# Delineamento e Análise de Experimentos

Professora Juliana Betini Fachini Gomes
e-mail: jfachini@unb.br

Brasília - 2023

**Experimentos Fatorias**

- O caso mais simples de experimentos fatoriais envolvem apenas dois fatores: fator $A$ com $a$ níveis e fator $B$ com $b$ níveis;
- Totalizando $ab$ combinações de tratatmentos;
- Cada replicação do experimento contém todas as combinações de tratamento ab. No geral, existem n réplicas.

- Seja $y_{ijk}$ a resposta observada quando o fator $A$ está no $i$-ésimo nível ($i = 1, 2, \ldots, a$) e o fator $B$ está no $j$-ésimo nível ($j = 1, 2, \ldots, b$) para a $k$-ésima repetição ($k = 1, 2, \ldots, n$);

- As *abn* observações são selecionadas aleatoriamente segundo um experimento inteiramente casualizado.

O modelo de efeitos é definido por:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}, \tag{1}$$

em que $i = 1, 2, ..., a; j = 1, 2, ..., b; k = 1, 2, ..., n$; $\mu$ é a média geral, $\tau_i$ é o efeito do $i$-ésimo nível do fator $A$; $\beta_j$ é o efeito do $j$-ésimo nível do fator $B$; $(\tau\beta)_{ij}$ é o efeito da interação entre $\tau_i$ e $\beta_j$ e $\varepsilon_{ijk}$ é o componente de erro aleatório.

$\rightarrow$ Ambos os fatores são assumidos ser fixos, o efeito de cada tratamento é um desvio da média geral, então $\sum_{i=1}^{a} \tau_i = 0$ e $\sum_{j=1}^{b} \beta_j = 0$;

$\rightarrow$ Similarmente os efeitos da interação são fixos e são definidos como $\sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$.

As hipótese de interesse são definidas por:

$$\begin{cases} H_0 : \tau_1 = \tau_2 = ... = \tau_a = 0, \\ H_1 : \exists \tau_i \neq 0 \end{cases}$$

$$\begin{cases} H_0 : \beta_1 = \beta_2 = ... = \beta_b = 0, \\ H_1 : \exists \beta_j \neq 0 \end{cases}$$

$$\begin{cases} H_0 : (\tau\beta)_{ij} = 0, \quad \text{para todo } i, j \\ H_1 : \exists (\tau\beta)_{ij} \neq 0 \end{cases}$$

**O que fazer quando existe um fator de incômodo?**

Nessa situação é indicado utilizar Experimentos em Blocos!!!

O modelo com dois fatores em blocos é definido por:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \varepsilon_{ijk}, \tag{2}$$

em que $i = 1, 2, ..., a; j = 1, 2, ..., b; k = 1, 2, ..., n$; $\mu$ é a média geral, $\tau_i$ é o efeito do $i$-ésimo nível do fator $A$; $\beta_j$ é o efeito do $j$-ésimo nível do fator $B$; $(\tau\beta)_{ij}$ é o efeito da interação entre $\tau_i$ e $\beta_j$, $\delta_k$ é o efeito do $k$-ésimo bloco e $\varepsilon_{ijk}$ é o componente de erro aleatório.

- Dentro do bloco, a ordem em que as combinações de tratamento são executados é completamente aleatório.

- O modelo (2) assume que a interação entre blocos e tratamentos é insignificante;
- Essa suposição foi assumida anteriormente na análise de experimentos em blocos casualizados.

# ANOVA

TABELA 1: Tabela de Análise de Variância para dois fatores em Blocos casualizados

| Fonte de Variação | SQ | g.l. | QM | E[QM] | F |
|---|---|---|---|---|---|
| Bloco | $SQ_{Bloco}$ | $n-1$ | $QM_{Bloco}$ | $\sigma^2 + ab\sigma^2_\delta$ | |
| Fator $A$ | $SQ_A$ | a - 1 | $QM_A$ | $\sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$ | $\frac{QM_A}{QM_{Res}}$ |
| Fator $B$ | $SQ_B$ | b - 1 | $QM_B$ | $\sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$ | $\frac{QM_B}{QM_{Res}}$ |
| Interação $AB$ | $SQ_{AB}$ | (a - 1)(b-1) | $QM_{AB}$ | $\sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)}$ | $\frac{QM_{AB}}{QM_{Res}}$ |
| Resíduo | $SQ_{Res}$ | (ab - 1)(n-1) | $QM_{Res}$ | $\sigma^2$ | |
| Total | $SQ_T$ | $abn-1$ | | | |

# ANOVA

- As somas de quadrados são definidas por:

$$SQ_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

$$SQ_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn}$$

$$SQ_B = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn}$$

$$SQ_{AB} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn} - SQ_A - SQ_B$$

# ANOVA

$$SQ_{Bloco} = \frac{1}{ab}\sum_{k=1}^{n} y_{..k}^2 - \frac{y_{...}^2}{abn}$$

$$SQ_{Res} = SQ_T - (SQ_A + SQ_B + SQ_{AB} + SQ_{Bloco})$$

# Comparações de médias

- Caso seja verificado pela ANOVA a diferença entre as médias de tratamentos;
- Pode-se utilizar todas as técnicas de comparações múltiplas estudadas anteriormente;
- Porém, deve-se realizar as modificações necessárias em relação ao número de repetições e número de graus de liberdade.

# Diagnóstico do Modelo

- Verificar se os pressupostos:
  - Normalidade;
  - Independência;
  - Homogeneidade;

  são válidos utilizando as técnicas estudadas anteriormente.

**Comos escolher o tamanho da amostra?**

# Tamanho da amostra

- A probabilidade do erro tipo II do modelo de efeitos fixos é definida por:

$$\begin{aligned} \beta &= P\{\text{não rejeitar} H_0 | H_0 \text{é falsa}\} \quad (3) \\ &= P\{F_0 < F_{\text{crítico}} | H_0 \text{é falsa}\} \end{aligned}$$

# Tamanho da amostra

- Para avaliar a probabilidade do erro tipo II, definida na equação (3), é preciso conhecer a distribuição da estatística do teste $F_0$ se a hipótese nula for falsa;

- Porém, em experimentos fatorias temos mais tem um teste $F$ de interesse:

Caso 1 Teste $F$ para fator $A$;
Caso 2 Teste $F$ para fator $B$;
Caso 3 Teste $F$ para interação $AB$;

# Tamanho da amostra - Caso 1

- O parâmetro de não centralidade da distribuição $F$ pode ser obtido ao calcular:

$$E\left[\frac{SQ_A}{\sigma^2}\right],$$

- E esse parâmetro será igual a:

$$\delta_A = \frac{bn\sum_{i=1}^{a}\tau_i^2}{\sigma^2}, \tag{4}$$

- É possível observar que sob $H_0$, a equação (4) é igual a 0.

# Tamanho da amostra - Caso 2

- O parâmetro de não centralidade da distribuição $F$ pode ser obtido ao calcular:

$$E\left[\frac{SQ_B}{\sigma^2}\right],$$

- E esse parâmetro será igual a:

$$\delta_B = \frac{an\sum_{j=1}^{b}\beta_j^2}{\sigma^2}, \tag{5}$$

- É possível observar que sob $H_0$, a equação (5) é igual a 0.

# Tamanho da amostra - Caso 3

- O parâmetro de não centralidade da distribuição $F$ pode ser obtido ao calcular:

$$E\left[\frac{SQ_{AB}}{\sigma^2}\right],$$

- E esse parâmetro será igual a:

$$\delta_{AB} = \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{\sigma^2}, \qquad (6)$$

- É possível observar que sob $H_0$, a equação (6) é igual a 0.

**É possível experimento fatorial com dois fatores de incômodo?**

Nessa situação é indicado utilizar Experimentos em Quadro Latino!!!

- Em Experimentos em Quadro Latino temos duas restrições de aleatorização (linha e coluna), cada uma com $p$ níveis;
- Em Experimentos Fatoriais temos $abc...m$ combinações de tratamentos;
- Então, para poder executar um Experimentos Fatorial em Quadro Latino é preciso que $p = abc...m$;

- Considere uma modificação do experimento de detecção de alvo por radar;
- Os fatores neste experimento são o tipo de filtro (dois níveis) e a desordem do solo (três níveis), e operadores são considerados como blocos.
- Suponha agora que, devido ao tempo de configuração necessário, apenas seis operações podem ser feitas por dia.
- Assim, os dias tornam-se uma segunda restrição de aleatorização, resultando no experimento em quadrado latino $6 \times 6$.

FIGURE: Tabela 3 - Dados do nível de intensidade na detecção do alvo em Quadrado Latino

**Radar Detection Experiment Run in a 6 × 6 Latin Square**

| | Operator | | | | | |
|---|---|---|---|---|---|---|
| **Day** | **1** | **2** | **3** | **4** | **5** | **6** |
| 1 | $A(f_1g_1 = 90)$ | $B(f_1g_2 = 106)$ | $C(f_1g_3 = 108)$ | $D(f_2g_1 = 81)$ | $F(f_2g_3 = 90)$ | $E(f_2g_2 = 88)$ |
| 2 | $C(f_1g_3 = 114)$ | $A(f_1g_1 = 96)$ | $B(f_1g_2 = 105)$ | $F(f_2g_3 = 83)$ | $E(f_2g_2 = 86)$ | $D(f_2g_1 = 84)$ |
| 3 | $B(f_1g_2 = 102)$ | $E(f_2g_2 = 90)$ | $G(f_2g_3 = 95)$ | $A(f_1g_1 = 92)$ | $D(f_2g_1 = 85)$ | $C(f_1g_3 = 104)$ |
| 4 | $E(f_2g_2 = 87)$ | $D(f_2g_1 = 84)$ | $A(f_1g_1 = 100)$ | $B(f_1g_2 = 96)$ | $C(f_1g_3 = 110)$ | $F(f_2g_3 = 91)$ |
| 5 | $F(f_2g_3 = 93)$ | $C(f_1g_3 = 112)$ | $D(f_2g_1 = 92)$ | $E(f_2g_2 = 80)$ | $A(f_1g_1 = 90)$ | $B(f_1g_2 = 98)$ |
| 6 | $D(f_2g_1 = 86)$ | $F(f_2g_3 = 91)$ | $E(f_2g_2 = 97)$ | $C(f_1g_3 = 98)$ | $B(f_1g_2 = 100)$ | $A(f_1g_1 = 92)$ |

- Na Tabela 3 as letras minúsculas $f_j$ e $g_k$ são usadas para representar a combinação dos tratamentos. Em que $f_j$ representa o $j$-ésimo nível de tipo de filtro e $g_k$ representa o $k$-ésimo nível de agrupamento de solo.

O modelo estatístico para esse experimento é definido por:

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + (\tau\beta)_{jk} + \theta_l + \varepsilon_{ijkl}, \tag{7}$$

em que $i = 1, 2, ..., p, j = 1, 2, ..., a; k = 1, 2, ..., b; l = 1, 2, ..., p;$ $\mu$ é a média geral, $alpha_i$ é o efeito da $i$-ésima linha, $\tau_j$ é o efeito do $j$-ésimo nível do fator $A$; $\beta_k$ é o efeito do $k$-ésimo nível do fator $B$; $(\tau\beta)_{jk}$ é o efeito da interação entre $\tau_j$ e $\beta_k$, $\theta_l$ é o efeito da $l$-ésima coluna e $\varepsilon_{ijk}$ é o componente de erro aleatório.