent.	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
2ª tent.	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15
sinais															

- **hipóteses**:
  H_0:
  H_1:

ou ainda, H_0: $E(X_i)$ **??** $E(Y_i)$ *vs* H_1: $E(X_i)$ **??** $E(Y_i)$

- T = **??**

O que podemos concluir? (resolvemos no R!)

Importante:

- Nos casos em que a variável (bivariada) é quantitativa (conhecemos a hierarquia e o tamanho da diferença entre os valores - (X,Y)):

  - as amostras não tem distribuição normal, são pequenas e pareadas => teste de Wilcoxon associado a postos

  - as amostras não são pareadas e são pequenas => teste de Mann-Whitney

  - se as amostras têm distribuição normal => teste t-Student.

Teste de McNemar

- variação do teste do sinal, sendo que as observações pertencem a uma **escala nominal** com duas categorias (0/1, N/S, etc.). As frequências das classes são conhecidas

- como verificar se existe diferença entre as duas categorias (**probabilidades** diferentes)?

- **ideia do teste**: uma amostra é submetida a um tratamento e o resultado é registrado (0/1). Em um segundo momento, essa mesma amostra é submetida a outro tratamento, e novamente os valores da variável são registrados (0/1).
  Geralmente, o tratamento é o mesmo porém, aplicado na amostra em dois momentos diferentes. O tratamento é aplicado a primeira vez, e após alguma intervenção, o mesmo tratamento é aplicado novamente

- o interesse é verificar se as **mudanças** (de respostas de uma categoria para a outra) **são significativas.**

- **Dados**:
  consistem de n observações independentes de uma v.a. bivariada (X_i,Y_i), $i = 1,..., n$, sendo classificados em uma entre
  duas categorias:

	Classificação dos Y's	
Classificação dos X's	0	1
0	a = número de pares (0,0)	b = número de pares (0,1)
1	c = número de pares (1,0)	d = número de pares (1,1)

- **Suposições**:
  - os pares (X_i, Y_i) são mutuamente independentes, e X representa o tratamento 1 ou "antes" e Y representa o tratamento 2 ou "depois" e Y ;
  - a escala de dados é nominal com duas categorias, para todo par (X_i, Y_i), $i = 1,..., n$;
  - a diferença $P(X_i = 0,Y_i = 1) - P(X_i = 1,Y_i = 0)$ é negativa, ou igual a zero ou é positiva, para todo i.

- **Hipóteses**:
  H_0: $P(X_i = 0,Y_i = 1) = P(X_i = 1,Y_i = 0)$, para todo i.
  H_1: $P(X_i = 0,Y_i = 1) \neq P(X_i = 1,Y_i = 0)$, para todo i.
  (H_0: declaramos que as mudanças que ocorreram nos dois sentidos não são significativas).

- **Teste estatístico:**

temos dois casos:

a) se $(b+c) > 20 \Rightarrow T_1 = (b-c)^2/(b+c)$

b) se $(b+c) \leq 20 \Rightarrow T_2 = b$

- **Regra de decisão**:

a) rejeite H_0 se $T_1 \geq w_{1-\alpha}$, obtida da distribuição qui-quadrado com 1 g.l.
b) rejeite H_0 se $T_2 \geq$ se $T_2 \leq t$ ou se $T_2 \geq n - t$, em que t é um valor obtido da distribuição binomial com parâmetros $n = b + c$ e, $p = 0{,}5$.

Aplicação: **Eficácia de ação educativa com reeducandas de cadeia pública de Mato Grosso sobre o vírus HPV**

(Corsino, P. K. D. *et al. Saúde e Pesquisa*, v. 11, n.1, pg 115-126, 2018)

- objetivo: analisar o impacto de ação educativa, sobre o vírus HPV, realizada com reeducandas de Cadeia Pública Feminina de Mato Grosso.

- estudo do tipo "antes" e depois"

- amostra: 37 mulheres (10/2016), em que

antes: explicação e sobre o tema da pesquisa e aplicação do questionário.

intervenção: duração de 30 minutos: atividade expositiva, utilizando cartazes com imagens relacionado ao HPV, modo de transmissão, diagnóstico, tratamento e prevenção, sendo possibilitado o moment reflexivo com as discussões e respostas aos questionamentos feitos pelas reeducandas.

após: após sete dias aplicou-se o mesmo questionário que precedeu a ação educativa.

Locais onde adquiriu conhecimento sobre HPV antes e após ação educativa realizada na Cadeia Pública Feminina.

Casa				Cadeia			
		Pós				Pós	
		S	N			S	N
Pré	S	2	5	Pré	S	2	1
	N	4	26		N	22	12

O que podemos concluir?

A realização da ação educativa contribuiu significativamente com a informação dessas mulheres sobre o HPV?

Vamos ao R!

Teste de Wilcoxon
(associado a postos)

Exemplo: “Tower Building”

É possível comparar, agora, os valores como quantidades e não apenas usando o sinal das diferenças

- teste aplicado para verificar se determinada amostra veio de uma população com **mediana** especificada (variável quantitativa)

- pode ser usado em situações nas quais as observações são **pareadas** (do tipo antes - depois), para verificar se as amostras têm mesma mediana

- as diferenças (entre observações do par e em módulo) são ordenadas e, **postos** são atribuídos a estas diferenças.

Observações:

a) **postos**: ordenação dos dados (do menor para o maior), atribuindo a “posição”

Exemplo

D= X-Y	-5	5	5	7	10	-13	13	15
|D|	5	5	5	7	10	13	13	15
Postos	2	2	2	4	5	6,5	6,5	8

- diferenças 5 e 13 **???**

b) **teste:**

- do sinal: aplicado nos caos em que apenas os sinais das diferenças são conhecidos
- de Wilcoxon: aplicado nos casos em que as diferenças, entre os valores dos tratamentos, são possíveis de serem obtidas (quantificar!).

- **Dados**:
  - consistem de *n* observações independentes de uma v.a. bivariada (Xi,Yi), i = 1,..., *n*.
  - defina as diferenças em módulo: |Di | = |Yi −Xi |, i=1,⋯,*n*.
  - seja ***n'*** o número de pares em que Xi ≠ Yi (o teste não considera pares nos quais Xi = Yi).
  - atribua postos, denotado por **Ri,** para as *n'* diferenças em módulo.

- **Suposições**:
  - Di's são mutuamente independentes;
  - Di's são v.as. contínuas com distribuição simétrica e tem a mesma mediana
  - a escala de medidas para os Di's é pelo menos intervalar.

- **Hipóteses**:

  a) H_0: $d_{.5} = 0$ vs H_1: $d_{.5} \neq 0$

  b) H_0: d.5 ≤ 0 vs H_1: d.5 > 0

  c) H_0: d.5 ≥ 0 vs H_1: d.5 < 0

- **Teste estatístico:**

  seja Ri o posto associado para cada par (se Xi ≠ Yi) definido por:

  se Di < 0 => Ri é o **negativo do posto** associado ao par (Xi, Yi)

  se Di > 0 => Ri é o **posto** associado ao par (Xi, Yi).

  O teste é dado por:

$$T = \frac{\sum_{i=1}^{n'} R_i}{\sqrt{\sum_{i=1}^{n'} R_i^2}}.$$

  Se não houver igualdade:

$$T^+ = \sum_{i=1}^{n'} R_i.$$

- **Regra de decisão:**

a) rejeite H_0, ao nível α de significância se $T > w_{1-\alpha/2}$ ou se $T < w_{\alpha/2}$

b) rejeite H_0 ao nível α de significância se $T < w_{\alpha}$

c) reject H_0 se $T > w_{1-\alpha}$

(w: tabela de quantis do teste de Wilcoxon)

Exemplo: construção das torres por 15 crianças

Crianças	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1ª tent.	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
2ª tent.	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15
diferenças															
postos															

Vamos para o R!