Delineamento e Análise de Experimentos

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

Brasília - 2023

ANOVA

- Para o modelo de efeitos (ANOVA):

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, ..., a; \quad j = 1, 2, ..., n \qquad (1)$$

  em que μ é a média geral, τ_i é a média ou efeito dos tratamentos e ε_{ij} componente de erro aleatório.

- As hipótese de interesse são definidas por:

$$\begin{cases} H_0 : \tau_1 = \tau_2 = ... = \tau_a = 0, & \text{(O efeito de tratamento é nulo)} \\ H_1 : \exists \tau_i \neq 0 \end{cases}$$

- Ou de forma equivalente:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = ... = \mu_a \\ H_1 : \exists \mu_i \neq \mu_j, \; i \neq j, \end{cases}$$

  em que $\mu_i = \mu + \tau_i$.

Anova

- Só é possível realizar a análise de variância se certas condições, ou seja, certas exigências do modelo matemático forem satisfeitas:
  - Os erros devem ter distribuição normal;
  - Os erros devem ser independentes;
  - Os erros deve ter a mesma variância, ou seja, deve existir homocedasticidade.
- Pode-se representar essas condições por:

$$\varepsilon_{ij} \sim N(0, \sigma^2). \tag{2}$$

ANOVA

$$E(QM_{Res}) = E\left(\frac{SQ_{Res}}{N-a}\right) = \frac{1}{N-a}E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y_{i.}})^2\right] = \sigma^2$$

e

$$E(QM_{Trat}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

- Ou seja, $QM_{Res} = SQ_{Res}/(N-a)$ estima σ^2;
- E $QM_{Trat} = SQ_{Trat}/(a-1)$ também estima σ^2, caso não há diferença entre as médias dos tratamentos (o que implica em $\tau_i = 0$).

ANOVA

- Como os graus de liberdade de SQ_{Trat} e SQ_{Res} somam $N-1$, que é o número de graus de liberdade da SQ_T, o Teorema de Cochran estabelece que SQ_{Trat}/σ^2 e SQ_{Res}/σ^2 são variáveis aleatórias independentes.

- Portanto, se a hipótese nula de nenhuma diferença nas médias de tratamento for verdadeira, a razão

$$F_0 = \frac{SQ_{Trat}/(a-1)}{SQ_{Res}/(N-a)} = \frac{QM_{Trat}}{QM_{Res}}, \qquad (3)$$

segue distribuição F com $a-1$ e $N-a$ graus de liberdade.

- A equação (3) é a estatística do teste para a hipótese de não haver diferenças nas médias de tratamento.

ANOVA

- A partir da esperança dos quadrados médios, verifica-se que, em geral, QM_{Res} é um estimador imparcial de σ^2;
- Além disso, sob a hipótese nula ser verdadeira, QM_{Trat} também é um estimador imparcial de σ^2;
- No entanto, se a hipótese nula for falsa, o valor esperado de QM_{Trat} é maior que σ^2;
- Portanto, sob a hipótese alternativa, o valor esperado do numerador da estatística de teste (3) é maior que o valor esperado do denominador, e deve-se rejeitar H_0 para valores da estatística do teste que são muito grandes;
- Portanto, deve-se rejeitar H_0 e concluir que existem diferenças nas médias de tratamento se $F_0 > F_{\alpha,(a-1),(N-a)}$.

ANOVA

Table: Tabela de Análise de Variância

Fonte de Variação	SQ	g.l.	QM	F
Tratamentos	SQ_{Trat}	a - 1	QM_{Trat}	QM_{Trat}/QM_{Res}
Resíduo	SQ_{Res}	N - a	QM_{Res}	
Total	SQ_T	N - 1		

em que $N = an$.

Diagnóstico

- É fundamental verificar a adequabilidade do modelo antes de tirar conclusões sobre o processo inferencial.

- Verificar as características do modelo.

- Para avaliar a adequabilidade do modelo serão usados:
  - Métodos gráficos
  - Testes estatísticos

Suposição de Normalidade

- Pode ser inicialmente verificada ao construir um histograma dos resíduos. Se a suposição sobre os erros for satisfeita, este gráfico deve se parecer com uma amostra de uma distribuição normal centrada em zero;
- Outro recurso gráfico importante é o **Gráfico de Probabilidade Normal dos resíduos** - cada resíduo é colocado em um gráfico versus seu valor esperado sob normalidade.
- Se o resultado gráfico for aproximadamente linear, há evidências da suposição de normalidade.
- Se o resultado gráfico afasta substancialmente da linearidade, não há evidências de que a distribuição do erro é normal.
- Teste de hipótese também pode ser utilizado, como o teste de Shapiro-Wilk.

Suposição de independência

- O objetivo é verificar se existe qualquer correlação entre os erros que são próximos um dos outros em uma sequência;
- Quando os erros são independentes espera-se que os resíduos versus a ordem das observações flutuem em um padrão mais ou menos aleatório em torno da linha de referência Zero.

Variância constante

- Gráficos dos resíduos versus valores ajustados podem ser usados para verificar se a variância dos erros é constante. Pois, os sinais dos resíduos não são importantes para examinar a constância da variância do erro;

- Testes de hipóteses.

Variância constante

Para verificar as hipóteses:

$$\begin{cases} H_0 : \sigma_1 = \sigma_2 = ... = \sigma_a = 0 \\ H_1 : \exists \sigma_i \neq 0 \end{cases}$$

pode-se utilizar os seguintes testes de hipóteses:

- F;
- Bartlett, e
- Levene

Estimação dos Parâmetros do Modelo

- Após verificar a qualidade de ajuste do modelo, o próximo passo é estimar os parâmetos do modelo;
- Ao considerar o modelo de efeitos:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, ..., a; \quad j = 1, 2, ..., n$$

  em que μ é a média geral, τ_i é a média ou efeito dos tratamentos e ε_{ij} componente de erro aleatório.

- Os parâmetros a serem estimados são: μ e τ_i;
- O método de estimação de Mínimos Quadrados será utilizado.

Estimação dos Parâmetros do Modelo

- Para encontrar os estimadores de mínimos quadrados de μ e τ_i, é necessário escrever a soma dos quadrados dos erros:

$$L = \sum_{i=1}^{a}\sum_{j=1}^{n}\varepsilon_{ij}^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \mu - \tau_i)^2 \qquad (4)$$

  e os valores de μ e τ_i que minimizam a equação (4) são os estimadores de mínimos quadrados, $\hat{\mu}$ e $\hat{\tau}_i$.
- Os valores apropriados seriam as soluções para as $a + 1$ equações simultâneas:

$$\frac{\partial L}{\partial \mu}|\hat{\mu}, \hat{\tau}_i = 0$$

  e

$$\frac{\partial L}{\partial \tau_i}|\hat{\mu}, \hat{\tau}_i = 0 \quad i = 1, 2, ..., a.$$

Estimação dos Parâmetros do Modelo

- Derivando a equação (4) em relação a μ e τ_i e igualando a zero, tem-se que:

$$-2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0$$

e

$$-2\sum_{j=1}^{n}(y_{ij} + \hat{\mu} - \hat{\tau}_i) = 0 \quad i = 1, 2, ..., a.$$

- Ao aplicar a restrição: $\sum_{i=1}^{a}\hat{\tau}_i = 0$, tem-se a seguinte solução para o sistema de equações normais:

$$\hat{\mu} = \bar{y}_{..}$$

e

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

Estimação dos Parâmetros do Modelo

- Um intervalo de confiança para a média do i-ésimo tratamento pode ser facilmente determinada;
- A média do i-ésimo tratamento é

$$\mu_i = \mu + \tau_i$$

- Então, um estimador pontual para μ_i é dado por:

$$\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i.}$$

- Ao assumir que a distribuição dos erro é normal, a média de cada tratamento, $\bar{y}_{i.}$ tem distribuição $N(\mu_i, \sigma^2/n)$;
- Se σ^2 fosse conhecido, a distribuição normal seria usada para construir o intervalo de confiança;

Estimação dos Parâmetros do Modelo

- Ao usar o QM_{Res} como estimador de σ^2, deve-se usar a distribuição t-Student para obter o intervalo de confiança;
- Sendo assim, o intervalo $100(1-\alpha)\%$ para a i-ésima média de tratamento, μ_i, é definido por:

$$\bar{y}_{i.} - t_{\alpha/2,N-a}\sqrt{\frac{QM_{Res}}{n}} \leq \mu_i \leq \bar{y}_{i.} + t_{\alpha/2,N-a}\sqrt{\frac{QM_{Res}}{n}} \quad (5)$$

- Quando o pesquisador quer construir o intervalo de confiança para várias médias de tratamento, deve-se fazer uma correção no nível de confiança usando o método de Bonferroni;
- Nesse caso deve-se usar $\alpha/(2r)$ na equação (5), em que r é a quantidade de intervalos de confiança simultâneos.

Comparações de médias

- Considere o experimento da gravação de plasma. Como a hipótese nula foi rejeitada, sabemos que algumas configurações de energia produzem taxas de gravação diferentes das outras, mas quais realmente causam essa diferença?

- Como verificar quais configurações de energia produzem taxas de gravação diferentes das outras?

Comparações de médias

Alguns teste podem ser utilizados para tal objetivo. São eles:

- Teste de Fisher;
- Teste de Tukey;
- Teste de Duncan;
- Teste de Scheffé.

Teste de Fisher

- Considere a hipótese $H_0 : \mu_i = \mu_j$, para $i \neq j$;
- A estatística do teste é definida por:

$$t_0 = \frac{\bar{y_i}. - \bar{y_j}.}{\sqrt{QM_{Res}(\frac{1}{n_i} + \frac{1}{n_j})}} \tag{6}$$

- Supondo uma hipótese alternativa bilateral ($H_1 : \mu_i \neq \mu_j$), o par de médias μ_i e μ_j terá diferença significativa se $|\bar{y_i}. - \bar{y_j}.| > t_{\alpha/2,N-a}\sqrt{QM_{Res}(1/n_i + 1/n_j)}$.
- A quantidade

$$LSD = t_{\alpha/2,N-a}\sqrt{QM_{Res}(1/n_i + 1/n_j)}, \tag{7}$$

é conhecida como diferença mínima significativa.

Teste de Fisher

- Ao usar o procedimento LSD de Fisher, simplesmente comparamos a diferença observada entre cada par de médias com o LSD correspondente;

- É possível observar que o procedimento LSD de Fisher controla a taxa de erro para cada comparação par a par individual, mas não controla a taxa de erro experimental.

Teste de Tukey

- Suponha que, após rejeitar a hipótese nula de igualdade de tratamento na ANOVA, deseja-se testar todas as comparações de médias:

$$
\begin{aligned}
H_0 &: \mu_i = \mu_j \\
H_1 &: \mu_i \neq \mu_j
\end{aligned}
$$

  para todo $i \neq j$.

- O procedimento de Tukey controla o erro experimental no nível selecionado. Este é um excelente procedimento de espionagem de dados quando o interesse está em pares de médias;

Teste de Tukey

- O teste de Tukey considerar a seguinte estatística de intervalo estudentizado:
$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{QM_{Res}/n}},$$
em que $\bar{y}_{max}$ e $\bar{y}_{min}$ são a maior e menor média, respectivamente, para o grupo de p amostras de médias.
- O valor de q deve ser comparado com valores de $q(p, f)$, em que $q()$ é o percentil da estatística de intervalo estudentizado;
- Em que f é o número de graus de liberdade associados ao QM_{Res};
- Para amostras de mesmo tamanho, o teste de Tukey declara duas médias significativamente diferentes se o valor absoluto de suas diferenças amostrais excedem:

$$T_\alpha = q_\alpha(a, f)\sqrt{\frac{QM_{Res}}{n}}. \tag{8}$$

Teste de Tukey

- De forma equivalente, pode-se construir um conjunto de $100(1-\alpha)\%$ intervalos de confiança para todos os pares de médias, ou seja:

$$\bar{y_i}. - \bar{y_j}. - q_\alpha(a,f)\sqrt{\frac{QM_{Res}}{n}} \leq \mu_i - \mu_j \leq \bar{y_i}. - \bar{y_j}. + q_\alpha(a,f)\sqrt{\frac{QM_{Res}}{n}}$$

Contrastes

- Considere o experimento da gravação de plasma. Como a hipótese nula foi rejeitada, sabemos que algumas configurações de energia produzem taxas de gravação diferentes das outras, mas quais realmente causam essa diferença?
- Podemos suspeitar no início do experimento que $200W$ e $220W$ produzem o mesma taxa de gravação, o que implica que gostaríamos de testar a hipótese:

$$H_0 : \mu_3 = \mu_4$$
$$H_1 : \mu_3 \neq \mu_4$$

ou equivalentemente:

$$H_0 : \mu_3 - \mu_4 = 0$$
$$H_1 : \mu_3 - \mu_4 \neq 0$$

Contrastes

- Ou, se tivéssemos suspeitado no início do experimento que a média dos níveis mais baixos de potência não diferia da média dos níveis mais altos de potência, então a hipótese teria sido:

$$H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4$$
$$H_1 : \mu_1 + \mu_2 \neq \mu_3 + \mu_4$$

ou

$$H_0 : \mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$$
$$H_1 : \mu_1 + \mu_2 - \mu_3 - \mu_4 \neq 0$$

Contrastes

- Muitos métodos de comparação múltipla usam a ideia de um contraste.
- Em geral, um contraste é uma combinação linear dos parâmetros na forma:

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i,$$

  em que a soma das constantes do contraste é igual a zero, isto é, $\sum_{i=1}^{a} c_i = 0$.
- Ambas as hipóteses acima podem ser expressas em termos dos contrates:

$$H_0 : \sum_{i=1}^{a} c_i \mu_i = 0$$

$$H_1 : \sum_{i=1}^{a} c_i \mu_i \neq 0$$

Qual a estatística do Teste nesse caso?

Contrastes

- Qual a estatística do Teste nesse caso?

- Há dois caminhos: teste t e teste F

CONTRASTES

- Primeiro vamos usar o teste t;
- Então, é preciso escrever o contraste de interesse em termos das médias de tratamento. Logo, tem-se que:

$$C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$$

- A variância de C é:

$$V(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2, \tag{9}$$

quando todos os tratamentos têm o mesmo tamanho.

Contrastes

- Sob a hipótese nula ser verdadeira, a estatística do teste é definida por:

$$\frac{\sum_{i=1}^{a} c_i \bar{y}_{i.} - 0}{\sqrt{\frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2}},$$

ou

$$\frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2}},$$

com distribuição $N(0, 1)$.

Contrastes

- Agora vamos substituir a variância desconhecida, σ^2, por sua estimativa, o erro quadrático médio e a estatística é definida por:

$$t_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{QM_{Res}}{n} \sum_{i=1}^{a} c_i^2}}, \tag{10}$$

- A hipótese nula será rejeitada se $|t_0| > t_{\alpha/2, N-a}$.

Contrastes

- A segunda abordagem usa o teste F;

- Então, ao elevar ao quadrado uma variável aleatória t com ν graus de liberdade, tem-se uma variável aleatória F com 1 grau de liberdade do numerador e ν graus de liberdade do denominador. Portanto, tem-se que:

$$F_0 = t_0^2 = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{QM_{Res}}{n} \sum_{i=1}^{a} c_i^2}. \tag{11}$$

- A hipótese nula será rejeitada se $F_0 > F_{\alpha,1,N-a}$.

Contrastes

- A estatística do teste (11) pode ser escrita como:

$$F_0 = \frac{QM_C}{QM_{Res}} = \frac{SQ_C/1}{QM_{Res}},$$

em que a soma de quadrados do contraste é:

$$SQ_C = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{1}{n}\sum_{i=1}^{a} c_i^2}, \tag{12}$$

e possui 1 grau de liberdade.

Contrastes

- Em vez de realizar teste de hipótese para o contraste, também é possível construir um Intervalo de Confiança;

- Considere o contraste de interesse:

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

- Ao substituir as médias dos tratamentos por seu respectivo estimador, tem-se que

$$C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$$

Contrastes

- Como

$$E(\sum_{i=1}^{a} c_i \bar{y}_{i.}) = \sum_{i=1}^{a} c_i \mu_i \qquad e \qquad V(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$$

- O intervalo de confiança para o contraste $\sum_{i=1}^{a} c_i \mu_i$ é:

$$IC = \left(\sum_{i=1}^{a} c_i \bar{y}_{i.} \pm t_{\alpha/2, N-a} \sqrt{\frac{QM_{Res}}{n} \sum_{i=1}^{a} c_i^2} \right). \qquad (13)$$

Contrastes

- Um caso especial de contraste é o contraste ortogonal;
- Dois contraste com coeficientes c_i e d_i são ortogonais se:

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

- Para a tratamentos, o conjunto de $a-1$ contrastes ortogonais particiona a soma de quadrados de tratamentos em $a-1$ componentes independentes com um único grau de liberdade para cada componente.
- Assim, os testes realizados com contrastes ortogonais são independentes.

Coeficiente de Determinação

- A medida R^2 é chamada de **Coeficiente de Determinação** e representa a proporção da variação total explicada pelo modelo ANOVA e é definida por:.

$$R^2 = \frac{SQ_{Modelo}}{SQ_T}$$

- Desde de que $0 \leq SQ_{Modelo} \leq SQ_T$, segue que

$$0 \leq R^2 \leq 1$$

- Valores grandes de R^2 indicam que a variação total é mais reduzida/explicada pelo efeito de tratamento.

Qual o tamanho da amostra?

- Em qualquer problema de planejamento experimental, uma decisão crítica é a escolha do tamanho da amostra, isto é, determinar o número de réplicas a serem executadas;

- Geralmente, se o pesquisador está interessado na detecção de pequenos efeitos, são necessárias mais réplicas do que se o pesquisador estiver interessado em detecção de grandes efeitos.

Tamanho da amostra

- A probabilidade do erro tipo II do modelo de efeitos fixos para o caso de amostras de tamanhos iguais para cada tratamento, é definida por:

$$\begin{aligned} \beta &= P\{\text{não rejeitar} H_0 | H_0 \text{é falsa}\} \qquad (14) \\ &= P\{F_0 < F_{\text{crítico}} | H_0 \text{é falsa}\} \end{aligned}$$

Tamanho da amostra

- Para avaliar a probabilidade do erro tipo II, definida na equação (14), é preciso conhecer a distribuição da estatística do teste F_0 se a hipótese nula for falsa;

- Pode-se mostrar que, se H_0 for falsa, a estatística $F_0 = QM_{Trat}/QM_{Res}$ tem distribuição F não central com $a-1$ e $N-a$ graus de liberdade e parâmetro de não centralidade δ;

- Se $\delta = 0$, a distribuição F não central torna-se a distribuição F usual (central).

Tamanho da amostra

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

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

- E esse parâmetro será igual a:

$$\delta = \frac{n\sum_{i=1}^{a}\tau_i^2}{\sigma^2}, \tag{15}$$

- É possível observar que sob H_0, a equação (15) é igual a 0.

Tamanho da amostra

- O pesquisador deve especificar os valores de τ e σ^2;
- A estimativa de σ^2 pode estar disponível a partir de uma experiência anterior, um experimento anterior ou uma estimativa de julgamento.
- Ao calcular δ, β e o poder do teste $(1 - \beta)$ para diferentes valores de n, é possível encontrar o tamanho da amostra.