UNIVERSIDADE DE SÃO PAULO
Faculdade de Zootecnia e Engenharia de Alimentos

ZAB0229 – ESTATÍSTICA EXPERIMENTAL

Prof. César Gonçalves de Lima

PROCEDIMENTOS DE COMPARAÇÕES MÚLTIPLAS

(I) Teste t para contrastes

2. PROCEDIMENTOS DE COMPARAÇÕES MÚLTIPLAS OU TESTES DE COMPARAÇÕES DE MÉDIAS

Exemplo 2.1. Em um experimento de alimentação de suínos, foram utilizadas quatro rações (A, B, C e D), cada uma fornecida a cinco animais. Os ganhos de peso, em kg, foram:

	A	B	C	D
	35	40	39	27
	19	35	27	12
	31	46	20	13
	15	41	29	28
	30	33	45	30
Média	26	39	32	22

Com as SQ's calculadas podemos construir o seguinte quadro:

Quadro de ANOVA

Fonte de Variação	g.l.	SQ	QM	F
Ração	3	823,75	274,58	3,99
Resíduo	16	1100,00	68,75	
Total	19	1923,75		

$\bar{y}_{\bullet\bullet} = 29{,}75$ $\quad s^2 = 68{,}75$ $\quad \text{CV} = 27{,}9\%$

- Hipóteses: $H_0: \mu_A = \mu_B = \mu_C = \mu_D$

  H_a: pelo menos duas médias diferem entre si.

- Para testar H_0 usamos a estatística $F = \frac{QMRação}{QMResiduo}$ que tem distribuição $F(3; 16)$.

- Da Tábua IV obtemos $F_{tab} = 3{,}24$ e como $F_{calc} = 3{,}99 > F_{tab}$, nós rejeitamos a hipótese H_0 e concluímos ($\alpha = 5\%$) que pelo menos duas médias de ganho de peso diferem entre si, ou que pelo menos duas rações proporcionam ganhos médios de peso diferentes nos suínos.

Dúvidas que persistem:

- Qual é a melhor ração?

- Qual é a ração que proporciona o maior ganho de peso dos suínos?

Quando a $H_0: \mu_1 = \mu_2 = \cdots = \mu_a$ é rejeitada e aceitamos H_a: Pelo menos duas médias diferem entre si, como verdadeira, **não conseguimos** identificar quais as médias que diferem entre si.

PROPOSTA: Utilizar algum <u>Procedimento de Comparação Múltipla</u> ou <u>Teste de Comparação de Médias</u> para identificar as médias que são estatisticamente iguais ou diferentes entre si, como:

- **Teste *t*-Student** e **teste de Scheffé** para contrastes de médias.
- **Teste de Tukey** para comparações de médias duas-a-duas.
- **Teste de Duncan** para comparações de médias duas-a-duas.
- **Teste de Dunnett** para comparações de médias com a de um tratamento padrão.

2.1. DEFINIÇÕES INICIAIS

Situação: Temos um experimento simples com a tratamentos, cujas médias populacionais são $\mu_1, \ldots, \mu_a$ e cujas estimativas $\bar{y}_1, \ldots, \bar{y}_a$ foram calculadas com amostras de tamanhos $n_1 = \ldots = n_a = n$.

Definição 1. Comparação ou **contraste de médias** é qualquer função linear do tipo

$$Y = c_1\mu_1 + c_2\mu_2 + \cdots + c_a\mu_a$$

em que $\sum_{i=1}^{a} c_i = 0$.

Formulado um contraste, estaremos interessados em testar

$$H_0: Y = 0 \;\; versus \;\; H_a: Y \neq 0$$

Exemplo: No Exemplo 2.1 podemos estar interessados em realizar as seguintes comparações de médias:

$$Y_1 = \mu_A - \mu_B$$

$$Y_2 = \mu_A + \mu_B - 2\mu_C$$

$$Y_3 = \mu_A + \mu_B - \mu_C - \mu_D$$

- Testar H_0: $Y_1 = 0$ é equivalente a testar $\mu_A - \mu_B = 0$ ou $\mu_A = \mu_B$, ou seja, testar se "os ganhos médios de peso dos animais que receberam as rações A e B são iguais".

- Testar H_0: $Y_2 = 0$ é equivalente a testar $\mu_A + \mu_B - 2\mu_C = 0$, ou que $\frac{\mu_A + \mu_B}{2} = \mu_C$, ou seja, testar se "o ganho médio de peso dos animais que receberam as rações A ou B é igual ao dos animais que receberam a ração C".

- Testar H_0: $Y_4 = 0$ é equivalente a testar $\mu_A + \mu_B - \mu_C - \mu_D = 0$ ou $\mu_A + \mu_B = \mu_C + \mu_D$, ou que $\frac{(\mu_A+\mu_B)}{2} = \frac{(\mu_C+\mu_D)}{2}$, ou seja, testar se "o ganho médio de peso dos animais que receberam as rações A ou B é igual ao dos animais que receberam as rações C ou D".

Supondo que todos os dados foram obtidos de grupos (populações) com variância σ^2, que as amostras têm o mesmo número de repetições ($n_i = n$) e que uma boa estimativa (imparcial e eficiente) de σ^2 é o quadrado médio do resíduo ($QMResiduo$) tem-se que:

$a)$ $\hat{Y} = c_1\bar{y}_1 + c_2\bar{y}_2 + \cdots + c_a\bar{y}_a$ é um estimador imparcial do contraste Y.

$b)$ $\widehat{var}(\hat{Y}) = \dfrac{QMResiduo}{n}\sum_{i=1}^{a} c_i^2$

Dentre todos os contrastes que podemos formular, existem alguns que apresentam uma propriedade interessante: são **ortogonais**.

- A ortogonalidade entre os contrastes de médias garante a **independência** nas conclusões dos testes do tipo H_0: $Y = 0$.

Definição 2: Dois contrastes $Y_1 = \sum_{i=1}^{a} c_i\, \mu_i$ e $Y_2 = \sum_{i=1}^{a} d_i\, \mu_i$ são considerados *ortogonais* se

- $\sum_{i=1}^{a} c_i\, d_i = 0$, quando o experimento é balanceado ($n_i = n$)
- $\sum_{i=1}^{a} \frac{c_i d_i}{n_i} = 0$, quando o experimento é desbalanceado ($n_i \neq n$)

Importante: Quando um experimento envolve a tratamentos nós só conseguiremos construir conjuntos de $(a-1)$ contrastes ortogonais e $a(a-1)/2$ contrastes envolvendo pares de médias.

Exemplo: Os quatro contrastes formulados para o Exemplo 2.1, $Y_1 = \mu_A - \mu_B$, $Y_2 = \mu_B - \mu_D$, $Y_3 = \mu_A + \mu_B - 2\mu_C$ e $Y_4 = \mu_A + \mu_B - \mu_C - \mu_D$, formam um grupo de contrastes ortogonais?

Dica: Para facilitar a conclusão sobre a ortogonalidade de contrastes, devemos montar uma tabela com os seus coeficientes e verificar se $\sum_{i=1}^{a} c_i \, d_i = 0$, para todos os pares de contrastes:

Contraste	μ_A	μ_B	μ_C	μ_D
$Y_1 = \mu_A - \mu_B$	1	-1	0	0
$Y_2 = \mu_B - \mu_D$	0	1	0	-1
$Y_3 = \mu_A + \mu_B - 2\mu_C$	1	1	-2	0
$Y_4 = \mu_A + \mu_B - \mu_C - \mu_D$	1	1	-1	-1

Você pode verificar facilmente que:

- Os contrastes Y_1 e Y_3, Y_1 e Y_4 são $ortogonais$.
- Os contrastes Y_1 e Y_2, Y_2 e Y_3, Y_2 e $Y_{4,}$ Y_3 e Y_4 <u>não são</u> $ortogonais$.

Exemplos de dois conjuntos de 3 contrastes ortogonais do Exemplo 2.1:

$$Y_1 = \mu_A - \mu_B \quad Y_2 = \mu_C - \mu_D \quad \text{e} \quad Y_3 = \mu_A + \mu_B - \mu_C - \mu_D$$

ou

$$Y_1 = \mu_A - \mu_B \quad Y_2 = \mu_A + \mu_B - 2\mu_C \quad \text{e} \quad Y_3 = \mu_A + \mu_B + \mu_C - 3\mu_D$$

(Verifique!!!)

Pergunta: Porque é interessante trabalhar com contrastes ortogonais?

Resposta:

i) As conclusões dos testes envolvendo contrastes ortogonais são independentes.

ii) $SQTrat$ pode ser particionada em $(a-1)$ somas de quadrados independentes associadas aos contrastes ortogonais.

Veremos uma grande utilidade desta partição quando o fator de tratamento for quantitativo e realizarmos testes de tendência de regressão.

A seguir são apresentados mais exemplos interessantes de contrastes (ortogonais).

Exemplo 2.3. Estudo de consorciação na cultura de abacaxi:

Tratamento	Descrição
A	Abacaxi (0,90×0,30m) monocultivo
B	Abacaxi (0,80×0,30m) monocultivo
C	Abacaxi (0,80×0,30m) + Amendoim
D	Abacaxi (0,80×0,30m) + Feijão

$\boldsymbol{Obs}$: Temos $a = 4$ tratamentos ⇒ podemos construir conjuntos com 3 (três) contrastes ortogonais.

Contraste	Significado
$Y_1 = \mu_A + \mu_B - \mu_C - \mu_D$	Compara monocultivo e consorciação
$Y_2 = \mu_A - \mu_B$	Compara os dois espaçamentos no monocultivo.
$Y_3 = \mu_C - \mu_D$	Compara os dois tipos de consorciação

H_0: $Y_1 = 0 \Rightarrow$ testar $(\mu_A + \mu_B)/2 = (\mu_C + \mu_D)/2$, ou se a produção média do monocultivo (A e B) é igual à da consorciação (C e D).

H_0: $Y_2 = 0 \Rightarrow$ testar $\mu_A = \mu_B$, ou se, no monocultivo, as produções médias nos espaçamentos 0,90×0,30m e 0,80×0,30m são iguais.

H_0: $Y_3 = 0 \Rightarrow$ testar $\mu_C = \mu_D$, ou s as produções médias das consorciações com amendoim e feijão no espaçamento 0,80×0,30m são iguais entre si.

Exemplo 2.4. Ensaio de adubação de uma forrageira com fósforo.

Adubo	Descrição
1	Testemunha (sem fósforo).
2	50 kg/ha de P_2O_5 – no outono.
3	50 kg/ha de P_2O_5 – ½ no outono e ½ na primavera.
4	100 kg/ha de P_2O_5 – no outono.
5	100 kg/ha de P_2O_5 – ½ no outono e ½ na primavera.

Algumas comparações (contrastes) de interesse podem ser escritas como:

Contraste	Significado
$Y_1 = 4\mu_1 - \mu_2 - \mu_3 - \mu_4 - \mu_5$	Avalia o efeito da adubação com fósforo.
$Y_2 = \mu_2 + \mu_3 - \mu_4 - \mu_5$	Compara dose simples (50 kg/ha) com a dose dupla (100 kg/ha).
$Y_3 = \mu_2 - \mu_3$	Compara a dose simples única com a dose simples parcelada.
$Y_4 = \mu_4 - \mu_5$	Compara a dose dupla única com a dose dupla parcelada.
$Y_5 = \mu_2 - \mu_3 + \mu_4 - \mu_5$	Compara dose única com a dose parcelada.

Dúvida: Parece que nem todos os contrastes desses dois exemplos são ortogonais. Vamos verificar...

Exemplo 2.3

	μ_1	μ_2	μ_3	μ_4
Y_1	1	1	–1	–1
Y_2	1	–1	0	0
Y_3	0	0	1	–1

Exemplo 2.4

	μ_1	μ_2	μ_3	μ_4	μ_5
Y_1	4	–1	–1	–1	–1
Y_2	0	1	1	–1	–1
Y_3	0	1	–1	0	0
Y_4	0	0	0	1	–1
Y_5	0	1	–1	1	–1

Após realizar a verificação $\sum_{i=1}^{a} c_i\, d_i = 0$, concluímos que:

- Os três contrastes do Exemplo 2.3 são mutuamente ortogonais.
- Os contrastes Y_1, Y_2, Y_3 e Y_4 do Exemplo 2.4 são mutuamente ortogonais.
- O contraste Y_5 só é ortogonal aos contrastes Y_1 e Y_2.

Relembrando:

- Quando $a = 5$ tratamentos $\Rightarrow$ só conseguimos montar grupos com $a - 1 = 4$ contrastes ortogonais!
- A ortogonalidade entre contrastes de médias garante a **independência** entre as conclusões dos testes do tipo $H_0: Y = 0$.

A seguir, vamos conhecer o teste t-Student e o teste de Scheffé que podem ser usados para testar hipóteses sobre contrastes que envolvem duas ou mais médias, como aqueles apresentados até agora.

2.2. O TESTE t-STUDENT PARA CONTRASTES DE MÉDIAS

- O teste t-Student pode ser usado para testar hipóteses do tipo

$$H_0: Y = 0$$

  Em que $Y = c_1\mu_1 + c_2\mu_2 + \cdots + c_a\mu_a$ é um contraste de médias definido *a priori*, ou seja, antes de examinar os dados ou as médias dos tratamentos.

- Embora não exista limite para o número de contrastes envolvendo as médias de tratamentos, o número máximo de contrastes ortogonais é $(a-1)$, onde a é o número de tratamentos.

- **Mais uma vez:** Se os contrastes usados na análise forem ortogonais, o teste t-Student garante a independência nas conclusões dos testes.

Para testar a hipótese H_0: $Y = \sum_{i=1}^{a} c_i\, \mu_i = 0$ *versus* H_a: $Y \neq 0$, nós usamos a estatística:

$$t = \frac{\hat{Y}}{\sqrt{\frac{QMResiduo}{n} \sum_{i=1}^{a} c_i^2}}\,,\ \text{com } \hat{Y} = c_1\bar{x}_1 + c_2\bar{x}_2 + \cdots + c_a\bar{x}_a$$

que tem distribuição t-Student com gl_{Res} graus de liberdade.

Fixando um nível de significância $\alpha = 5\%$, por exemplo, buscamos o valor t_{tab} na Tábua III e comparamos com o valor de t_{calc}, calculado para o contraste.

- Se $|t_{calc}| > t_{tab}$ nós rejeitamos a hipótese H_0 e aceitamos H_a como verdadeira, ao nível α de significância.
- Se $|t_{calc}| \leq t_{tab}$, não rejeitamos a hipótese H_0.

- É comum indicarmos com um asterisco à direita do valor da estatística (t_{calc}) a rejeição da hipótese H_0, ao nível $\alpha = 5\%$ de significância.

- No caso da não rejeição de H_0 nós colocamos $n.s.$ à direita do t_{calc} para indicar um resultado "não significativo".

IMPORTANTE:

- Todos os resultados dos testes envolvendo contrastes devem ser comentados, se rejeitarmos ou não a hipótese $H_0: Y = 0$.

- O teste t-Student é indicado para contrastes complexos, que envolvem mais de duas médias e são elaborados $a\ priori$.

- Para comparar pares de médias existem testes específicos, que serão conhecidos nas próximas aulas.

Exemplo 2.5. Avaliar o efeito da administração de raízes e tubérculos como suplementação de inverno na alimentação de vacas em lactação. Utilizou-se um DIC com 4 tratamentos e 6 repetições por tratamento. As produções médias diárias (kg) de leite a 4% de gordura das vacas que receberam os tratamentos foram:

Tratamento	Testemunha	Mandioca	Araruta	Batata doce
Média (kg)	22,2	27,8	32,2	22,8

Sabe-se que $s^2 = QMResiduo = 12{,}467$ ($20\ g.l.$) e que o teste F da ANOVA rejeitou a hipótese de que as médias dos tratamentos são iguais entre si.

Teste a hipótese $H_0\colon Y_i = 0$ para os seguintes contrastes ortogonais:

$$Y_1 = 3\mu_1 - \mu_2 - \mu_3 - \mu_4, \quad Y_2 = 2\mu_3 - \mu_2 - \mu_4 \ \text{ e } \ Y_3 = \mu_2 - \mu_4$$

Trat	1	2	3	4			
Média	22,2	27,8	32,2	22,8	$\hat{Y}_i$	Σc_i^2	$\|t_{calc}\|$
Y_1	3	-1	-1	-1	$-16{,}2$	12	3,24 *
Y_2	0	-1	2	-1	13,8	6	3,91 *
Y_3	0	1	0	-1	5,0	2	2,45 *

Note: Para calcular $\hat{Y}_i$, nós usamos $\hat{Y}_i = c_1\bar{x}_1 + c_2\bar{x}_2 + c_3\bar{x}_3 + c_4\bar{x}_4$, em que $\bar{x}_i$ é a média aritmética das observações do tratamento i.

Para $\alpha = 5\%$, a Tábua III (distribuição t) nos fornece ($20\ gl$) o valor $t_{tab} = 2{,}086$. Como $|t_{calc}| > 2{,}086$ rejeitamos a nulidade de cada um dos três contrastes.

Precisamos relatar as conclusões dos testes!!!

Como todos os contrastes são significativamente diferentes de zero, avaliando o valor de cada $\hat{Y}_i$, podemos comentar que:

- $\mu_1 < \frac{(\mu_2+\mu_3+\mu_4)}{3}$, ou seja, a produção média de leite das vacas que receberam suplementação com raízes e tubérculos foi superior à das vacas não suplementadas.

- $\mu_3 > \frac{(\mu_2+\mu_4)}{2}$, ou seja, a suplementação com araruta resultou em uma produção média de leite superior à dos animais que receberam a suplementação com mandioca ou batata-doce.

- $\mu_2 > \mu_4$, ou seja, a produção média das vacas que receberam a suplementação com mandioca foi superior à das vacas que receberam batata-doce.

Exercício: Formule 3 contrastes ortogonais para os dados do Exemplo 2.1 (Experimento de alimentação de suínos com 4 rações, cada uma fornecida a 5 animais) e teste a hipótese H_0: $Y_i = 0$ utilizando o teste t. Apresente os resultados "por extenso".

Trat	A	B	C	D	$\hat{Y}_i$	Σc_i^2	$\lvert t_{calc} \rvert$
Média	26	39	32	22			
Y_1							
Y_2							
Y_3							

Lembrando que: $gl_{trat} = 3$, $gl_{res} = 16$ e $QMRes = 68{,}75$