

Testes de Comparação Múltipla:

Bonferroni, Duncan e Dunnet

29 de Novembro de 2016

André Felipe B. Menezes
Victor Hugo Nagahama

Departamento de Estatística
Universidade Estadual de Maringá – UEM





Organização

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências



Introdução

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

2

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

Estrutura da Apresentação

1. Explicação do teste;
2. Regra de decisão;
3. Código do SAS;
4. Aplicação do teste.



Introdução

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

3

Banco de dados

Como exemplificação utilizamos o banco de dados *Cholesterol* que possui 50 observações de duas variáveis observadas em um estudo clínico cujo objetivo foi avaliar o efeito de três formulações da mesma droga na redução do colesterol. Foi utilizado também duas drogas concorrentes como grupo controle (drugD e drugE).

- ▶ *Response*: redução do colesterol (mg/dl).
- ▶ *Trt*: tipo/dosagem da droga.



Banco de dados

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

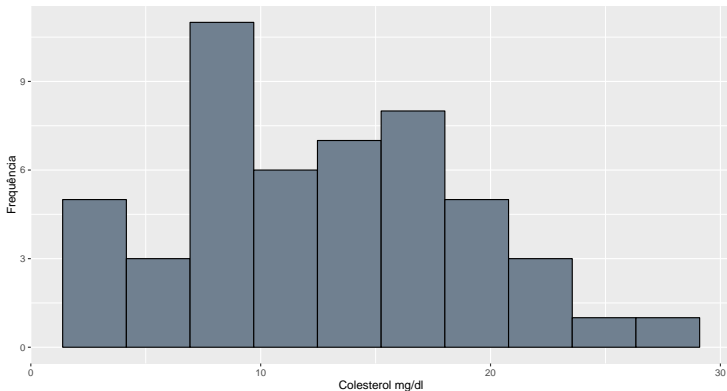
Teste de Duncan

Teste de Dunnett

Recomendações

Referências

4





Banco de dados

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

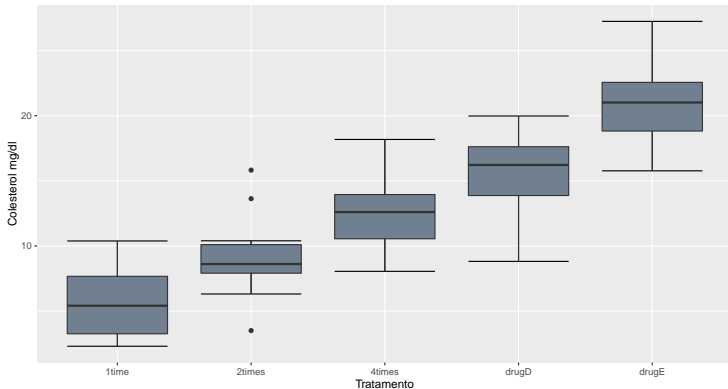
Teste de Duncan

Teste de Dunnett

Recomendações

Referências

5





Banco de dados

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

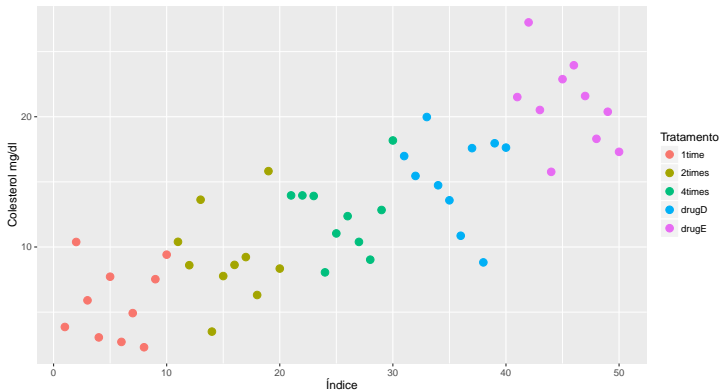
Teste de Duncan

Teste de Dunnett

Recomendações

Referências

6





Introdução

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

7

Modelo do Delineamento Inteiramente Casualizado

Um modelo linear pode ser expresso por:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij},$$

com $i = 1, 2, \dots, k$ e $j = 1, 2, \dots, n_i$. Sendo:

- ▶ Y_{ij} : variável resposta, isto é, j -ésima repetição do i -ésimo tratamento;
- ▶ μ : média global;
- ▶ τ_i : efeito do i -ésimo tratamento;
- ▶ ε_{ij} : erro residual da j -ésima repetição do i -ésimo tratamento.



Introdução

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

8

Pressupostos

As k variáveis aleatórias Y_{ij} são:

- ▶ Mutuamente independentes.
- ▶ Provenientes de uma distribuição Normal.
- ▶ E possuem a mesma variância (homoscedasticidade).



Hipóteses

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

9

Sem tratamento controle

Realiza-se a comparação entre todos os pares de médias, isto é, todos os tratamentos. Logo temos as seguintes hipóteses:

$$\begin{cases} H_0 : \mu_i = \mu_j \\ H_1 : \mu_i \neq \mu_j \end{cases} \quad (1)$$

para $i \neq j$. O procedimento é repetido para todas as $\frac{k(k-1)}{2}$ comparações.



Hipóteses

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Banco de dados

Hipóteses

Erro Tipo I

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

10

Tratamento controle

As comparações são realizadas entre a média do tratamento controle e todas as outras médias. Assim as hipóteses são:

$$\begin{cases} H_0 : \mu_i = \mu_a \\ H_1 : \mu_i \neq \mu_a \end{cases} \quad (2)$$

O procedimento é repetido para todas as $k - 1$ comparações.



Erro Tipo I

- ▶ **Erro Tipo I por comparação (CER):** Probabilidade de ocorrer um Erro Tipo I em qualquer comparação, denotaremos por α .
- ▶ **Erro Tipo I por experimento (EER):** Probabilidade de um ou mais Erro Tipo I em um conjunto (família) de comparações, denotaremos por α_f
- ▶ A relação entre estas taxas de erros quando os testes são independentes é dada por:

$$\alpha_f = 1 - (1 - \alpha)^c \quad (3)$$

em que c é número de comparação no experimento.



Teste de Bonferroni

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

12

Explicação

Consiste na realização de um teste t para cada par de médias, sendo que a inequação de Bonferroni é utilizada para controlar **CER**, isto é,

$$\mathbf{CER} = \frac{\alpha}{c} \quad (4)$$

em que c é o número de testes realizados, então o **EER** é no máximo α .



Teste de Bonferroni

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

13

Regra de Decisão

Dado o interesse em testar as hipóteses definidas em (1), o teste de Bonferroni com **EER** $< \alpha$, declara duas médias significativamente diferentes se

$$|\bar{y}_{i.} - \bar{y}_{j.}| \geq t(\epsilon; N - k) \sqrt{\frac{1}{2} \text{QME} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}. \quad (5)$$

em que

$$\epsilon = \frac{2\alpha}{k(k-1)}. \quad (6)$$



Teste de Bonferroni

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

14

Código no SAS

O procedimento GLM possui a opção BON no *statement* MEANS.
Ou ainda, a opção ADJUST=BON no *statement* LSMEANS.

```
proc glm data=dados;  
    class    trat;  
    model    resp = trat;  
    means    trat / bon;  
    lsmeans  trat / adjust=bon;  
  
run;
```



Teste de Bonferroni

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Aplicação

Introdução

Teste de Bonferroni

15

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

Comparison	Difference	LowerCL	UpperCL	Significance
drugE – drugD	5.587	1.326	9.848	***
drugE – 4time	8.574	4.313	12.835	***
drugE – 2time	11.722	7.461	15.983	***
drugE – 1time	15.167	10.906	19.428	***
drugD – drugE	–5.587	–9.848	–1.326	***
drugD – 4time	2.987	–1.274	7.248	0
drugD – 2time	6.135	1.874	10.396	***
drugD – 1time	9.580	5.319	13.841	***
4time – drugE	–8.574	–12.835	–4.313	***
4time – drugD	–2.987	–7.248	1.274	0
4time – 2time	3.148	–1.113	7.409	0
4time – 1time	6.593	2.332	10.854	***
2time – drugE	–11.722	–15.983	–7.461	***
2time – drugD	–6.135	–10.396	–1.874	***
2time – 4time	–3.148	–7.409	1.113	0
2time – 1time	3.445	–0.816	7.706	0
1time – drugE	–15.167	–19.428	–10.906	***
1time – drugD	–9.580	–13.841	–5.319	***
1time – 4time	–6.593	–10.854	–2.332	***
1time – 2time	–3.445	–7.706	0.816	0



Teste de Bonferroni - Aplicação

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

16

Forest plot

Differences and Adjusted 95% Confidence Intervals

1time vs 2times, $p = 0.213$

1time vs 3times, $p < .001$

1time vs drugD, $p < .001$

1time vs drugE, $p < .001$

2times vs 3times, $p = 0.344$

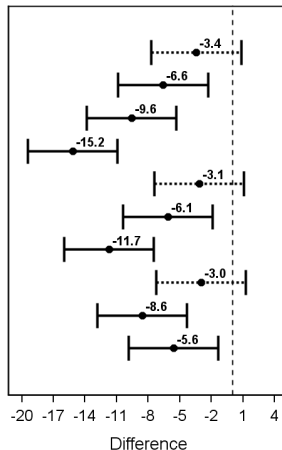
2times vs drugD, $p = 0.001$

2times vs drugE, $p < .001$

3times vs drugD, $p = 0.443$

3times vs drugE, $p < .001$

drugD vs drugE, $p = 0.003$





Teste de Bonferroni - Aplicação

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

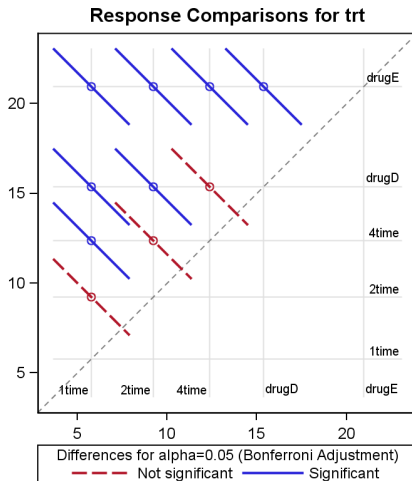
Teste de Dunnett

Recomendações

Referências

17

Difograma





Teste de Bonferroni

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

18

Discussão

- ▶ Controla **EER** em no máximo α .
- ▶ Pode ser utilizado para testar contrastes ou outros testes de hipóteses simultâneos.
- ▶ Possui menor poder comparado a outros procedimentos.
- ▶ Recomendado para testar contrastes planejados (HOWELL, 2010).



Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

19

Explicação

- ▶ Proposto em 1955 a partir de uma modificação no teste de Newman-Keuls o teste de Duncan é considerado um teste de múltiplos estágios.
- ▶ Testes de múltiplos estágios avaliam as hipóteses, (1), em etapas determinadas conforme a amplitude das médias dos tratamentos.



Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

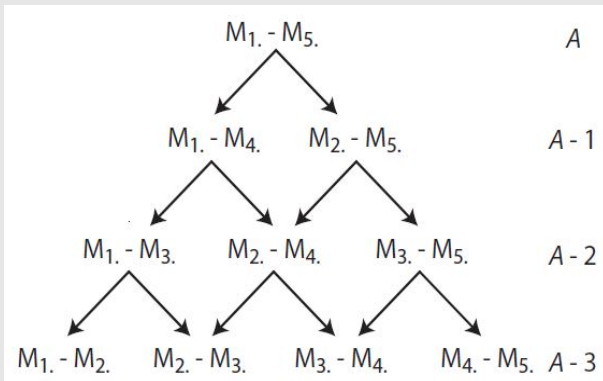
Teste de Dunnett

Recomendações

Referências

20

Explicação





Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

21

Regra de Decisão

Em nível de significância α , rejeita-se a hipótese nula, (1), se

$$|\bar{y}_i. - \bar{y}_j.| > Q(\gamma_p, p, N - k) \cdot \sqrt{\frac{1}{2} \text{QME} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \quad (7)$$

Em que:

- ▶ $p = 2, 3, \dots, k$.
- ▶ $\gamma_p = 1 - (1 - \alpha)^{p-1}$.
- ▶ $Q(\gamma_p, k, N - k)$ é o $100 \times \gamma_p$ percentil da distribuição Studentized Range com alcance k e $N - k$ graus de liberdade.



Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

22

Código no SAS

Através do procedimento GLM utiliza-se a opção DUNCAN no *statement* MEANS.

```
proc glm data=dados;  
    class    trat;  
    model    resp = trat;  
    means    trat / duncan;  
  
run;
```



Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

23

Aplicação

Alpha	0.05
Error Degrees of Freedom	45
Error Mean Square	10.41633

Number of Means	2	3	4	5
Critical Range	2.907	3.057	3.156	3.227

Means with the same letter
are not significantly different.

Duncan Grouping	Mean	N	Trt
A	20.948	10	drugE
B	15.361	10	drugD
C	12.374	10	4time
D	9.226	10	2time
E	5.781	10	1time



Teste de Duncan

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

24

Discussão

- ▶ Controla o **CER** ao nível α .
- ▶ Considerado o teste com maior poder, todavia possui alta taxa do Erro Tipo I por experimento (**EER**).
- ▶ Não computa intervalos de confiança simultâneos, pois utiliza diferentes valores críticos.
- ▶ Diante disso, manual do SAS não recomenda seu uso.



Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

25

Explicação

- ▶ O teste de Dunnett é um procedimento para comparar a média do tratamento controle e os demais tratamentos. O procedimento desenvolvido por Dunnett (1964) tem como as hipóteses especificadas em (2).
- ▶ Além dos pressupostos previamente citados, o teste de Dunnett exige que haja também a presença de um tratamento controle nos dados.



Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

26

Regra de decisão

Dada as hipóteses especificadas em (2), a média do i -ésimo tratamento difere significativamente em relação ao tratamento controle se

$$| \bar{y}_{i.} - \bar{y}_{a.} | \geq d_{\alpha}(k-1, N-k) \sqrt{QME \left(\frac{1}{n_i} + \frac{1}{n_a} \right)} \quad (8)$$

Em que

- ▶ a é o tratamento controle.
- ▶ α é o nível de significância conjunta associada com os $k-1$ testes.
- ▶ $d_{\alpha}(k-1, N-k)$ é conhecido como “*many-to-one t statistics*”.



Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

27

Código em SAS

O procedimento GLM possui a opção DUNNET no *statement* MEANS. Ou ainda, a opção ADJUST = DUNNETT no *statement* LSMEANS.

```
proc glm data=dados;  
    class    trat;  
    model    resp = trat;  
    means    trat / dunnett("control");  
    lsmeans  trat / adjust = dunnett  
                pdiff = control("level");  
  
run;
```



Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

28

Aplicação

Comparison	Difference	LowerCL	UpperCL	Significance
drugE – drugD	5.587	1.933	9.241	***
4time – drugD	–2.987	–6.641	0.667	
2time – drugD	–6.135	–9.789	–2.481	***
1time – drugD	–9.580	–13.234	–5.926	***



Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

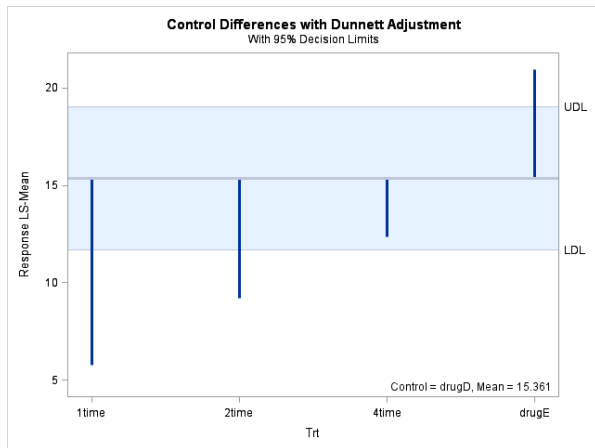
Teste de Dunnett

Recomendações

Referências

29

Control plot





Teste de Dunnett

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

30

Discussão

- ▶ O teste de Dunnett é mais poderoso que outros procedimentos, uma vez que é comparado o efeito dos tratamentos com um único grupo controle.
- ▶ O SAS permite realizar o teste de Dunnett unilateral através das opções DUNNETTL e DUNNETTU no procedimento GLM.



Recomendações

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

31

Recomendações SAS

Em resumo o manual do SAS orienta da seguinte forma:

- ▶ Se o interesse está em realizar diversas comparações individuais sem preocupação com os efeitos da inferência simultânea, recomenda-se o uso do teste t ou Fisher's LSD.
- ▶ Se o interesse está em comparar todos os pares ou comparar todos os tratamentos com um controle, sugere-se o uso, respectivamente, dos testes Tukey ou Dunnett.
- ▶ Se preferir por suposições inferências fracas e não está interessado em intervalos de confiança para as diferenças médias, pode-se optar pelo procedimento REGWQ.
- ▶ Pode-se escolher a abordagem Bayesiana de Waller e Duncan, caso concorde com suas suposições.



Recomendações

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

32

Resumo

Table: Principais Testes de Comparação Múltipla

Teste	Erro Tipo I	Comparação	Estatística
Fisher's LSD	CER	Dois a dois	t
Bonferroni	EER	Contraste	t [†]
Tukey HSD	EER	Dois a dois	Range
Newman-Keuls	CER	Dois a dois	Range [†]
Duncan	CER	Dois a dois	Range [†]
REGWQ	EER	Dois a dois	Range [†]
Scheffé	EER	Contraste	F [†]
Dunnet	EER	Controle	t [†]

[†] Modificado.



Referências

Testes de Comparação
Múltipla

André F. B. Menezes
Victor H. Nagahama

Introdução

Teste de Bonferroni

Teste de Duncan

Teste de Dunnett

Recomendações

Referências

33

[1] HOWELL, David C. **Statistical methods for psychology**. Cengage Learning, 2012.

[2] MONTGOMERY, Douglas C., **Design and Analysis of Experiments**, John Wiley & Sons, 2006.

[3] SAS INSTITUTE, INC., **SAS/STAT 9.2 User's Guide**, Cary, NC: SAS Institute, Inc., 2011.

[4] WESTFALL, Peter H.; TOBIAS, Randall D.; WOLFINGER, Russell D. **Multiple comparisons and multiple tests using SAS**. SAS Institute, 2011.