ANOVA Não Paramétrica

Anderson Rodrigo da Silva

Instituto Federal Goiano

Conteúdo

Estatística não paramétrica

- Em geral os métodos são simples e versáteis
- Há pouca ou nenhuma especificação ou exigência acerca dos parâmetros da população da qual a amostra foi obtida
- Dispensam, por exemplo: normalidade, homocedasticidade
- São especialmente úteis em situações em que é difícil estabelecer uma escala quantitativa (variáveis ordinais)
- São em geral mais poderosos que os paramétricos quando as exigências do modelo não são atendidas

Restrições

- Em geral não levam em conta a magnitude dos dados
- Em geral não nos permite testar interações
- Requerem, frequentemente, o uso de tabelas um pouco mais complexas

ANOVA não paramétrica

Quando as pressuposições do modelo de ANOVA (normalidade, homocedasticidade) não são verificadas, ou quando há *outliers*, a hipótese de igualdade de tratamentos pode ainda ser avaliada por meio dos seguintes testes não paramétricos:

- Teste de Kruskal-Wallis (DIC)
- Teste de Friedman (DBC)

Teste de Kruskal-Wallis

Os centros (medianas, $\tilde{\mu}$) de de k amostras independentes (tratamentos) podem ser comparados por meio do teste de Kruskal-Wallis.
As seguintes hipóteses são avaliadas:

$$H_0 : \tilde{\mu}_1 = \tilde{\mu}_2 = ... = \tilde{\mu}_k$$

contra a hipótese H_1 de que ao menos duas amostras diferem entre si

Método

- Considere k amostras, cada uma com n_i observações. Temos que $N = \sum_{i=1}^{k} n_i$
- Atribua ranks de 1 a N, conjuntamente, às observações. Havendo impates, calcule um rank médio.
- Calcule a estatística:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

em que R_i é a soma dos ranks da i-ésima amostra.
Para $k \geq 3$ e $n_i \geq 6$, a seguinte aproximação é feita:

$$H \sim \chi^2_{k-1}$$

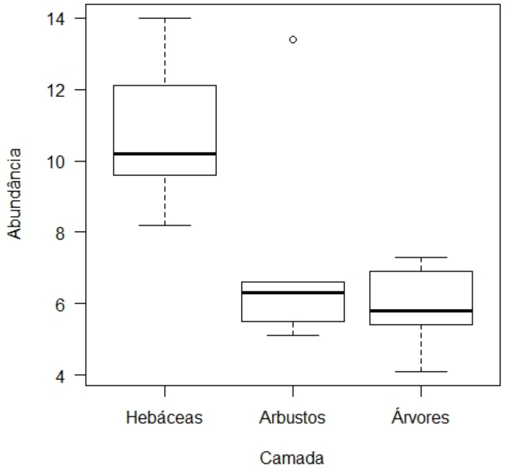
Obs.: para $k = 2$ o teste de Wilcoxon deve ser aplicado; para $k = 3$, tabelas específicas devem ser utilizadas (ver Campos, 1983)

Exemplo

(Adaptado de Zar, 1999). Um entomologista está estudando a distribuição vertical de uma espécie de mosca em uma floresta decídua e obtém cinco amostras de moscas em três diferentes camadas de vegetação: herbáceas, arbustos e árvores. O entomologista deseja avaliar a seguinte hipótese H_0: A abundância de moscas é a mesma nas três camadas de vegetação.

Herbáceas	Arbustos	Árvores
14	13.4	6.9
12.1	5.1	7.3
9.6	5.5	5.8
8.2	6.6	4.1
10.2	6.3	5.4

Exemplo



Teste de Friedman

- O teste de Friedman é análogo ao teste F da ANOVA para dados provenientes de um DBC.
- Além dos tratamentos, há mais um fator de classificação (blocos)
- As mesmas hipóteses apresentadas no teste de Kruskal-Wallis são avaliadas
- O procedimento consiste em atribuir ranks às observações de cada bloco, separadamente.

Teste de Friedman

Estatística de teste:

$$X^2 = \frac{12}{bk(k+1)} \sum_{i=1}^{k} R_i^2 - 3b(k+1)$$

em que b é o número de blocos, k é o número de tratamentos, R_i é a soma dos ranks do i-ésimo tratamento
H_0 é rejeitada quando: $X^2 \geq \chi^2_{k-1,\ 1-\alpha}$

Examplo

Considere um experimento delineado para avaliar o efeito do fotoperíodo sobre o crescimento (mm) de plantas de seis genótipos diferentes.

	Fotoperíodo			
Genótipo	Muito curto	Curto	Longo	Muito longo
A	2	3	3	4
B	3	4	5	6
C	1	2	1	2
D	1	1	2	2
E	2	2	2	2
F	1	1	2	3

Comparações múltiplas de tratamentos - teste de Kruskal-Wallis

A hipótese $H_0 : \tilde{\mu}_i = \tilde{\mu}_j$ de que dois tratamentos são iguais, será rejeitada quando

$$|\bar{R}_i - \bar{R}_j| \geq t_{N-k} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

sendo $\bar{R}_i$ o rank médio do i-ésimo tratamento, t_{N-k} é o valor da tabela de t-Student ao nível $1 - \alpha/2$

Comparações múltiplas de tratamentos - teste de Friedman

A hipótese $H_0 : \tilde{\mu}_i = \tilde{\mu}_j$ de que dois tratamentos são iguais, será rejeitada quando

$$|R_i - R_j| \geq t_{(k-1)(b-1)} \sqrt{\frac{bk(k+1)}{12}}$$

sendo R_i a soma dos ranks do i-ésimo tratamento, $t_{(k-1)(b-1)}$ é o valor da tabela de t-Student ao nível $1 - \alpha/2$