# MAF 261 - Estatística Experimental

Prof. Fernando de Souza Bastos

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

# Sumário

O estudo de contrastes é muito importante na Estatística Experimental, principalmente quando o experimento em análise é composto por mais do que dois tratamentos. Com o uso de contrastes é possível ao pesquisador estabelecer comparações, entre tratamentos ou grupos de tratamentos, que sejam de interesse. Vamos estudar os fundamentos para estabelecer grupos de contrastes, obter a estimativa para cada contraste estabelecido, bem com estimar a variabilidade associada a cada um destes contrastes.

**Definição:** Uma função linear de médias populacionais de tratamentos dada por

$$C = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

é dita ser um contraste se satisfizer:

$$\sum_{i=1}^{n} a_i = 0$$

Na prática, geralmente não se conhece os valores das médias populacionais $\mu_i$, mas suas estimativas. Daí, em Estatística Experimental, não trabalhamos com o contraste $C$ mas com o seu estimador $\hat{C}$, que também é uma função linear de médias obtidas por meio de experimentos ou amostras. Assim tem-se que o estimador para o contraste de médias é dado por:

$$\hat{C} = a_1\hat{\mu}_1 + a_2\hat{\mu}_2 + \cdots + a_n\hat{\mu}_n$$

# Exemplo

É contraste?

1. $C_1 = \mu_1 - \mu_2$

# Exemplo

É contraste?

1. $C_1 = \mu_1 - \mu_2$
2. $C_2 = \mu_1 + 2\mu_2 + 5\mu_3 - 4\mu_4 - 3\mu_5$

# Exemplo

É contraste?

1. $C_1 = \mu_1 - \mu_2$
2. $C_2 = \mu_1 + 2\mu_2 + 5\mu_3 - 4\mu_4 - 3\mu_5$
3. $C_1 = \mu_1 + \mu_2 - 2\mu_3$

Num experimento de consórcio na cultura do abacaxi, com 5 repetições, as médias de produção de frutos de abacaxi (em t/ha), foram as seguintes:

| Tratamentos | $\hat{\mu}_i$ |
|---|---|
| 1−Abacaxi $(0,90\ x0,30m)$ monocultivo | 53,5 |
| 2−Abacaxi $(0,80x0,30m)$ monocultivo | 56,5 |
| 3−Abacaxi $(0,80x0,30m)+$ amendoim | 62,0 |
| 4−Abacaxi $(0,80x0,30m)+$ feijão | 60,4 |

Pede-se obter as estimativas dos seguintes contrastes:

$$C_1 = \mu_1 + \mu_2 - \mu_3 - \mu_4; \;\; C_2 = \mu_1 - \mu_2; \;\; C_3 = \mu_3 - \mu_4$$

Considere o estimador do contraste C, dado por:

$$\hat{C} = a_1\hat{\mu}_1 + a_2\hat{\mu}_2 + \cdots + a_n\hat{\mu}_n$$

A variância do estimador do contraste é dada por:

$$V(\hat{C}) = V(a_1\hat{\mu}_1 + a_2\hat{\mu}_2 + \cdots + a_n\hat{\mu}_n)$$

Admitindo independência entre as médias, temos:

$$V(\hat{C}) = V(a_1\hat{\mu}_1) + V(a_2\hat{\mu}_2) + \cdots + V(a_n\hat{\mu}_n)$$

Logo,

$$V(\hat{C}) = a_1^2 V(\hat{\mu}_1) + a_2^2 V(\hat{\mu}_2) + \cdots + a_n^2 V(\hat{\mu}_n)$$

Lembrando que $\hat{\mu}_i = \dfrac{\sum_{j=1}^{r_i} x_j}{r_i}$ temos que $V(\hat{\mu}_i) = \dfrac{\sigma_i^2}{r_i}$, assim:

$$V(\hat{C}) = a_1^2 \frac{\sigma_1^2}{r_1} + a_2^2 \frac{\sigma_2^2}{r_2} + \cdots + a_n^2 \frac{\sigma_n^2}{r_n}$$

Admitindo-se homogeneidade de variâncias, ou seja, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$, então

$$V(\hat{C}) = \left( \frac{a_1^2}{r_1} + \frac{a_2^2}{r_2} + \cdots + \frac{a_n^2}{r_n} \right) \sigma^2 = \sigma^2 \sum_{i=1}^{n} \frac{a_i^2}{r_i}$$

Na prática, geralmente, não se conhece a variância $\sigma^2$, mas sua estimativa a qual é obtida por meio de dados experimentais. Esta estimativa é denominada como estimador comum $S_c^2$. Então, o que normalmente se obtém é o valor do estimador da variância do estimador do contraste, que é obtida por:

$$\hat{V}(\hat{C}) = S_c^2 \sum_{i=1}^{n} \frac{a_i^2}{r_i}$$

Em algumas situações desejamos testar um grupo de contrastes relacionados com o experimento em estudo. Alguns tipos de testes indicados para este objetivo, necessitam que os contrastes, que compõem o grupo a ser testado, sejam ortogonais entre si. A ortogonalidade entre os contrastes indica independência linear na comparação estabelecida por um contraste com a comparação estabelecida pelos outros contrastes.

Sejam os estimadores dos contrastes de $C_1$ e $C_2$ dados, respectivamente, por:

$$\hat{C}_1 = a_1\hat{\mu}_1 + a_2\hat{\mu}_2 + \cdots + a_n\hat{\mu}_n$$
$$\hat{C}_2 = b_1\hat{\mu}_1 + b_2\hat{\mu}_2 + \cdots + b_n\hat{\mu}_n$$

A covariância entre $\hat{C}_1$ e $\hat{C}_2$, supondo independência entre tratamentos, é obtida por $Cov(\hat{C}_1, \hat{C}_2) = a_1b_1V(\hat{\mu}_1) + a_2b_2V(\hat{\mu}_2) + \cdots + a_nb_nV(\hat{\mu}_n)$. E como já vimos, a variância da média amostral é dada por $V(\hat{\mu}_i) = \frac{\sigma_i^2}{r_i}$, para $i = 1, 2, \cdots, n$. Logo,

$$Cov(\hat{C}_1, \hat{C}_2) = a_1b_1\frac{\sigma_1^2}{r_1} + a_2b_2\frac{\sigma_2^2}{r_2} + \cdots + a_nb_n\frac{\sigma_n^2}{r_n}.$$

Admitindo que exista homogeneidade de variâncias entre os tratamentos, ou seja, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$, então,

$$Cov(\hat{C}_1, \hat{C}_2) = \left(\frac{a_1 b_1}{r_1} + \frac{a_2 b_2}{r_2} + \cdots + \frac{a_n b_n}{r_n}\right)\sigma^2 = \sigma^2 \sum_{i=1}^{n} \frac{a_i b_i}{r_i}$$

Sabe-se ainda que, se duas variáveis aleatórias são independentes, a covariância entre elas é igual a zero. Assim, se $\hat{C}_1$ e $\hat{C}_2$ são independentes, a covariância entre eles é igual a zero, isto é:

$$Cov(\hat{C}_1, \hat{C}_2) = 0$$

Para que a covariância seja nula, é necessário, portanto que:

$$\sum_{i=1}^{n} \frac{a_i b_i}{r_i} = 0$$

Esta é a condição de ortogonalidade entre dois contrastes para um experimento com número diferente de repetições para os tratamentos. Para um experimento com o mesmo número de repetições, satisfazendo as mesmas pressuposições (médias independentes e homogeneidade de variâncias), a condição de ortogonalidade se resume a:

$$\sum_{i=1}^{n} a_i b_i = 0$$

Para um experimento com $I$ tratamentos, podem ser formados vários grupos de contrastes ortogonais, no entanto cada grupo deverá conter no máximo $(I - 1)$ contrastes ortogonais, o que corresponde ao número de graus de liberdade para tratamentos. Dentro de um grupo de contrastes ortogonais, todos os contrastes tomados dois a dois, serão também ortogonais.

# Obtenção por Meio de Sistema de Equações Lineares

Neste método, deve-se estabelecer, a princípio, um contraste que seja de interesse e, a partir deste é que os demais são obtidos. Por meio da imposição da condição de ortogonalidade e da condição para ser um contraste, obtém-se equações lineares, cujas incógnitas são os coeficientes das médias que compõem o contraste. Como o número de incógnitas é superior ao número de equações existentes, será sempre necessário atribuir valores a algumas incógnitas. É desejável que os valores a serem atribuídos, permitam que os coeficientes sejam números inteiros.

Por meio desta metodologia, é possível estabelecer facilmente um grupo de contrastes ortogonais. A metodologia pode ser resumida nos seguintes passos (BANZATTO e KRONKA, 1989):

1. Divide-se o conjunto das médias de todos os tratamentos do experimento em dois grupos. O primeiro contraste é obtido pela comparação das médias de um grupo contra as médias do outro grupo. Para isso atribui-se sinais positivos para membros de um grupo e negativos para membros do outro grupo.

Por meio desta metodologia, é possível estabelecer facilmente um grupo de contrastes ortogonais. A metodologia pode ser resumida nos seguintes passos (BANZATTO e KRONKA, 1989):

1. Divide-se o conjunto das médias de todos os tratamentos do experimento em dois grupos. O primeiro contraste é obtido pela comparação das médias de um grupo contra as médias do outro grupo. Para isso atribui-se sinais positivos para membros de um grupo e negativos para membros do outro grupo.
2. Dentro de cada grupo formado no passo anterior, que possui mais que uma média, aplica-se o passo 1, subdividindo-os em subgrupos. Repete-se este passo até que se forme subgrupos com apenas uma média. Ao final, deveremos ter formado $(I-1)$ comparações.

Para se obter os coeficientes que multiplicam cada média que compõem os contrastes estabelecidos, deve-se, para cada contraste:

1. Verificar o número de parcelas experimentais envolvidas no $1^\circ$ grupo, digamos $g_1$, e o número de parcelas experimentais envolvidas no $2^\circ$ grupo, digamos $g_2$. Calcula-se o mínimo múltiplo comum (m.m.c.) entre $g_1$ e $g_2$.

Para se obter os coeficientes que multiplicam cada média que compõem os contrastes estabelecidos, deve-se, para cada contraste:

1. Verificar o número de parcelas experimentais envolvidas no $1^\circ$ grupo, digamos $g_1$, e o número de parcelas experimentais envolvidas no $2^\circ$ grupo, digamos $g_2$. Calcula-se o mínimo múltiplo comum (m.m.c.) entre $g_1$ e $g_2$.
2. Dividir o m.m.c. por $g_1$. O resultado será o coeficiente de cada média do $1^\circ$ grupo.

Para se obter os coeficientes que multiplicam cada média que compõem os contrastes estabelecidos, deve-se, para cada contraste:

1. Verificar o número de parcelas experimentais envolvidas no $1^\circ$ grupo, digamos $g_1$, e o número de parcelas experimentais envolvidas no $2^\circ$ grupo, digamos $g_2$. Calcula-se o mínimo múltiplo comum (m.m.c.) entre $g_1$ e $g_2$.
2. Dividir o m.m.c. por $g_1$. O resultado será o coeficiente de cada média do $1^\circ$ grupo.
3. Dividir o m.m.c. por $g_2$. O resultado será o coeficiente de cada média do $2^\circ$ grupo.

Para se obter os coeficientes que multiplicam cada média que compõem os contrastes estabelecidos, deve-se, para cada contraste:

1. Verificar o número de parcelas experimentais envolvidas no $1^\circ$ grupo, digamos $g_1$, e o número de parcelas experimentais envolvidas no $2^\circ$ grupo, digamos $g_2$. Calcula-se o mínimo múltiplo comum (m.m.c.) entre $g_1$ e $g_2$.
2. Dividir o m.m.c. por $g_1$. O resultado será o coeficiente de cada média do $1^\circ$ grupo.
3. Dividir o m.m.c. por $g_2$. O resultado será o coeficiente de cada média do $2^\circ$ grupo.
4. Multiplicar os coeficientes obtidos pelo número de repetições da respectiva média. Se possível, simplificar os coeficientes obtidos por uma constante. No caso em que o número de repetições é igual para todos os tratamentos, este passo pode ser eliminado.

# Exemplo

$$
\begin{aligned}
C_1 &= (T_1, T_2, T_3, T_4) \ vs \ T_5 \\
C_2 &= (T_1, T_2, T_3) \ vs \ T_4 \\
C_3 &= (T_1, T_2) \ vs \ T_3 \\
C_4 &= (T_1) \ vs \ T_2
\end{aligned}
$$

Em que $T_1$ possui 6 repetições, $T_2$ possui 6, $T_3$ possui 4, $T_4$ possui 5 e $T_6$ possui 6.