MAF 261 - Estatística Experimental

Prof. Fernando de Souza Bastos

Instituto de Ciências Exatas e Tecnológicas
Universidade Federal de Viçosa
Campus UFV - Florestal

Sumário

Como já vimos, a análise da variância serve para verificar se há alguma diferença significativa entre as médias dos níveis de um fator a um determinado nível de significância. No caso em que o teste F for significativo, ou seja, a hipótese de nulidade for rejeitada, vimos que existe pelo menos um contraste entre médias estatisticamente diferente de zero.

Os procedimentos de comparações múltiplas que veremos, visam identificar quais são estes contrastes de forma que possamos identificar qual é o nível do fator em estudo que apresentou maior média.

Dentre os vários procedimentos existentes para comparações múltiplas, veremos:

1. Procedimentos para testar todos os possíveis contrastes entre duas médias dos níveis do fator em estudo:
    - Teste de Tukey;

Dentre os vários procedimentos existentes para comparações múltiplas, veremos:

1. Procedimentos para testar todos os possíveis contrastes entre duas médias dos níveis do fator em estudo:
    - Teste de Tukey;
    - Teste de Duncan;

Dentre os vários procedimentos existentes para comparações múltiplas, veremos:

1. Procedimentos para testar todos os possíveis contrastes entre duas médias dos níveis do fator em estudo:
    - Teste de Tukey;
    - Teste de Duncan;
2. Prodedimentos para testar todos os possíveis contrastes entre médias dos níveis do fator em estudo:
    - Teste t;

Dentre os vários procedimentos existentes para comparações múltiplas, veremos:

1. Procedimentos para testar todos os possíveis contrastes entre duas médias dos níveis do fator em estudo:
    - Teste de Tukey;
    - Teste de Duncan;
2. Prodedimentos para testar todos os possíveis contrastes entre médias dos níveis do fator em estudo:
    - Teste t;
    - Teste de Scheffé.

Todos os procedimentos se baseiam no cálculo de uma diferença mínima significativa (dms). A dms representa o menor valor que a estimativa de um contraste deve apresentar para que se possa considerá-lo como significativo.

Teste de Tukey

Usaremos o teste de Tukey para comparar a totalidade dos contrastes entre duas médias, ou seja, $C = \mu_i - \mu_u, \quad 1 \leq i < u \leq I$. Este teste baseia-se na diferença mínima significativa (d.m.s.), dada por:

$$\Delta = q\sqrt{\frac{1}{2}\hat{V}(\hat{C})},$$

em que $q = q_\alpha(I, n_2)$ é o valor tabelado da amplitude total estudentizada, na qual α é o nível de significância, I é o número de níveis do fator em estudo, n_2 são os graus de liberdade do resíduo e $\hat{V}(\hat{C}) = QMRes\left(\frac{1}{r_i} + \frac{1}{r_u}\right)$.

No caso em que todos os tratamentos apresentarem o mesmo número de repetições, ou seja, $r_i = r_u = K$, então o valor de Δ é simplificado para a seguinte expressão:

$$\Delta = q\sqrt{\frac{QMRes}{K}}$$

Para realizar o teste de Tukey, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular Δ;

Para realizar o teste de Tukey, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular Δ;
2. ordenar as médias do fator em estudo em ordem decrescente;

Para realizar o teste de Tukey, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular Δ;
2. ordenar as médias do fator em estudo em ordem decrescente;
3. montar grupos de comparação entre os contrastes e obter as estimativas dos contrastes, com base nos valores amostrais;

Para realizar o teste de Tukey, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular Δ;
2. ordenar as médias do fator em estudo em ordem decrescente;
3. montar grupos de comparação entre os contrastes e obter as estimativas dos contrastes, com base nos valores amostrais;
4. concluir usando a seguinte relação: se $|\hat{C}| \geq \Delta$, rejeita-se H_0 e se $|\hat{C}| < \Delta$, não rejeita-se H_0. No último caso, indicar as médias iguais, seguidas por uma mesma letra.

Para comparar a produtividade de quatro variedades de milho, um agrônomo tomou vinte parcelas similares e distribuiu, inteiramente ao acaso, cada uma das 4 variedades em 5 parcelas experimentais. A partir dos dados experimentais fornecidos, é possível concluir que existe diferença significativa entre as variedades com relação a produtividade, utilizando o nível de significância de 5%?

	Variedades			
	A	B	C	D
	25	31	22	33
	26	25	26	29
	20	28	28	31
	23	27	25	34
	21	24	29	28
Totais	115	135	130	155
Médias	23	27	26	31

Teste de Duncan

Assim como o teste de Tukey, o teste de Duncan será válido para a totalidade dos contrastes de duas médias, ou seja, $C = m_i - m_u$, $1 \leq i < u \leq I$. Este teste baseia-se na amplitude total mínima significativa dada por:

$$D_n = z_n \sqrt{\frac{1}{2}\hat{V}(\hat{C})},$$

em que $z_n = z_\alpha(n, n_2)$ é o valor tabelado da amplitude total estudentizada, na qual α é o nível de significância, n é o número de médias ordenadas abrangidas pelo contraste entre os níveis do fator em estudo, n_2 são os graus de liberdade do resíduo e $\hat{V}(\hat{C}) = QMRes\left(\frac{1}{r_i} + \frac{1}{r_u}\right)$.

No caso em que todos os tratamentos apresentarem o mesmo número de repetições, ou seja, $r_i = r_u = K$, então o valor de D_n é simplificado para a seguinte expressão:

$$D_n = z_n \sqrt{\frac{QMRes}{K}}$$

Para realizar o teste de Duncan, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular o valor D_n, para $n = 2, \cdots, I$;

Para realizar o teste de Duncan, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular o valor D_n, para $n = 2, \cdots, I$;
2. ordenar as médias do fator em estudo em ordem decrescente;

Para realizar o teste de Duncan, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular o valor D_n, para $n = 2, \cdots, I$;
2. ordenar as médias do fator em estudo em ordem decrescente;
3. montar grupos de comparação entre os contrastes e obter as estimativas dos contrastes, com base nos valores amostrais;

Para realizar o teste de Duncan, a um nível de signifiância α, deve-se seguir os seguintes passos:

1. calcular o valor D_n, para $n = 2, \cdots, I$;
2. ordenar as médias do fator em estudo em ordem decrescente;
3. montar grupos de comparação entre os contrastes e obter as estimativas dos contrastes, com base nos valores amostrais;
4. concluir usando a seguinte relação: se $|\hat{C}| \geq D_n$, rejeita-se H_0 e se $|\hat{C}| < D_n$, não rejeita-se H_0. No último caso, indicar as médias iguais, seguidas por uma mesma letra.

Para comparar a produtividade de quatro variedades de milho, um agrônomo tomou vinte parcelas similares e distribuiu, inteiramente ao acaso, cada uma das 4 variedades em 5 parcelas experimentais. A partir dos dados experimentais fornecidos, é possível concluir que existe diferença significativa entre as variedades com relação a produtividade, utilizando o nível de significância de 5%?

	Variedades			
	A	B	C	D
	25	31	22	33
	26	25	26	29
	20	28	28	31
	23	27	25	34
	21	24	29	28
Totais	115	135	130	155
Médias	23	27	26	31

O teste t pode ser usado para testar contrastes envolvendo duas ou mais médias. Porém, este teste exige:

- as comparações a serem realizadas devem ser determinadas antes dos dados serem examinados;
- podem-se testar no máximo, tantos contrastes quantos são os graus de liberdade para tratamentos e estes contrastes devem ser ortogonais;

Consideremos um contraste entre médias, dado por:

$$C = a_1\mu_1 + \cdots + a_I\mu_I$$

do qual obtemos a estimativa por meio do estimador

$$\hat{C} = a_1\hat{\mu}_1 + \cdots + a_I\hat{\mu}_I$$

Considere a estatística t, dada por:

$$t_{cal} = \frac{\hat{C} - C}{\sqrt{QMRes \cdot \sum_{i=1}^{I} \frac{a_i^2}{r_i}}}$$

que tem distribuição t com n_2 graus de liberdade, sendo n_2 o número de graus de liberdade do resíduo.

Caso o número de repetições seja o mesmo para todos os tratamentos, ou seja, $r_1 = \cdots = r_I = K$, então a fórmula se resume a:

$$t_{cal} = \frac{\hat{C} - C}{\sqrt{\frac{QMRes}{K} \cdot \sum_{i=1}^{I} a_i^2}}$$

Quando aplicamos o teste t a um contraste C, geralmente o interesse é testar as hipóteses: $H_0 : C = 0$ contra $H_a : C \neq 0$.
O valor tabelado de t é obtido por $t_{tab} = t_{\alpha}(n_2)$ e a regra decisória é a seguinte:

- se $|t_{cal}| \geq t_{tab}$, rejeita-se H_0
- caso contrário, não rejeita-se H_0

Teste de Scheffé

Este teste pode ser aplicado para testar todo e qualquer contraste entre médias, mesmo quando sugerido pelos dados. O teste de Scheffé não exige que os contrastes sejam ortogonais e nem que estes contrastes sejam estabelecidos antes de se examinar os dados.

A estatística do teste, denotada por S, é calculada por:

$$S_{cal} = \sqrt{(I-1) \cdot F_{tab} \cdot QMRes \cdot \sum_{i=1}^{I} \frac{a_i^2}{r_i}}$$

em que I é o número de níveis do fator em estudo, $F_{tab} = F_{\alpha}(I-1, n_2)$.

Caso o número de repetições seja o mesmo para todos os tratamentos, ou seja, $r_1 = \cdots = r_I = K$, então a fórmula se resume a:

$$S_{cal} = \sqrt{(I-1) \cdot F_{tab} \cdot \frac{QMRes}{K} \cdot \sum_{i=1}^{I} a_i^2}$$

Prosseguindo, deve-se calcular a estimativa do contrates C, ou seja, $\hat{C}$, e verificar se $|\hat{C}| \geq S$, concluindo que o contraste é significativamente diferente de zero ao nível de α de probabilidade, indicando que os grupos de médias confrontados no contraste diferem entre si a esse nível de probabilidade.

Exemplo 1 (Exercício 5.6, pág. 52):

Quatro padarias da cidade de São Paulo, foram fiscalizadas para verificar a quantidade de bromato de potássio existente nos pães franceses que elas produzem. Com esta finalidade foi tomada uma amostra de pães, inteiramente ao acaso, de cada padaria e para cada um deles foi avaliado o teor de bromato de potássio (mg de bromato de potássio por 1 kg de pão). O resumo da avaliação é fornecido a seguir:

Padaria	1	2	3	4
Teor Médio	10	11	8	9
Núm. de pães avaliados	7	8	7	8

Usando $SQRes = 52$ e $\alpha = 5\%$:

1. Pode-se concluir que existe diferença significativa no teor médio de bromato de potássio no pão entre as padarias avaliadas?
2. Suponha que as padarias 1 e 2 suprem a classe A, a padaria 3 a classe B e a 4 a classe C. Verifique, por meio de um contraste, pelos testes de Scheffé e t, se existe diferença no teor médio de bromato de potássio entre as padaria que suprem as classes A e C.