Comparando Duas Amostras **(abordagem não-paramétrica)**

Aula de hoje

Testes de Tendência Central (média, mediana, proporção)	Classificação	Variável 1	Variável 2	Número Grupos	Dependência	Premissas
Teste Z	Paramétrico	Quantitativa	-	1	-	Variância pop. * conhecida
Teste t	Paramétrico	Quantitativa	-	1	-	Distribuição normal
Wilcoxon (teste dos sinais, Wilcoxon p/ 1 amostra)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	-	1	-	
Teste t p/ 2 amostras	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias iguais
Teste t p/ 2 amostras com variâncias diferentes	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias diferentes
Teste t pareado	Paramétrico	Quantitativa	Nominal	2	Dependentes	Distribuição normal
ANOVA	Paramétrico	Quantitativa	Nominal	3 ou mais	Independentes	Distribuição normal; variâncias iguais
ANOVA p/ medidas repetidas	Paramétrico	Quantitativa	Nominal	3 ou mais	Dependentes	Distribuição normal; Esfericidade
Mann-Whitney (Wilcoxon não-pareado)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Independentes	
Wilcoxon (Wilcoxon pareado, teste dos sinais)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Dependentes	
Kruskal-Wallis	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Independentes	
Friedman	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Dependentes	
Teste p/ 1 proporção	Paramétrico	Nominal	-	1	-	
Teste p/ 2 proporções	Paramétrico	Nominal	Nominal	2	Independentes	

Estatística Não-Paramétrica

- Parâmetro: é medida usada para descrever uma característica de uma população
  - Ponto – estimativa por ponto (média, mediana, moda)
  - Intervalo – estimativa por intervalo (amplitude, desvio padrão, ...)
- Testes não-paramétricos: testes estatísticos "livres de distribuição", ou seja, não é necessário que se façam suposições sobre a distribuição de probabilidades das variáveis em estudo.

Diferenças entre testes paramétricos e não-paramétricos

- Testes paramétricos:
  - Baseados em parâmetros da amostra (média e desvio-padrão).
  - Fazem pressupostos sobre a distribuição dos dados (ex distribuição normal)
- Testes não-paramétricos:
  - Baseiam-se em postos (*ranks*) dos dados.
  - Pouco influenciados por valores extremos
  - Não fazem pressupostos sobre a distribuição dos dados

O que acontece se as pressuposições dos testes paramétricos não forem satisfeitas?

- Análise incorreta: falha na aderência às pressuposições
- Valor de *p* pode não ser correto
- Alguns testes são robustos: testes *t* e *t*-pareado (valor de *p* é pouco afetado se a distribuição dos dados apresenta pequenos desvios em relação à Normalidade)
- Transformações podem ser feitas (log, raiz quadrada, etc.) em uma tentativa de obter uma distribuição aproximadamente Normal

Vantagens e desvantagens

- Testes não paramétricos podem ser usados em situações especiais, quando os paramétricos não são apropriados:
  - Tamanho da amostra pequeno (≤5 ou 6): difícil estabelecer o tipo de distribuição
  - Pressuposições de distribuição dos testes paramétricos não são atendidas
  - Mensuração dos dados é ordinal ou nominal (variáveis qualitativas)

Vantagens e desvantagens

- **Poder dos testes**: testes não-paramétricos apresentam *poder* ligeiramente menor que seus correspondentes paramétricos. Ou seja, há casos em que o teste paramétrico rejeita a hipótese nula e o teste não-paramétrico não a rejeita.
- No entanto, quando a amostra é pequena (n < 30) e a distribuição dos dados não é Normal, o teste paramétrico deixa de ser confiável. Daí a preferência por testes não-paramétricos nestas condições.

Teste U de Mann-Whitney

- É o equivalente não-paramétrico do teste *t* não-pareado (para amostras independentes)
- Única suposição é de que a variável seja ordenável
- É possível utilizar com variáveis ordinais, como, por exemplo, uma escala de cruzes.
- É também conhecido como teste de somatória de postos de Wilcoxon

Teste U de Mann-Whitney

- Etapas:
  - Ordenar valores da variável em ordem crescente ou decrescente
  - Calcular os postos das observações
  - Calcular a soma dos postos de cada um dos grupos de tamanhos n_1 e n_2 (respectivamente, R_1 e R_2).

Exemplo 1

- Dezessete filhotes de cão foram treinados para defecar fora de casa desde o desmame até 6 semanas de idade através de:
  - Postura 1 – **Reforço positivo** (elogio quando animal defeca fora de casa) ($n_1 = 8$)
  - Postura 2 – **Reforço negativo** (castigo quando animal defeca dentro de casa) (n_2 =9)
- Mediu-se o tempo necessário para o estabelecimento do hábito (7 dias consecutivos sem defecar dentro de casa), em dias.*
- Pergunta-se: As duas posturas são igualmente efetivas?

*Petrie e Watson, Statistics for Veterinary and Animal Science, 1999

Exemplo 1

- H_0: Quanto ao tempo para adquirir o hábito, os 2 grupos **não são** estatisticamente diferentes (são iguais)
- H_1: Quanto ao tempo para adquirir o hábito, os 2 grupos **são** estatisticamente diferentes

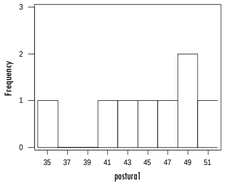
Em alguns textos de Estatística (e no Minitab):

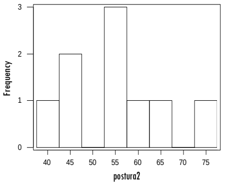
- H_0: mediana da postura 1 = mediana da postura 2
- H_1: mediana da postura 1 ≠ mediana da postura 2

1° PASSO: ordenar os valores segundo a variável de interesse.

Animal	Tempo (dias)	Postura
1	43	1
2	41	1
3	48	1
4	44	1
5	51	1
6	48	1
7	47	1
8	35	1
9	42	2
10	47	2
11	57	2
12	53	2
13	74	2
14	59	2
15	65	2
16	54	2
17	46	2

Animal	Tempo (dias)	Postura
8	35	1
2	41	1
9	42	2
1	43	1
4	44	1
17	46	2
10	47	2
7	47	1
3	48	1
6	48	1
5	51	1
12	53	2
16	54	2
11	57	2
14	59	2
15	65	2
13	74	2





2° PASSO: calcular os postos (atentando para as repetições) e calcular a soma dos postos de cada um dos dois grupos (R_1 e R_2).

Animal	Tempo (d)	Postura	Posto	Posto	Posto (P.1)	Posto (P.2)
8	35	1	1	1	1	
2	41	1	2	2	2	
9	42	2	3	3		3
1	43	1	4	4	4	
4	44	1	5	5	5	
17	46	2	6	6		6
10	47	2	7	7,5		7,5
7	47	1	8	7,5	7,5	
3	48	1	9	9,5	9,5	
6	48	1	10	9,5	9,5	
5	51	1	11	11	11	
12	53	2	12	12		12
16	54	2	13	13		13
11	57	2	14	14		14
14	59	2	15	15		15
15	65	2	16	16		16
13	74	2	17	17		17

$R_1=49{,}5$ $R_2=103{,}5$

Exemplo 1

- Comparação: há tabelas para distribuição das somas dos postos (vejam em referências).
- As somas dos postos são utilizadas para comparação com os valores da tabela.
- Programas de análise estatística fornecem o valor de *p* correspondente.

3° PASSO: Cálculo do valor de p (Minitab)

Exemplo 1

Mann-Whitney Test and CI: postura1; postura2

```
postura1   N =    8     Median =        45,50
postura2   N =    9     Median =        54,00
Point estimate for ETA1-ETA2 is       -10,00
95,1 Percent CI for ETA1-ETA2 is (-18,00;-2,00)
W = 49,5
Test of ETA1 = ETA2  vs  ETA1 not = ETA2 is significant at 0,0343
The test is significant at 0,0340 (adjusted for ties)
```

Notem que a estatística W corresponde a R_1

Exemplo 1

- Como o valor de p = 0,034 = 3,4%, para um nível de significância de 5%, rejeita-se H_0.
- Conclusão: há diferença estatística quanto ao tempo para se adquirir o hábito entre as posturas adotadas

Teste U de Mann-Whitney

Para n_1 e $n_2 \geq 10$ (ou ≥ 20 (Siegel, 1975)):

- Calculam-se

$$U_1 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$U_2 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

- Determina-se U como o menor entre U_1 e U_2
- Calcula-se a variável padronizada

$$z_U = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 + 1)}{12}}}$$

- Observa-se a tabela da distribuição Normal reduzida
- No Minitab, não é necessário se preocupar com esses passos

Teste U de Mann-Whitney

- Poder do teste:
  - Se comparado ao teste *t*: ~95% (Siegel, 1975)
  - Excelente alternativa ao teste *t*
  - Dispensa suposições restritivas e exigências inerentes ao teste *t*

Exemplo 2

- Cães semelhantes (com respeito a raça, peso e tamanho), apresentando escores de dor idênticos após serem submetidos ao mesmo tipo de cirurgia, foram alocados em 2 grupos: um grupo recebeu o analgésico A e outro o analgésico B. Alguns minutos após a medicação ter sido administrada, um médico veterinário fez uma avaliação do escore de dor. Há diferença, quanto ao escore de dor, entre os grupos?

Exemplo 2

- Hipóteses:
- H_0: Os dois grupos **não são** estatisticamente diferentes quanto ao escore de dor
- H_1: Os dois grupos **são** estatisticamente diferentes quanto ao escore de dor

Em alguns textos de Estatística (e no Minitab):

- H_0: mediana do grupo A = mediana do grupo B
- H_1: mediana do grupo A ≠ mediana do grupo B

Exemplo 2

Escores de dor em escala arbitrária

Animal	Analgésico A	Animal	Analgésico B
1	+++	10	+++
2	++	11	+
3	+	12	++
4	+++	13	++
5	+++	14	+
6	++	15	++
7	++	16	+
8	++		
9	+++		

Exemplo 2

Animal	Escore	Analgésico	Posto	Posto (empates)	Postos de A	Postos de B
3	+	A	1	2,5	2,5	
11	+	B	2	2,5		2,5
14	+	B	3	2,5		2,5
16	+	B	4	2,5		2,5
2	++	A	5	8	8	
6	++	A	6	8	8	
7	++	A	7	8	8	
8	++	A	8	8	8	
12	++	B	9	8		8
13	++	B	10	8		8
15	++	B	11	8		8
1	+++	A	12	14	14	
4	+++	A	13	14	14	
5	+++	A	14	14	14	
9	+++	A	15	14	14	
10	+++	B	16	14		14

R_A=90,5 R_B=45,5

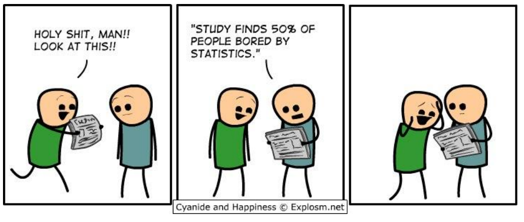
Exemplo 2

```
Mann-Whitney Test and CI: Vas1, Vas2

Vas1        N =   9     Median =       2,000
Vas2        N =   7     Median =       2,000
Point estimate for ETA1-ETA2 is        1,000
95.6 Percent CI for ETA1-ETA2 is (0,000, 1,000)
W = 90,5
Test of ETA1 = ETA2  vs  ETA1 not = ETA2 is significant at 0,1530
The test is significant at 0,1263 (adjusted for ties)

Cannot reject at alpha = 0,05
```

Com base nesses dados, não se pode rejeitar a hipótese nula. Deste modo, não foi observada diferença estatística significativa (p=0,1263) nos escores de dor dos dois grupos comparados, para um nível de significância de 5%.



Teste de Wilcoxon

- Alternativa não-paramétrica ao teste *t*-pareado
- Pode ser aplicado a qualquer variável ordenável (qualitativa ou quantitativa)

Teste de Wilcoxon

- Etapas:
  - Calcular, para cada indivíduo, a diferença entre o valor "antes" e o "depois" (d_i). Aqueles em que d_i=0 são excluídos da análise.
  - Ordenar os $|d_i|$ em ordem crescente ou decrescente e dar o valor dos postos. No caso de empate, tirar a média dos postos.
  - Calcular a soma dos postos dos $d_i > 0$ (T^+) e dos $d_i < 0$ (T^-) e verificar qual é a soma com sinal menos freqüente (T^*).

Exemplo 3

- Deseja-se saber se uma nova ração para suínos, em um galpão de terminação, promove ganho de peso. Os pesos dos animais foram medidos em diferentes momentos (dados hipotéticos):
  - Momento 0
  - Momento 1: 2 meses após a introdução da nova ração.
- Houve alteração significativa nos pesos dos suínos?

Exemplo 3

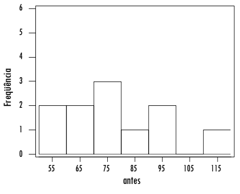
Hipóteses do teste:

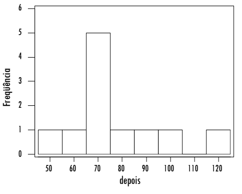
- H_0: Nova ração **não produz** efeito estatisticamente significativo
- H_1: Nova ração **produz** efeito estatisticamente significativo

Em alguns textos de Estatística (e no Minitab):

- H_0: mediana antes = mediana depois
- H_1: mediana antes ≠ mediana depois

Animal	Peso Antes (kg)	Peso Depois (kg)
1	55	48
2	72	68
3	95	73
4	88	89
5	74	72
6	66	66
7	59	58
8	97	119
9	112	99
10	69	73
11	75	80





Animal	Peso Antes (kg)	Peso Depois (kg)	d_i	$\lvert d_i \rvert$	Posto de d_i
1	55	48	– 7	7	7
2	72	68	– 4	4	4,5
3	95	73	– 22	22	9,5
4	88	89	+ 1	1	1,5
5	74	72	– 2	2	3
6	66	66	0	0	-
7	59	58	– 1	1	1,5
8	97	119	+ 22	22	9,5
9	112	99	– 13	13	8
10	69	73	+4	4	4,5
11	75	80	+5	5	6

*depois – antes

1º Passo: Calcular a diferença entre o depois e o antes (d_i);
2º Passo: Considerar o módulo de d_i;
3º Passo: Atribuir postos aos valores do 2º Passo, ignorando os zeros;
4º Passo: Calcular a soma dos postos cuja diferença foi positiva (T+) e cuja diferença foi negativa (T-). E a soma total dos postos (T).

T+ =21,5
T- =33,5
T=55

Exemplo 3

$T^+ = 21{,}5$

$T^- = 33{,}5$

Wilcoxon Signed Rank Test: dif

```
Test of median = 0,000000 versus median not = 0,000000

              N for   Wilcoxon            Estimated
         N    Test   Statistic        P      Median
dif     11      10        33,5    0,575       1,250
```

- Não se rejeita H_0
  - Conclusão: há animais ganhando e perdendo peso na mesma magnitude

Teste de Wilcoxon

Total de postos: $T = \dfrac{n \cdot (n+1)}{2}$

$$T^- = T - T^+ = \frac{n \cdot (n+1)}{2} - T^+$$

n: número de medidas onde se observou alguma mudança, isto é, $d_i \neq 0$

Teste de Wilcoxon

- Quando n ≥ 25:
  - $$z_W = \frac{T^* - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}}$$
  - T^* correspondente ao sinal menos frequente. Isto é, se $n^+ < n^-$, escolher T^+ ; se não, escolher T^-.
  - Tabela da Distribuição Normal Reduzida

Teste de Wilcoxon

- Teste de Wilcoxon: considera direção das diferenças
- Poder do teste para pequenas amostras: ~95% do teste *t*-pareado
- Dispensa suposições restritivas e exigências inerentes ao teste *t*

Referências

- A. Petrie e P. Watson, “Statistics for Veterinary and Animal Science”, Oxford, Blackwell, 1999.
- S. Siegel, “Estatística Não-Paramétrica”, São Paulo, McGraw-Hill, 1975.
- W. J. Conover, “Practical Nonparametric Statistics”, 3.ed., New York, Wiley, 1999.
- D. Salsburg, “Uma Senhora Toma Chá”, Rio de Janeiro, Zahar, 2009.