MATD49-Estatística não paramétrica

Testes para k amostras

Testes para k amostras independentes

- Em geral são extensões dos testes de 2 amostras;
- Úteis quando as suposições paramétricas não forem atendidas;
- Não exigem que as k amostras sejam de mesmo tamanho.
- Verificam se as k amostras foram extraídas ou não da mesma população ou de populações idênticas;
- Os dados geralmente seguem a seguinte estrutura:

Amostras ou Grupos			
1	**2**	$\ldots$	k
X_{11}	X_{21}	$\ldots$	X_{k1}
X_{12}	X_{22}	$\ldots$	X_{k2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
X_{1n_1}	X_{2n_2}	$\ldots$	X_{kn_k}

Teste de Kruskal-Wallis

Introdução

- É uma extensão do teste de Mann-Whitney (soma dos postos de Wilcoxon), i.e. para $k = 2$ populações, este teste é equivalente ao teste de Wilcoxon da soma de postos;
- É equivalente ao teste de Mantel-Haenszel aplicado aos postos dos dados;
- Pode ser interpretado como a versão não-paramétrica do teste F da ANOVA com 1 fator;
- Não leva em consideração formas específicas de distribuição;
- Deseja-se testar se k amostras aleatórias(possivelmente tamanhos diferentes) de uma v.a. possuem a mesma distribuição.

Pressupostos:

- As amostras são aleatórias e independentes entre si;
- A escala de mensuração é no mínimo ordinal.

Hipóteses do teste:

- H_0 : Todas as k populações têm funções de distribuição idênticas.
- H_1 : Pelo menos uma das populações difere das demais.

Construção do teste I

- Considere a $i-$ésima amostra aleatória de tamanho n_i: $X_{i1}, X_{i2}, \ldots, X_{in_i}$, $i = 1, 2, \ldots, k$.
- Seja N o número total de observações: $N = \sum_{i=1}^{k} n_i$.

Os dados podem ser organizados em colunas:

Amostra 1	**Amostra 2**	$\ldots$	**Amostra k**
X_{11}	X_{21}	$\ldots$	X_{k1}
X_{12}	X_{22}	$\ldots$	X_{k2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
X_{1n_1}	X_{2n_2}	$\ldots$	X_{kn_k}

Construção do teste II

- Atribua posto 1 à menor observação do total de N observações, posto 2 a segunda menor, e assim por diante, até a maior de todas as N observações, que recebe posto N (quando não há empates).
- Quando há observações iguais, calcular a média dos postos.
- Sejam :
  - R_{ij} o posto da observação de X_{ij}.
  - $R_{i\bullet}$ a soma dos postos da $i-$ésima amostra:

$$R_{i\bullet} = \sum_{j=1}^{n_i} R_{ij}, \ \ i = 1, 2, \ldots, k,$$

  - e $R_{\bullet\bullet}$ a média geral dos postos:

$$R_{\bullet\bullet} = \frac{\sum_{i=1}^{k} R_{i\bullet}}{N}. \tag{1}$$

- Se não há empates, teremos que a soma total de postos(numerador de (1)) é a soma de uma PA e portanto $R_{\bullet\bullet} = (N+1)/2$.

Estatística de teste I

A estatística de teste T é definida como:

$$T = \frac{1}{S^2}\left(\sum_{i=1}^{k}\frac{R_{i\bullet}^2}{n_i} - \frac{N(N+1)^2}{4}\right),$$

em que $S^2 = \frac{1}{N-1}\left(\sum_{i,j} R_{ij}^2 - \frac{N(N+1)^2}{4}\right).$

- Esta estatística mede a razão entre :
  - A soma diferenças quadráticas das médias dos tratamentos para a média geral (soma de quadrados entre tratamentos) e
  - As diferenças quadráticas dos postos em relação à média geral (quadrado médio total, variância);

Estatística de teste II

- Se não há empates, teremos que $S^2 = N(N+1)/12$ e a estatística de teste reduz a:

$$T = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_{i\bullet}^2}{n_i} - 3(N+1);$$

- A distribuição exata de T é tabelada (veja [Conover, 1996], Tabela A8) para $k = 3$ e $n_i \leq 5$ (no caso sem empates). No entanto, a distribuição exata é complexa e nos demais casos utiliza-se uma aproximação de T por a qui-quadrado com $k-1$ graus de liberdade;

- Rejeita H_0 ao nível α se T for maior que o quantil $1-\alpha$ (dos valores tabelados ou da χ_{k-1}^2).

Comparações múltiplas

- Se rejeitarmos a hipótese nula no teste de Kruskal-Wallis, é necessário realizar comparações múltiplas para detectar quais pares de populações podem ser considerados diferentes.
- As populações i e j são consideradas diferentes se a seguinte inequação é satisfeita:

$$\left|\frac{R_{i\bullet}}{n_i} - \frac{R_{j\bullet}}{n_j}\right| > t_{1-\alpha/2}\left(S^2\frac{N-1-T}{N-k}\right)^{1/2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)^{1/2},$$

  em que $R_{i\bullet}$ e $R_{j\bullet}$ são as somas dos postos das amostras i e j, respectivamente, $t_{1-\alpha/2}$ é o quantil $(1-\alpha/2)$ da distribuição $t-Student$ com $(N-k)$ graus de liberdade.
- que é equivalente ao intervalo:

$$\left[\left(\frac{R_{i\bullet}}{n_i} - \frac{R_{j\bullet}}{n_j}\right) \pm t_{1-\alpha/2}\left(S^2\frac{N-1-T}{N-k}\right)^{1/2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)^{1/2}\right]$$

  não conter o zero;

Aspecto computacional

No R: Pode ser realizado pelo pacote `kruskal.test` ou pelo pacote `PMCMRplus`

```
stats::kruskal.test(resposta,tratamento)
    kruskal.test(resposta~tratamento)
PMCMRplus::kruskalTest(resposta, tratamento, dist = "Chisquare")
      ::kruskalTest(resposta, tratamento)
# dist =  "KruskalWallis" p/ teste com a distr. exata.
### comparacoes multiplas ###
      kwAllPairsConoverTest(resposta~tratamento)
```

No SAS: Há duas formas para realizar o teste de Kruskal-Wallis. Ou com o comando para o teste da soma de postos de Wilcoxon (NPAR1WAY) ou através do teste de Mantel-Haenszel.

```
/*Abordagem 1: NPAR1WAY*/
PROC NPAR1WAY DATA=dados WILCOXON DSCF;
CLASS tratamento;
VAR resposta;
EXACT;
RUN;
/*Abordagem 2: via Mantel-Haenszel*/
PROC FREQ DATA=database;
TABLES tratamento*resposta /
CMH2 SCORES=RANK;
RUN;
```

Exemplo 1

Verificar a influência do fator *Idade* sobre a variável *tempo (em dias) para conseguir um emprego*, considerando as seguintes amostras:

Acima de 40 anos	Entre 25 e 40	Abaixo de 25
63	33	25
20	42	31
43	27	6
58	28	14
	51	18

Ao nível de 5% de significância, é possível afirmar que o fator *idade* tem influência sobre o *tempo para encontrar trabalho*?

Exemplo 2 - [Korosteleva, 2014]

Em uma competição de arco e flecha, 3 competidores disputam o primeiro lugar, que será definido estatisticamente. A pontuação por acerto pode ser 10 para o círculo menor, 5 a 9 para os círculos intermediários e 1 a 4 para os círculos mais externos. Cada competidor tem direito a 10 flechas. Decida se houve ganhador, ou se todos obtiveram o mesmo desempenho, com base nos dados abaixo.

Monica		Bob		Jeff	
Pontuação	Posto	Pontuação	Posto	Pontuação	Posto
3	6.5	2	4	1	1.5
4	10.5	2	4	1	1.5
4	10.5	3	6.5	2	4
5	16	4	10.5	4	10.5
5	16	4	10.5	4	10.5
5	16	5	16	5	16
10	22	10	22	10	22
10	22	10	22	10	22
				10	22

Exemplo 3 - [Beall, 1942]

Este exemplo contem dados de contagem de insetos sobreviventes após expostos a 6 tipos de inseticidas (A,B C, D, E, F) e deseja-se avaliar se existe diferença em suas eficácias. Os dados foram coletados por [Beall, 1942] e podem ser obtidos no pacote `datasets` do **R**: `datasets::InsectSprays`. Apenas ilustrativamente, a contagem, médias e desvios-padrão são apresentadas na tabela abaixo:

Inseticida	N	Média	Desv.Pad.
A	174	14.50	4.72
B	184	15.33	4.27
C	25	2.08	1.98
D	59	4.92	2.50
E	42	3.50	1.73
F	200	16.67	6.21

Decida ao nível 5% se há diferença entre os inseticidas.

Testes para k amostras relacionadas

Introdução I

Estrutura geral dos dados:

Amostras (Blocos)	Tratamentos			
	1	2	$\dots$	c
1	X_{11}	X_{12}	$\dots$	X_{1c}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\dots$
k	X_{k1}	X_{k2}	$\dots$	X_{kc}

Tabela 1: Estrutura geral dos dados nos testes de k amostras relacionadas.

- Outra forma é observar k unidades em c períodos de tempos. Neste caso teríamos um estudo de medidas repetidas, como acontece em avaliações do tipo *antes* e *depois* da intervenção na mesma unidade amostral;
- Estes testes são versões não-paramétricas da análise da variância com blocos completamente casualizados;

Introdução II

- Relembre que as seguintes condições devem ser satisfeitas para que possamos analisar um conjunto de dados utilizando análise de variância (ANOVA), também chamada de *teste F*:
  - As observações devem ser independentes e extraídas de populações normais;
  - As populações devem ter a mesma variância (homocedasticidade);
  - Os efeitos devem ser aditivos.
- Se estas condições forem satisfeitas, os testes não-paramétricos geralmente são uma alternativa viável para a análise, pois não exigem normalidade ou igualdade das variâncias e são aplicáveis a amostras de menores;
- Para amostras pequenas, mesmo que as condições do teste F sejam satisfeitas, sua aplicação não é recomendável;
- Falaremos do teste Q de Cochran e do teste de Friedman.

Teste Q de Cochran

Introdução

- Este teste é uma extensão do teste de McNemar para k amostras relacionadas, que chamaremos de *blocos*;
- Não avalia a extensão da mudança, apenas se a mudança ocorreu;
- Avalia se houve ou não sucesso na aplicação dos c tratamentos nos k blocos.

Pressupostos:

- A variável de interesse é dicotômica;
- As categorias das variáveis explicativas (blocos e tratamentos) são mutuamente excludentes;
- Cada um dos k tratamentos são aplicados independentemente para cada um dos c blocos;
- Em amostras muito pequenas não pode ser utilizado;
- Utiliza frequências e não os postos.

Estrutura geral dos dados

- Seja X_{ij} váriável aleatória dicotômica(1=*sucesso* ou 0=*fracasso*) representando o resultado para o tratamento j no grupo i. O valor $n = \sum_{i,j} X_{ij}$ representa o número total de 1's na tabela.
- Os resultados são apresentados em forma de tabela com k linhas, representando os blocos, e c colunas, representando os tratamentos:
  - $R_{i\bullet}$ é o total da linha i, $i = 1, 2, \ldots, k$;
  - $C_{\bullet j}$ é o total da coluna j, $j = 1, 2, \ldots, c$.
- Os dados podem então ser organizados como na tabela a seguir:

Amostras (Blocos)	**Tratamentos**				**Total**
	1	**2**	$\ldots$	c	
1	X_{11}	X_{12}	$\ldots$	X_{1c}	$R_{1\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
k	X_{k1}	X_{k2}	$\ldots$	X_{kc}	$R_{k\bullet}$
Totais	$C_{\bullet 1}$	$C_{\bullet 2}$	$\ldots$	$C_{\bullet c}$	n

Hipotéses do teste

Hipótese nula:

$H_0 : p_1 = p_2 = \ldots = p_c$, i.e. os tratamentos são igualmente eficazes.

Hipótese alternativa:

$H_1 : p_{j_1} \neq p_{j_2}$, para algum tratamento j_1 e j_2, i.e. ao menos 1 tratamento é diferente dos demais.

Estatística de teste e regra de decisão

A estatística de teste T é definida por:

$$Q = c(c-1)\frac{\sum_{j=1}^{c}\left(C_{\bullet j} - \frac{n}{c}\right)^2}{\sum_{i=1}^{k} R_{i\bullet}\,(c - R_{i\bullet})} = \frac{c(c-1)\sum_{j=1}^{c} C_{\bullet j}^2 - (c-1)n^2}{cn - \sum_{i=1}^{k} R_{i\bullet}^2}$$

Sob H_0, Q tem distribuição aproximadamente Qui-Quadrado com $c-1$ graus de liberdade.

Decisão do teste:

- Rejeita-se H_0 ao nível de significância α se o valor da estatística Q for maior do que o quantil $1-\alpha$ da distribuição Qui-Quadrado com $c-1$ graus de liberdade.
- Se H_0 é rejeitada, comparações entre os tratamentos podem ser feita usando o teste de McNemar.

Aspecto computacional

No R:

```
qcochran.test<-function(db.test, pcut=NA, alpha=0.05, OnlyPrint=TRUE)
#    db.test: data.frame com c colunas e (trats)
#             e  k linhas(blocos) .
#       pcut: Ponto de corte se db.test for composto por proporcoes.
#             Se especificado, o R categoriza os dados de db.test em
#             1 se maior ou igual pcut e 0 se menor que pcut
#      alpha: Nivel de significancia do teste.
# OnlyPrint: Se TRUE apenas imprime os resultados.
#             Se FALSE imprime e devolve uma lista com os resultados.
```

No SAS: a sintaxe é a mesma que o teste de McNemar.

```
PROC FREQ data = tabela_dados ;
TABLES trat1*trat2*trat3 /AGREE;
WEIGHT cont;
RUN;
```

Exemplo 1 - [Conover, 1996], adaptado

Três modelos preditivos (A, B e C) foram utilizados para prever os resultados de jogos de basquete colegial. Doze jogos foram selecionados de forma aleatória e o resultado da previsão de cada modelo é apresentado na tabela ao lado usando 1 para previsão correta (sucesso) e 0 para previsão errada (fracasso).

Verifique todos os modelos são igualmente eficazes na capacidade de prever os resultados dos jogos de basquete. Considere $\alpha = 5\%$.

Jogo	Modelo			Total
	A	B	C	
1	1	1	1	3
2	1	1	1	3
3	0	1	0	1
4	1	1	0	2
5	0	0	0	0
6	1	1	1	3
7	1	1	1	3
8	1	1	0	2
9	0	0	1	1
10	0	1	0	1
11	1	1	1	3
12	1	1	1	3
Total	8	10	7	25

Exemplo 2 - [Agresti, 2007]

46 pacientes foram tratados com 3 drogas diferentes para uma doença crônica e a resposta foi contabilizada como favorável(F) ou desfavorável (U).

Droga				Droga			
A	B	C	Freq	A	B	C	Freq
F	F	F	6	U	F	F	2
F	F	U	16	U	F	U	4
F	U	F	2	U	U	F	6
F	U	U	4	U	U	U	6

Avalie a eficácia dos tratamentos.

Teste de Friedman

Introdução I

- É uma extensão do teste dos sinais de Wilcoxon, logo, também utiliza os postos das observações. É também um caso especial do teste de Mantel-Haenszel geral, o qual vimos anteriormente o caso $k \times 2$;
- É útil para estudos de medidas repetidas ou delineamento em blocos;
- Neste tipo de estudo, observa-se o mesmo grupo de indivíduos sob cada um das k tratamentos, ou então formam-se conjuntos de indivíduos homogêneos que são alocados aleatoriamente a cada um dos tratamentos;
- É uma alternativa não paramétrica para a ANOVA com blocos casualizados;
- O teste examina os postos (*ranks*) dos dados em cada tratamento para determinar se as distribuições das variáveis são provenientes da mesma população.

Introdução II

Suposições:

- A variável de interesse é medida no mínimo em escala ordinal;
- Os blocos são independentes, i.e. a variabilidade dentro de um bloco não influencia os resultados de outro bloco.
- Em cada bloco (amostra), as observações podem ser ordenadas de acordo com algum critério de interesse.

Dados:

- Consistem de b vetores aleatórios independentes $k-$variados $(X_{i1}, X_{i2}, \ldots, X_{ik})$, chamados *blocos* (ou amostras), $i = 1, \ldots, b$.
- A variável aleatória X_{ij} representa a observação associada ao bloco i e ao tratamento j. Os dados podem ser organizados na mesma forma da Tabela 1 que foi ilustrada também para o teste de Cochran.

Construção do teste

- Atribuir postos de 1 a k para as observações do bloco i, $i = 1, 2, \ldots, b$, ditos R_{ij};
- Em caso de empates, atribuir a média dos postos;
- A soma dos postos para cada tratamento, $R_{\bullet j}$, é definida por:

$$R_{\bullet j} = \sum_{i=1}^{b} R_{ij}, \text{ para todo } j = 1, 2, \ldots, k$$

- Teremos assim uma nova tabela de postos:

Amostras (Blocos)	**Tratamentos**				**Postos**			
	1	**2**	$\ldots$	k	**1**	**2**	$\ldots$	k
1	X_{11}	X_{12}	$\ldots$	X_{1k}	R_{11}	R_{12}	$\ldots$	R_{1k}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
b	X_{b1}	X_{b2}	$\ldots$	X_{bk}	R_{b1}	R_{b2}	$\ldots$	R_{bk}
Total	-	-	-	-	$R_{\bullet 1}$	$R_{\bullet 2}$	$\ldots$	$R_{\bullet k}$

Hipóteses do teste

Hipótese nula:

H_0 : Os tratamentos produzem o mesmo efeito.

Hipótese alternativa:

H_1 : Pelo menos um dos tratamentos apresenta efeito diferente dos demais.

Estatística de teste I

- Diferente do teste de Kruskal-Wallis, aqui temos postos $1, \ldots, k$ para cada bloco. Logo, a soma total dos postos, se não há empates, será b vezes a soma de uma PA de k termos.
- A média geral dos postos será então:

$$R_{..} = \frac{1}{k} b \left(\frac{k(k+1)}{2} \right) = \frac{b(k+1)}{2};$$

- No caso sem empates, Friedman propôs a estatística do teste como:

$$\begin{aligned} T_1 &= \frac{12}{bk(k+1)} \sum_{j=1}^{k} \left(R_{.j} - \frac{b(k+1)}{2} \right)^2 \\ &= \frac{12}{bk(k+1)} \sum_{j=1}^{k} R_{.j}^2 - 3b(k+1), \end{aligned}$$

Note que, assim como no teste de Kruskal-Wallis, esta estatística também está relacionada com a soma das diferenças quadráticas das médias dos tratamentos para a média geral.

Estatística de teste II

- Se há empates, um ajustamento na estatística T_1 precisa ser feito:
  - Seja A_1 a soma dos quadrados dos postos: $A_1 = \sum_{i=1}^{b}\sum_{j=1}^{k}[R_{ij}]^2$;
  - Calcule o fator de correção dado por: $C_1 = bk(k+1)^2/4$;
- Então, a estatística T_1, modificada na presença de empates, é:

$$
\begin{aligned}
T_1^* &= \frac{(k-1)\left(\sum_{j=1}^{k} R_{\bullet j}^2 - bC_1\right)}{A_1 - C_1} \\
&= \frac{(k-1)\sum_{j=1}^{k}\left(R_{\bullet j} - \frac{b(k+1)}{2}\right)^2}{A_1 - C_1};
\end{aligned}
$$

  - A distribuição exata de T_1 e T_1^* é difícil ser encontrada e uma aproximação é comumente usada;
  - T_1 e T_1^* têm, sob H_0 distr. aprox. Qui-quadrado com $(k-1)$ graus de liberdade;

Estatística de teste III

- Uma estatística alternativa é a estatística dos experimentos em blocos ao acaso na ANOVA calculada sobre os postos $R(X_{ij})$:

$$T_2 = \frac{(b-1)T_1}{b(k-1)-T_1};$$

- A estatística T_2 tem, aproximadamente, distribuição F com $k_1 = (k-1)$ e $k_2 = (b-1)(k-1)$ graus de liberdade.

Decisão do teste:

- Fixado α, rejeitamos H_0 se T_1 exceder o quantil $(1-\alpha)$ da distribuição Qui-quadrado com $(k-1)$ graus de liberdade.
- De maneira similar, rejeitamos H_0 ao nível de significância α se T_2 exceder o quantil $(1-\alpha)$ da distribuição F com $k_1 = (k-1)$ e $k_2 = (b-1)(k-1)$ graus de liberdade.

Comparações múltiplas

- Ao rejeitarmos a hipótese nula no teste de Friedman, faremos comparações múltiplas para detectar quais pares de tratamentos podem ser considerados diferentes.
- Considere dois tratamentos i e j, para todo $i \neq j$ e $i, j = 1 \ldots, k$.
- Os tratamentos i e j são considerados diferentes se

$$|R_{\bullet j} - R_{\bullet i}| > t_{1-\alpha/2} \left[\frac{2(bA_1 - \sum R_{\bullet j}^2)}{(b-1)(k-1)} \right]^{1/2},$$

em que $t_{1-\alpha/2}$ é o quantil $(1-\alpha/2)$ da distribuição $t - Student$ com $(b-1)(k-1)$ graus de liberdade.

- Alternativamente, a equação $|R_{\bullet j} - R_{\bullet i}|$ pode ser expressa com uma função de T_1:

$$|R_{\bullet j} - R_{\bullet i}| > t_{1-\alpha/2} \left[\frac{(A_1 - C_1)2b}{(b-1)(k-1)} \left(1 - \frac{T_1}{b(k-1)} \right) \right]^{1/2}.$$

Aspecto computacional

No R:

```
friedman.test(resp, trat, bloco)
friedman.test(resp ~ trat | bloco, data)
```

No SAS: Há duas formas de se realizar o teste de Friedman: utilizando o teste de *Cochran-Mantel-Haenszel* para os escores, já que os testes são equivalentes, ou utilizando a estatística T_2, isto é, construindo uma ANOVA para os postos:

```
/*Abordagem 1: via CHM*/
PROC FREQ DATA=database;
TABLES bloco*trat*resp /
        CMH2 SCORES=RANK;
RUN;
/*Abordagem 2: ANOVA DOS POSTOS:*/
PROC RANK;
BY bloco;
VAR resp;
RANKS varrank;
RUN;
PROC ANOVA;
CLASS bloco trat;
MODEL varrank = bloco trat;
RUN;
```

Exemplo 1

Um teste de consumo de combustível envolvendo carros produzidos por três fabricantes foi realizado e os resultados, em quilômetros por litro de combustível estão apresentados na tabela abaixo. Estabelecer e testar a hipótese adequada. Considere $\alpha = 5\%$.

Modelo	Fabricante		
	G	F	C
Pequeno	9.0	11.3	10.6
Médio- 6 cil.	9.4	10.9	10.2
Médio- 8 cil.	8.1	8.6	9.1
Grande-8 cil.	8.3	8.6	8.8
Esporte	8.2	9.2	9.5

Exemplo 2 - [Lehmann and D'abrera, 2006]

Num estudo sobre hipnose, 8 sujeitos tiveram a tensão elétrica na superfície da pele medida (em milivolts) em 4 situações emocionais distintas: *medo*, *alegria*,*tristeza* e *calma*. Avalie se existe diferença na tensão entre os diferentes estados emocionais.

Sujeito	Medo	Alegria	Tristeza	Calma
1	23.1	22.7	22.5	22.6
2	57.6	53.2	53.7	53.1
3	10.5	9.7	10.8	8.3
4	23.6	19.6	21.1	21.6
5	11.9	13.8	13.7	13.3
6	54.6	47.1	39.2	37.0
7	21.0	13.6	13.7	14.8
8	20.3	23.6	16.3	14.8

Exemplo 3 - Velocidades de atletas

6 atletas de ciclismo tiveram suas velocidades médias calculadas ao longo de 4 trechos de uma prova. Avalie se algum atleta se destacou dos demais.

	Trecho			
Atleta	A	B	C	D
1	32.60	36.40	29.50	29.40
2	42.70	47.10	32.90	40.00
3	35.30	40.10	33.60	35.00
4	35.20	40.30	35.70	40.00
5	33.20	34.30	33.20	34.00
6	33.10	34.40	33.10	34.10

Acknowledgements

Agradecemos ao prof. Anderson Ara pela disponibilização de seu material didático, no qual nos baseamos para a elaboração destes slides. Alguns trechos desta apresentação são replicados de seu material.

Referências I

Agresti, A. (2007).
An introduction to categorical data analysis.
Wiley-Interscience.

Beall, G. (1942).
The transformation of data from entomological field experiments so that the analysis of variance becomes applicable.
Biometrika, 32:243.

Conover, W. J. (1996).
Practical nonparametric statistics.
John Wiley and sons, 3 ed. edition.

Korosteleva, O. (2014).
Nonparametric methods in statistics with SAS applications.
Crc Press.

Referências II

Lehmann, E. L. and D'abrera, H. J. M. (2006).
Nonparametrics : statistical methods based on ranks.
Springer.