

ME720 - Modelos Lineares Generalizados

Parte 4 - Teste de hipótese em modelos normais lineares

Profa. **Larissa Avila Matos**

ANOVA

Exemplo 1

No primeiro modelo, o interesse primário, de certa forma, é testar se a carga não contribui para explicar o consumo de oxigênio, ou seja:

$$H_0 : \beta_1 = 0 \text{ vs } H_1 : \beta_1 \neq 0, i = 1, \dots, p$$

No modelo 2, temos igual interesse. Contudo, neste caso, queremos verificar a existência de tal contribuição para todos os grupos. Ou seja, desejamos testar:

$$H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = 0 \text{ vs } H_1 : \text{há pelo menos uma diferença}$$

Exemplo 2

Para o exemplo 2, temos interesse em testar se as médias dos níveis de absorbância dos solventes são iguais, ou seja, testar se:

$$H_0 : \alpha_2 = \alpha_3 = \alpha_4 = 0 \text{ vs } H_1 : \text{há pelo menos uma diferença}$$

Algumas dessas hipóteses podem (e devem) ser testadas antes de nos concentrarmos em outras hipóteses.

A primeira e a terceira podem ser testadas a partir da decomposição das somas de quadrados (ANOVA).

A segunda, requer uma abordagem mais específica.

ANOVA para modelos de regressão

Suponha o seguinte modelo:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_{p-1} x_{(p-1)i} + \xi_i, \quad \xi_i \stackrel{i.i.d.}{\sim} N(0, \sigma^2).$$

Logo, $Y_i \stackrel{ind.}{\sim} N(\beta_0 + \sum_{j=1}^{p-1} \beta_j x_{ji}, \sigma^2)$.

O modelo acima define uma média (condicional aos valores de x_{ji} , $j = 1, \dots, p-1$; $i = 1, \dots, n$) para cada observação Y_i .

Defina $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_{p-1} x_{(p-1)i}$ (valor predito pelo modelo).

O resíduo é definido por $R_i = \hat{\xi}_i = Y_i - \hat{Y}_i$.

Nosso objetivo é considerar um modelo que explique adequadamente a variabilidade dos dados, ou seja, um modelo para o qual os resíduos sejam “pequenos”.

A soma de quadrados dos resíduos é dada por $SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$.

Lema: $\sum_{i=1}^n R_i = \sum_{i=1}^n Y_i - \hat{Y}_i = 0$.

Prova: ??

Corolário: $\frac{1}{n} \sum_{i=1}^n \hat{Y}_i = \bar{Y}$.

Prova: ??

Pode-se provar que, a soma de quadrados total $SQT = \sum_{i=1}^n (Y_i - \bar{Y})^2$, pode ser decomposta como:

$$SQT = \underbrace{\sum_{i=1}^n (\bar{Y} - \hat{Y}_i)^2}_{SQM} + \underbrace{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}_{SQR}.$$

Assim, quanto maior for o valor de SQM em relação à SQR, maior será a contribuição da parte sistemática para explicar a variabilidade dos dados. Portanto, mais “provável” que exista (pelo menos um)

$$\beta_j \neq 0, \quad j = 1, \dots, p-1.$$

SQM: soma de quadrados do modelo ; SQR: soma de quadrados dos resíduos.

Portanto, como estatística de teste, podemos comparar, de alguma forma, as somas de quadrado (do modelo e dos resíduos).

Pergunta: como construir uma estatística de teste adequada?

Adequada: que serve para testar as hipóteses de interesse, que possua distribuição conhecida (sob H_0 e sob H_1) e que tenha um “razoável” poder (assumindo-se ser possível fixar o nível de significância α).

Lembrando da forma matricial do modelo: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$, pode-se demonstrar que:

$$\text{SQT} = \mathbf{Y}' (\mathbf{I} - n^{-1}\mathbf{J}) \mathbf{Y}, \text{ em que } \mathbf{J} = \mathbf{1}\mathbf{1}'.$$

$$\text{SQM} = \mathbf{Y}' (\mathbf{H} - n^{-1}\mathbf{J}) \mathbf{Y} = \mathbf{Y}' \mathbf{A} \mathbf{Y}, \text{ em que } \mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

$$\text{SQR} = \mathbf{Y}' (\mathbf{I} - \mathbf{H}) \mathbf{Y} = \mathbf{Y}' \mathbf{B} \mathbf{Y}.$$

Pode-se provar que as matrizes $\mathbf{A} = \mathbf{H} - n^{-1}\mathbf{J}$ e $\mathbf{B} = \mathbf{I} - \mathbf{H}$ são ortogonais, ou seja, $\mathbf{AB} = \mathbf{0}$ (provando-se que $n^{-1}\mathbf{HJ} = n^{-1}\mathbf{J}$).

Dizemos que $\hat{\mathbf{Y}}$ e \mathbf{R} projetam \mathbf{Y} em dois subespaços ortogonais.

Um caminho para a construção de estatísticas do teste é o estudo das propriedades de formas quadráticas aleatórias (normais).

Distribuição de formas quadráticas

Distribuição qui-quadrado não central

Sejam $X_1, X_2, \dots, X_n \stackrel{ind}{\sim} N(\mu_i, \sigma_i^2)$.

Defina $Y = \sum_{i=1}^n \left(\frac{X_i}{\sigma_i}\right)^2$. Dizemos então que Y tem distribuição qui-quadrado não central com n graus de liberdade e parâmetro de não-centralidade $\delta = \sum_{i=1}^n \left(\frac{\mu_i}{\sigma_i}\right)^2$.

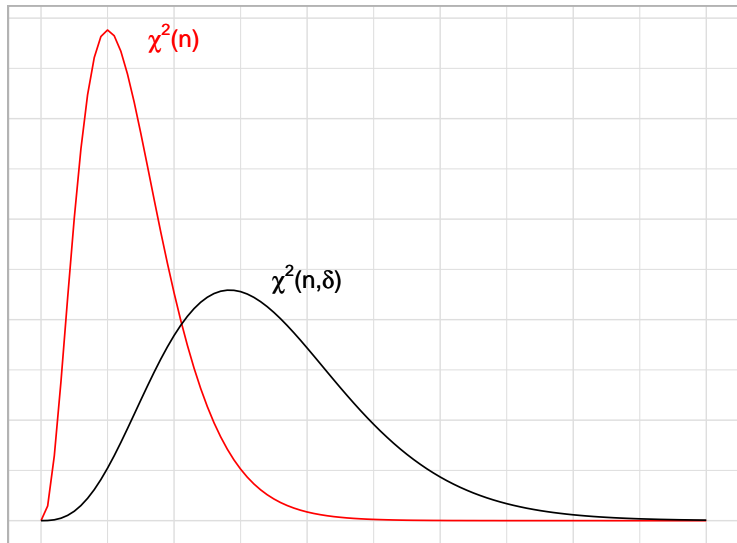
Notação: $Y \sim \chi_{(n,\delta)}^2$, cuja fdp é dada por

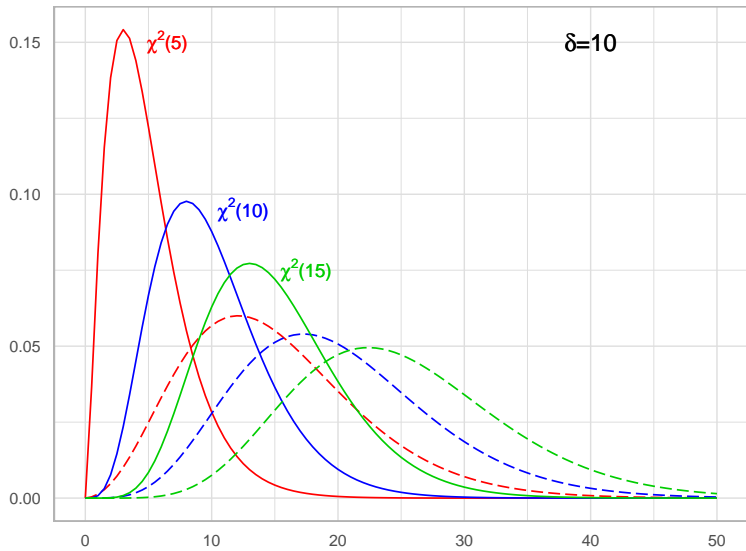
$$f_Y(y) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} f_{W_{n+2i}}(y) \mathbb{1}_{(0,\infty)}(y),$$

em que $W_{n+2i} \sim \chi_{(n+2i)}^2$.

Se $\delta = 0$, então $Y \sim \chi_{(n)}^2$.

Densidades da qui-quadrado central e não central





Distribuição F não central

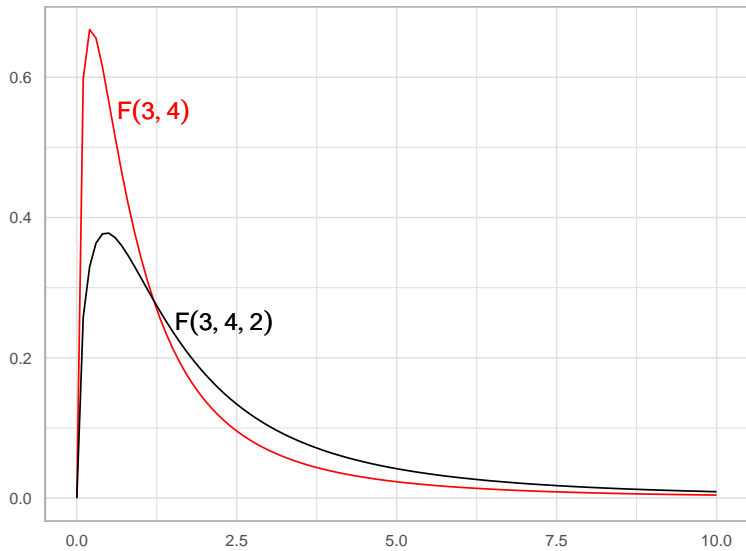
Seja V uma outra v.a., independente de Y , $V \sim \chi^2_{(m)}$.

Defina $F = \frac{Y/n}{V/m}$. Então, F tem distribuição F não central com graus de liberdade, n e m e parâmetro de não centralidade δ .

Notação: $F \sim F_{(n,m,\delta)}$, cuja fdp é dada por

$$\begin{aligned} f_F(f) &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{\beta(m/2, n/2 + i) i!} \left(\frac{n}{m}\right)^{n/2+i} \left(\frac{m}{m + nf}\right)^{(n+m)/2+i} \\ &\times f^{n/2-1+i} \mathbb{1}_{(0,\infty)}(f), \end{aligned}$$

em que $\beta(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$.



Distribuição F não central dupla

Sejam U e V variáveis aleatórias independentes tais que $U \sim \chi^2_{(n, \delta_1)}$ e $V \sim \chi^2_{(m, \delta_2)}$.

Defina,

$$F = \frac{U/n}{V/m}.$$

Então, F tem distribuição F não central dupla com n graus de liberdade no numerador, m graus de liberdade no denominador e parâmetros de não centralidade δ_1 e δ_2 .

Notação: $F \sim F_{(n, m, \delta_1, \delta_2)}$.

Distribuição t de Student não-central

Defina $T = \frac{Z+\mu}{\sqrt{V/\nu}}$, em que $Z \perp V$, $Z \sim N(0, 1)$ e $V \sim \chi^2_\nu$. Dizemos que T tem distribuição t de student não central, com ν graus de liberdade e parâmetro de não centralidade μ .

Uma forma de se apresentar a densidade é:

$$\begin{aligned} f_T(t) &= \frac{\nu^{\nu/2} e^{-\frac{\nu\mu^2}{2(t^2+\nu)}}}{\sqrt{\pi}\Gamma(\nu/2)2^{\frac{\nu-1}{2}}(t^2 + \nu)^{(\nu+1)/2}} \\ &\times \int_0^\infty y^\nu \exp\left\{-\frac{1}{2}\left(y - \frac{\mu t}{\sqrt{t^2 + \nu}}\right)^2\right\} dy \end{aligned}$$

Distribuição de formas quadráticas normais

Seja $\mathbf{Z} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dizemos, então, que $Y = \mathbf{Z}'\mathbf{A}\mathbf{Z}$ é uma forma quadrática normal. Em geral, \mathbf{A} é uma matriz não aleatória simétrica e real.

Pode-se provar, por exemplo, que:

$$Y = (\mathbf{Z} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Z} - \boldsymbol{\mu}) \sim \chi_{(p)}^2. \quad (1)$$

Vamos demonstrar a validade do resultado dado pela equação (1).

Por simplicidade, admita que $\boldsymbol{\mu} = \mathbf{0}$.

Assim, temos que (fazendo $\widehat{\boldsymbol{\theta}} = \mathbf{z}$)

$$\begin{aligned}
 M_Y(t) &= \mathcal{E}(e^{Yt}) = \int_{\mathcal{R}^p} e^{t'y} f_Z(\mathbf{z}) d\mathbf{z} \\
 &= \int_{\mathcal{R}} \int_{\mathcal{R}} \dots \int_{\mathcal{R}} e^{t\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}} |\boldsymbol{\Sigma}|^{-1/2} (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} \mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z} \right\} d\mathbf{z} \\
 &= \int_{\mathcal{R}} \int_{\mathcal{R}} \dots \int_{\mathcal{R}} |\boldsymbol{\Sigma}|^{-1/2} (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} \mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}(1-2t) \right\} d\mathbf{z}
 \end{aligned}$$

Considerando a mesma transformação anteriormente utilizada ($\mathbf{W} = \boldsymbol{\Psi}^{-1}\mathbf{z}$), em que $\boldsymbol{\Sigma} = \boldsymbol{\Psi}\boldsymbol{\Psi}'$, temos que:

$$\begin{aligned}
 M_Y(t) &= \prod_{i=1}^p (1-2t)^{-1/2} \int_{\mathcal{R}} \underbrace{\frac{(1-2t)^{-1/2}}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} w_i^2 (1-2t) \right\}}_{N(0, (1-2t)^{-1})} dw_i \\
 &= \prod_{i=1}^p (1-2t)^{-1/2} = (1-2t)^{-p/2}
 \end{aligned}$$

A qual corresponde à fgm de uma distribuição χ_p^2 .

Principais teoremas de formas quadráticas normais

1 Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \mathbf{I})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A})]}$$

se, e somente se \mathbf{A} for idempotente.

2 Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \mathbf{\Sigma})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A})]}$$

se, e somente se $\mathbf{A}\mathbf{\Sigma}$ for idempotente.

3 Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \mathbf{I})$ e \mathbf{A} uma matriz simétrica. Então,

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A}), \delta=\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}]}$$

se, e somente se \mathbf{A} for idempotente.

4 Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A}), \delta=\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}]}$$

se, e somente se $\mathbf{A}\boldsymbol{\Sigma}$ for idempotente.

Obs. Naturalmente, se $\delta = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = 0$, as distribuições passam a ser qui-quadrados centrais.

- 5 Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} uma matriz simétrica e \mathbf{B} uma matriz qualquer.

Então, $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ e $\mathbf{B}\mathbf{Y}$ são independentes se, e somente se

$$\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{0}.$$

- 6 Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, \mathbf{A} e \mathbf{B} matrizes simétricas.

Então, $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ e $\mathbf{Y}'\mathbf{B}\mathbf{Y}$ são independentes se, e somente se

$$\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{0}.$$

Demonstrações: pesquisar nas referências.

Teorema de Cochran

Seja $\mathbf{X} \sim N_n(\mathbf{0}, \mathbf{I})$ e suponha que

$$Q = \mathbf{X}'\mathbf{X} = \sum_{i=1}^n X_i^2 = Q_1 + Q_2 + \dots + Q_k,$$

em que

$$Q_i = \mathbf{X}'\mathbf{A}_i\mathbf{X}, i = 1, 2, \dots, k,$$

é uma forma quadrática de posto r_i , i.e., $\text{rank}(\mathbf{A}_i) = r_i$, e \mathbf{A}_i é simétrica para todo i , $i = 1, 2, \dots, k$. Então, as seguintes condições são equivalentes:

1 $Q_i \perp Q_j, \forall i \neq j.$

2 $Q_i \sim \chi_{(r_i)}^2.$

3 $\sum_{i=1}^k r_i = n.$

Voltando ao problema

Lembremos que:

$SQM = \mathbf{Y}' (\mathbf{H} - n^{-1} \mathbf{J}) \mathbf{Y}$, em que $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, e

$SQR = \mathbf{Y}' (\mathbf{I} - \mathbf{H}) \mathbf{Y}$,

com $\mathbf{A} = \mathbf{H} - n^{-1} \mathbf{J}$ e $\mathbf{B} = \mathbf{I} - \mathbf{H}$.

Sabemos que $\mathbf{Y} \sim N_n(\mathbf{X}\beta, \sigma^2 \mathbf{I})$.

Defina $W = SQM/\sigma^2 = \mathbf{Y}' \left(\frac{\mathbf{H} - n^{-1} \mathbf{J}}{\sigma^2} \right) \mathbf{Y}$ e

$V = SQR/\sigma^2 = \mathbf{Y}' \left(\frac{\mathbf{I} - \mathbf{H}}{\sigma^2} \right) \mathbf{Y}$.

Postulamos que $\frac{\mathbf{A}}{\sigma^2}(\sigma^2 \mathbf{I}) \frac{\mathbf{B}}{\sigma^2} = \frac{1}{\sigma^2} \mathbf{AB} = \mathbf{0}$ (são ortogonais).

Então, $W \perp V$.

Como \mathbf{A} e \mathbf{B} são matrizes idempotentes, logo \mathbf{AI} e \mathbf{BI} , também o são.

Pode-se provar que $\mathcal{E}(V) = n - p$ e que, sob H_0 , $\mathcal{E}(W) = p - 1$ (parâmetros de não centralidade são iguais à 0).

Assim, $W \sim \chi^2_{(p-1)}$ (sob H_0) e $V \sim \chi^2_{(n-p)}$, em que $W \perp V$.

Logo, $F = \frac{W/(p-1)}{V/(n-p)} \sim F_{[(p-1), (n-p)]}$, sob H_0 .

Ainda, sob H_1 , $F \sim F_{[(p-1), (n-p), \delta = \frac{1}{\sigma^2} (\beta' \mathbf{X}' \mathbf{X} \beta - n^{-1} \beta' \mathbf{X}' \mathbf{J} \mathbf{X} \beta)]}$.

Assim, ao se optar por usar formas quadráticas para se testar hipóteses, é necessário verificar:

- A distribuição de \mathbf{Y} .
- As propriedades das matrizes núcleo (como a matriz de covariâncias de \mathbf{Y}).
- As esperanças das formas quadráticas (sob H_0 e H_1).

Portanto, rejeita-se H_0 se $f_c > f_{critico}$ ou, analogamente, se $pvalor = P(F > f_c | H_0) < \alpha$, em que f_c é o valor calculado da estatística F e $\alpha = P(F > f_{critico} | H_0)$.

Tabela de ANOVA (matricial)

Para testar $H_0 : \beta_1 = \beta_2 = \dots = \beta_{(p-1)} = 0$ vs $H_1 : \text{Há pelo menos uma diferença.}$

FV	SQ	GL	QM	$\mathcal{E}(QM)$	Estat. F	p-valor
Modelo	$SQM = \mathbf{Y}'\mathbf{A}\mathbf{Y}$	p-1	$= \frac{SQM}{p-1}$ (QMM)	$\sigma^2 + \frac{1}{p-1} \boldsymbol{\beta}' \mathbf{X}' (\mathbf{I} - \mathbf{n}^{-1} \mathbf{J}) \mathbf{X} \boldsymbol{\beta}$	$\frac{QMM}{QMR}$	p^*
Resíduo	$SQR = \mathbf{Y}'\mathbf{B}\mathbf{Y}$	n-p	$\frac{SQR}{n-p}$ (QMR)	σ^2		
Total	SQT	n-1				

$p^* = \min(P(F > f|H_0), P(F < f|H_0))$ FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio.

Quando não há intercepto no modelo ou quando há mais de um, eventualmente, o procedimento descrito anteriormente pode não ser adequado para se testar as hipóteses de interesse.

Alternativas: o procedimento pode ser adaptado ou pode-se usar outros tipo de testes como os “do tipo” $C\beta$ (o qual veremos mais adiante).

Os testes do tipo $C\beta$ são úteis (como veremos) para testar outras hipóteses de interesse como, por exemplo, descobrir quais componentes do vetor β são diferentes de 0.

Exemplo do Solvente (generalizando para k grupos)

Relembrando, o modelo é dado por

$$Y_{ij} = \mu + \alpha_i + \xi_{ij}, \quad i = 1, \dots, k \quad \text{e} \quad j = 1, \dots, n_i.$$

Resumindo:

y_{ij} = j -ésima observação do i -ésimo solvente;

μ = média geral dos dados;

α_i = efeito do solvente i ;

ξ_{ij} = componente aleatória do modelo.

Neste caso, as hipóteses de interesse são

$H_0 : \mu_1 = \mu_2 = \dots \mu_k \Leftrightarrow \alpha_2 = \alpha_3 = \dots = \alpha_k = 0$ vs H_1 : há pelo menos uma diferença.

Primeiramente, a partir dos dados utilizaremos a seguinte notação:

$y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$: soma das observações do solvente i ;

$\overline{y_{i\cdot}} = \sum_{j=1}^{n_i} y_{ij} / n_i$ média das observações do solvente i ;

$y_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}$: soma de todas as observações;

$\overline{y_{\cdot\cdot}} = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} / n$: média das observações;

onde $n = \sum_{i=1}^k n_i$ total de observações.

A soma de quadrados pode ser expressa como:

$$\begin{aligned} SQT &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \overline{y_{\cdot\cdot}})^2 = \underbrace{\sum_{i=1}^k n_i (\overline{y_{i\cdot}} - \overline{y_{\cdot\cdot}})^2}_{\text{entre os grupos}} + \underbrace{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \overline{y_{i\cdot}})^2}_{\text{dentro de cada grupo}} \\ &= SQF + SQR \end{aligned}$$

Se SQF for significativamente maior do que SQR , conclui-se que as médias populacionais são diferentes.

Somas de quadrados dos resíduos

Temos que,

■ Erro: $\xi = Y - X\beta$.

■ Valor predito: $Y = X\hat{\beta}$.

■ Resíduo (valor predito para o erro):

$$\hat{\xi} = Y - X\hat{\beta} = Y - X(X'X)^{-1}X'Y = [I - H]Y,$$

onde $H = X(X'X)^{-1}X'$ (matriz “hat” ou matriz de projeção).

Então, a $SQR = \hat{\xi}'\hat{\xi} = Y'[I - H][I - H]Y = Y'BY$, onde a matriz $B = I - H$ é simétrica, idempotente, com $posto(B) = tr(B) = n - k$.

Obs. A SQR é a parte da variabilidade não explicada pelo modelo (devida à outros fatores).

Somas de quadrados do modelo

Considere o modelo,

$$Y_{ij} = \mu + \xi_{ij},$$

ou seja,

$$\mathbf{Y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\xi} \quad (\text{modelo a ser testado}).$$

Note que $\boldsymbol{\beta}^* = \mu$ e $\mathbf{X}^* = \mathbf{1}_n = \mathbf{1}$, com $n = \sum_{i=1}^k n_i$, temos

$$SQR^* = \mathbf{Y}' \left[\mathbf{I} - \mathbf{1} (\mathbf{1}' \mathbf{1})^{-1} \mathbf{1}' \right] \mathbf{Y} = \mathbf{Y}' [\mathbf{I} - n^{-1} \mathbf{J}] \mathbf{Y},$$

$$\mathbf{J} = \mathbf{1} \mathbf{1}'.$$

Objetivo na comparação de médias: medir $SQF = SQR^* - SQR$ (soma de quadrados dos fatores ou do modelo) (exercício).

Portanto $SQF = \mathbf{Y}' [\mathbf{I} - n^{-1} \mathbf{J}] \mathbf{Y} - \mathbf{Y}' [\mathbf{I} - \mathbf{H}] \mathbf{Y} = \mathbf{Y}' [\mathbf{H} - n^{-1} \mathbf{J}] \mathbf{Y}$.

Fato: independentemente da parametrização (modelo de médias, casela de referência e desvios com restrição), o valor predito é dado por

$$\hat{\mathbf{Y}} = \mathbf{X} \hat{\boldsymbol{\beta}} = \begin{bmatrix} \bar{Y}_{1.} \\ \vdots \\ \bar{Y}_{1.} \\ \bar{Y}_{2.} \\ \vdots \\ \bar{Y}_{2.} \\ \vdots \\ \bar{Y}_{k.} \\ \vdots \\ \bar{Y}_{k.} \end{bmatrix}_{n \times 1}.$$

Ou seja,

$$H = X \left(X' X \right)^{-1} X' = \begin{bmatrix} n_1^{-1} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} & \cdots & \mathbf{0}'_{n_k} \\ \vdots & \vdots & \ddots & \vdots \\ n_1^{-1} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} & \cdots & \mathbf{0}'_{n_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{n_1} & n_2^{-1} \mathbf{1}'_{n_2} & \cdots & \mathbf{0}'_{n_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{n_1} & n_2^{-1} \mathbf{1}'_{n_2} & \cdots & \mathbf{0}'_{n_k} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \cdots & n_k^{-1} \mathbf{1}'_{n_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \cdots & n_k^{-1} \mathbf{1}'_{n_k} \end{bmatrix}$$

Seja $\mathbf{A} = [\mathbf{H} - n^{-1}\mathbf{J}]$.

Além disso, se $\mathbf{P} = n^{-1}\mathbf{J}$, temos que $\mathbf{PP} = \mathbf{P}$.

Logo $\mathbf{AA} = [\mathbf{H} - \mathbf{HP} - \mathbf{HP} + \mathbf{PP}] = [\mathbf{H} - \mathbf{P}] = \mathbf{A}$.

Assim, \mathbf{A} é uma matriz simétrica, idempotente de posto $k - 1$.

Além disso, $\mathbf{AB} = [\mathbf{H} - \mathbf{H} - \mathbf{P} + \mathbf{PH}] = \mathbf{0}_{n \times n}$.

Anova matricial: resumo

Tem-se que $SQF = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ e $SQR = \mathbf{Y}'\mathbf{B}\mathbf{Y}$, em que $\mathbf{A} = \mathbf{H} - n^{-1}\mathbf{J}$ e $\mathbf{B} = \mathbf{I} - \mathbf{H}$.

Sob H_0 , temos que

$$Y_{ij} \sim N(\mu, \sigma^2) \text{ (modelo reduzido);}$$

$$SQF/\sigma^2 \sim \chi^2_{(k-1)}, \quad SQR/\sigma^2 \sim \chi^2_{(n-k)} \text{ e } SQF \perp SQR;$$

$$\text{Então, } F = \frac{SQF/(k-1)}{SQR/(n-k)} \sim F_{(k-1, n-k)}.$$

$$\text{Sob } H_1, \quad F \sim F_{[(k-1), (k-p), \delta = \frac{1}{\sigma^2} (\beta' \mathbf{X}' \mathbf{X} \beta - n^{-1} \beta' \mathbf{X}' \mathbf{J} \mathbf{X} \beta)]}.$$

Tabela de ANOVA (matricial)

Para testar a igualdade simultânea das médias

FV	SQ	GL	QM	$\mathcal{E}(QM)$	Estat. F	p-valor
Fatores	$SQF = \mathbf{Y}'\mathbf{A}\mathbf{Y}$	k-1	$\frac{SQF}{k-1}$ (QMF)	$\sigma^2 + \frac{1}{k-1} \beta' \mathbf{X}' (\mathbf{I} - \mathbf{n}^{-1} \mathbf{J}) \mathbf{X} \beta$	$\frac{QMF}{QMR}$	p^*
Resíduo	$SQR = \mathbf{Y}'\mathbf{B}\mathbf{Y}$	n-k	$\frac{SQR}{n-k}$ (QMR)	σ^2		
Total	SQT	n-1				

$p^* = P(F > f | H_0)$ FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio.

Exemplo do Solvente

Tabela ANOVA

FV	SQ	GL	QM	Estatística F	pvalor
Solvente	0,541	4	0,135	212,81	< 0,0001
Resíduo	0,012	20	< 0,001		
Total	0.553	24			

Rejeita-se H_0 .

```
fit.model <- lm(mabsor~solvfac)
anova(fit.model)
```

Analysis of Variance Table

Response: mabsor

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
solvfac	4	0.54127	0.135317	212.81	4.378e-16 ***
Residuals	20	0.01272	0.000636		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Estimativas dos parâmetros do modelo.

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ (E50)	0,539	0,011	[0,517; 0,561]	47,826	< 0,0001
α_2 (E70)	0,069	0,0160	[0,037 ; 0,010]	4,298	0,0003
α_3 (EAW)	0,028	0,0160	[-0,004 ; 0,059]	1,726	0,0998
α_4 (M1M)	-0,343	0,0160	[-0,374; -0,311]	-21,481	< 0,0001
α_5 (MAW)	-0,090	0,0160	[-0,121 ; -0,058]	-5,624	< 0,0001

Parâmetro α_3 não significativo. Isto sugere uma possível equivalência entre os solventes E50 e EAW.

Mais testes de hipóteses em modelos normais lineares

Teste para um subconjunto dos β 's

Suponha agora que queremos testar a hipótese que um subconjunto dos X 's não é útil para prever Y .

Podemos particionar β e X da seguinte maneira $\beta = (\beta_1', \beta_2')'$ e $X = (X_1, X_2)$, o qual podemos reescrever o modelo como

$$Y = X\beta + \xi = X_1\beta_1 + X_2\beta_2 + \xi, \quad \xi \sim N(\mathbf{0}, \sigma^2 I),$$

onde β_2 contém os β 's a serem testados. O intercepto está incluso em β_1 .

Então, temos interesse em testar

$$H_0 : \beta_2 = \mathbf{0} \text{ vs } H_1 : \beta_2 \neq \mathbf{0}.$$

Sob H_0 , o modelo reduzido é dado por

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1^* + \boldsymbol{\xi}^*,$$

onde temos que a estimativa para $\widehat{\boldsymbol{\beta}}_1^* = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}$.

Pode-se mostrar que, a soma de quadrados total

$$SQT = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \mathbf{Y}'\mathbf{Y} - n\bar{Y}^2,$$

pode ser decomposta como

$$SQT = \underbrace{\mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y}}_{SQR} + \underbrace{\mathbf{Y}'(\mathbf{H} - \mathbf{H}_1)\mathbf{Y}}_{SQ(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)} + \underbrace{\mathbf{Y}'(\mathbf{H}_1 - n^{-1}\mathbf{J})\mathbf{Y}}_{SQM(reduzido)},$$

onde $\mathbf{H}_1 = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$.

SQM(reduzido): soma de quadrados do modelo reduzido; SQR: soma de quadrados dos resíduos; $SQ(\beta_2|\beta_1)$ soma de quadrados devido a β_2 depois de ajustar β_1 .

Podemos reescrever $SQ(\beta_2|\beta_1)$ como a diferença entre a soma de quadrados do modelo completo menos a soma de quadrados do modelo reduzido, ou seja,

$$\begin{aligned} SQ(\beta_2|\beta_1) &= SQM(completo) - SQM(reduzido) \\ &= \mathbf{Y}'(\mathbf{H} - n^{-1}\mathbf{J})\mathbf{Y} - \mathbf{Y}'(\mathbf{H}_1 - n^{-1}\mathbf{J})\mathbf{Y} \end{aligned}$$

Assim, se $H_0 : \beta_2 = \mathbf{0}$ é verdadeira, esperamos que a $SQ(\beta_2|\beta_1)$ seja pequena e a SQT seja composta essencialmente pela SQ do modelo reduzido e a SQ dos resíduos.

Já vimos que, $(\mathbf{I} - \mathbf{H})$ é idempotente com posto $n - p$ e que $\frac{\mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y}}{\sigma^2} \sim \chi^2_{(n-1)}$.

Podemos provar que $(\mathbf{H} - \mathbf{H}_1)$ é idempotente de posto h , onde h é o número de elementos em β_2 .

Pode-se provar também que

$$\frac{\mathbf{Y}'(\mathbf{H} - \mathbf{H}_1)\mathbf{Y}}{\sigma^2} \sim \chi^2_{(h, \delta/\sigma^2)},$$

onde $\delta = \beta_2' \mathbf{X}_2' \mathbf{X}_2 \beta_2 - \beta_2' \mathbf{X}_2' \mathbf{X}_1 (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{X}_2 \beta_2$.

Além disso, $SQR \perp SQ(\beta_2 | \beta_1)$.

Logo, sob H_0 ,

$$F = \frac{SQ(\beta_2|\beta_1)/h}{SQR/(n-p)} \sim F_{[h,(n-p)]}.$$

Ainda, sob H_1 ,

$$F \sim F_{[h,(n-p),\delta=\frac{1}{\sigma^2}(\beta_2'\mathbf{X}_2'\mathbf{X}_2\beta_2-\beta_2'\mathbf{X}_2'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\beta_2)]}.$$

Portanto, rejeita-se H_0 se $f_c > f_{critico}$ ou, analogamente, se $pvalor = P(F > f_c|H_0) < \alpha$, em que f_c é o valor calculado da estatística F e $\alpha = P(F > f_{critico}|H_0)$.

Tabela de ANOVA (matricial)

FV	SQ	GL	QM	Estatística F
$R(\beta_2 \beta_1)$	$SQ(\beta_2 \beta_1)$	h	$\frac{SQ(\beta_2 \beta_1)}{h}$	$F = \frac{SQ(\beta_2 \beta_1)/h}{SQR/(n-p)}$
$R(\beta_1)$	$SQM(reduzido)$	p-h-1	$\frac{SQM(reduzido)}{p-h-1}$	
Resíduo	SQR	n-p	$\frac{SQR}{n-p}$	
Total	SQT	n-1		

FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio.

Exemplo 1: considerando as etiologias cardíacas

O modelo considerando as etiologias cardíacas é dado por

$$Y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + \xi_{ij}, i = 1, \dots; j = 1, \dots, n_i,$$

com

- Etiologias = CH ($i = 1$), ID ($i = 2$), IS ($i = 3$), C: ($i = 4$);
- $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$;
- $(\beta_{01}, \beta_{02}, \beta_{03}, \beta_{04}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \sigma^2)$ parâmetros desconhecidos;
- x_{ij} : carga à que o paciente j que apresenta a etiologia cardíaca i foi submetido (conhecido e não aleatório);
- Parte sistemática: $\mathcal{E}(Y_{ij}) = \beta_{0i} + \beta_{1i}x_{ij}$;
- Parte aleatória: ξ_{ij} .

O modelo acima implica que $Y_{ij} \stackrel{ind.}{\sim} N(\beta_{0i} + \beta_{1i}x_{ij}, \sigma^2)$.

Ao ajustarmos o modelo anterior no R, ele fornece a seguinte “Tabela ANOVA”:

FV	GL	SQ	QM	Estatística F	p-valor
(1)	4	13749,0	3437,5	1015,726	<0,0001
(2)	4	473,3	118,3	34,964	<0,0001
Resíduos	116	392,6	3,38		

Que hipóteses estão sendo testadas em cada linha da tabela acima?

```
fit.model<-lm(vo2~-1+etiofac+carga:etiofac)
anova(fit.model)
```

Analysis of Variance Table

Response: vo2

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
etiofac	4	13750.0	3437.5	1015.726	< 2.2e-16 ***
etiofac:carga	4	473.3	118.3	34.964	< 2.2e-16 ***
Residuals	116	392.6	3.4		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Para responder à estas perguntas, precisamos saber como as somas de quadrados foram calculadas (matricialmente, de preferência) e estudar suas propriedades.

Exercício: obter as expressões matriciais das somas de quadrados em questão.

Para responder à estas perguntas, precisamos saber como as somas de quadrados foram calculadas (matricialmente, de preferência) e estudar suas propriedades.

Exercício: obter as expressões matriciais das somas de quadrados em questão.

Solução:

(1): $H_0 : \beta_{01} = \beta_{02} = \beta_{03} = \beta_{04} = 0$ vs H_1 : há pelo menos uma diferença.

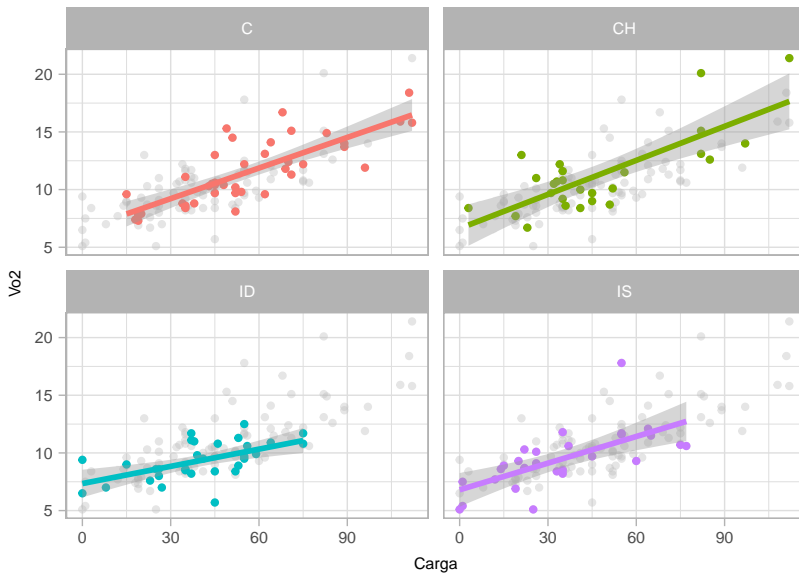
(2): $H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = 0$ vs H_1 : há pelo menos uma diferença.

Estimativas dos parâmetros do modelo.

Parâmetro	Estimativa	EP	Estat. t	IC(95%)	p-valor
$\beta_{01}(C)$	6,56	0,71	9,18	[5,16 ; 7,96]	<0,0001
$\beta_{02}(CH)$	6,63	0,75	8,88	[5,17 ; 8,10]	<0,0001
$\beta_{03}(ID)$	7,35	0,78	9,45	[5,82 ; 8,87]	<0,0001
$\beta_{04}(IS)$	6,80	0,66	10,33	[5,51 ; 8,09]	<0,0001
$\beta_{11}(C)$	0,09	0,01	7,62	[0,07 ; 0,11]	<0,0001
$\beta_{12}(CH)$	0,10	0,01	7,14	[0,07 ; 0,13]	<0,0001
$\beta_{13}(ID)$	0,05	0,02	2,82	[0,02 ; 0,08]	0,0056
$\beta_{14}(IS)$	0,08	0,02	4,78	[0,05 ; 0,11]	<0,0001

O consumo de oxigênio dos pacientes para carga 0 parecem ser semelhantes entre os grupos. O aumento no consumo parecer ser menor que os demais, para pacientes idiopáticos e igual para os outros três tipos.

Consumo de oxigenio em funcao da carga



Temos interesse em saber os consumos de oxigênio, para pacientes submetidos à uma carga nula, são os mesmos entre os grupos. Ou seja, desejamos testar se:

$$H_0 : \beta_{01} = \beta_{02} = \beta_{03} = \beta_{04} \text{ vs } H_1 : \text{há pelo menos uma diferença.} \quad (2)$$

Temos interesse em saber se os aumentos no consumo de oxigênio, são todos nulos entre os grupos. Ou seja, desejamos testar:

$$H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = 0 \text{ vs } H_1 : \text{há pelo menos uma diferença.} \quad (3)$$

Em sendo não nulos, temos interesse em saber se os aumentos no consumo de oxigênio, são os mesmos entre os grupos. Ou seja, desejamos testar:

$$H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} \text{ vs } H_1 : \text{há pelo menos uma diferença.} \quad (4)$$

Ao se detectar a existência de pelo menos uma diferença (rejeitar H_0), devemos identificar os padrões dela (comparações dois a dois, por exemplo, sempre procedendo-se com cautela).

Em geral, a grande maioria das hipóteses de interesse, podem ser descritas como

$$H_0 : C_{(q \times p)} \beta_{(p \times 1)} = \mathbf{0} \text{ vs } H_1 : C_{(q \times p)} \beta_{(p \times 1)} \neq \mathbf{0}, \quad (5)$$

em que, via de regra, $q \leq p$.

Como podemos testar as hipóteses acima?

Lembremos que no exemplo 1 considerando as etiologias cardíacas

$$\boldsymbol{\beta} = (\beta_{01}, \beta_{02}, \beta_{03}, \beta_{04}, \beta_{11}, \beta_{21}, \beta_{31}, \beta_{41})'.$$

A hipótese (nula) (2), pode ser escrita como:

$$H_0 : \begin{cases} \beta_{01} - \beta_{02} = 0 \\ \beta_{01} - \beta_{03} = 0 \\ \beta_{01} - \beta_{04} = 0 \end{cases} \Leftrightarrow H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0},$$

em que

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A hipótese (nula) (3), pode ser escrita como:

$$H_0 : \begin{cases} \beta_{11} = 0 \\ \beta_{12} = 0 \\ \beta_{13} = 0 \\ \beta_{14} = 0 \end{cases} \Leftrightarrow H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0},$$

em que

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(4 \times 4)} & \mathbf{I}_4 \end{bmatrix}.$$

A hipótese (nula) (4), pode ser escrita como:

$$H_0 : \begin{cases} \beta_{11} - \beta_{12} = 0 \\ \beta_{11} - \beta_{13} = 0 \\ \beta_{11} - \beta_{14} = 0 \end{cases} \Leftrightarrow H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0},$$

em que

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}.$$

Testes para $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$

Sabemos que

$$\hat{\boldsymbol{\theta}} = \mathbf{C}\hat{\boldsymbol{\beta}} \sim N_q(\mathbf{C}\boldsymbol{\beta}, \sigma^2 \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}').$$

Como $\hat{\boldsymbol{\beta}} \perp \hat{\sigma}^2$, então $\mathbf{C}\hat{\boldsymbol{\beta}} \perp \hat{\sigma}^2$, em que

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \frac{1}{n-p}\mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y} = \frac{SQR}{n-p} = QMR.$$

Portanto, sob $H_0(\mathbf{C}\boldsymbol{\beta} = \mathbf{0})$ e usando alguns resultados de distribuições de formas quadráticas (provar), temos que

$$Q^* = \frac{1}{\sigma^2} \left(\mathbf{C}\hat{\boldsymbol{\beta}} \right)' \left(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} \left(\mathbf{C}\hat{\boldsymbol{\beta}} \right) \sim \chi_{(q)}^2.$$

Construção da Estatística do Teste

Além disso, sabemos que $(n - p)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{(n-p)}$.

Portanto, pelos resultados acima, temos, sob H_0 , que

$$F = \frac{Q^*/q}{\hat{\sigma}^2/\sigma^2} = \frac{1}{q\hat{\sigma}^2} \left(\mathbf{C}\hat{\beta} \right)' \left(\mathbf{C} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}' \right)^{-1} \left(\mathbf{C}\hat{\beta} \right) \sim F_{(q,n-p)}$$

p -valor = $P(F > f | H_0)$, em que f é o valor calculado da estatística definida acima, e $F \sim F_{(q,n-p)}$.

Sob H_1 , $F \sim F_{\left[q, n-p, \delta = \frac{1}{\sigma^2} \left((\mathbf{C}\beta)' (\mathbf{C} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}')^{-1} (\mathbf{C}\beta) \right) \right]}$.

Voltando ao exemplo 1

Para o teste de nulidade simultânea de todos os interceptos, temos (estatística (p-valor)): 89,95 ($< 0,0001$).

Para o teste de nulidade simultânea de todos os incrementos, temos (estatística (p-valor)): 34,96 ($< 0,001$).

Para o teste de igualdade simultânea de todos os interceptos, temos (estatística (p-valor)): 0,22 (0,8842).

Para o teste de igualdade simultânea de todos os incrementos, temos (estatística (p-valor)): 1,72 (0,1666).

À rigor, após (ou mesmo antes) de ajustar o modelo, devemos verificar se as hipóteses se verificam (homocedasticidade, ausência de correlação e normalidade dos erros). Faremos isso mais adiante.

Exemplo 2: Modelo (casela de referência)

$$Y_{ij} = \mu + \alpha_i + \xi_{ij}, \quad i = 1, 2, \dots, 5 \text{ (grupos)}, \quad j = 1, \dots, 5 \text{ (u.e.)};$$

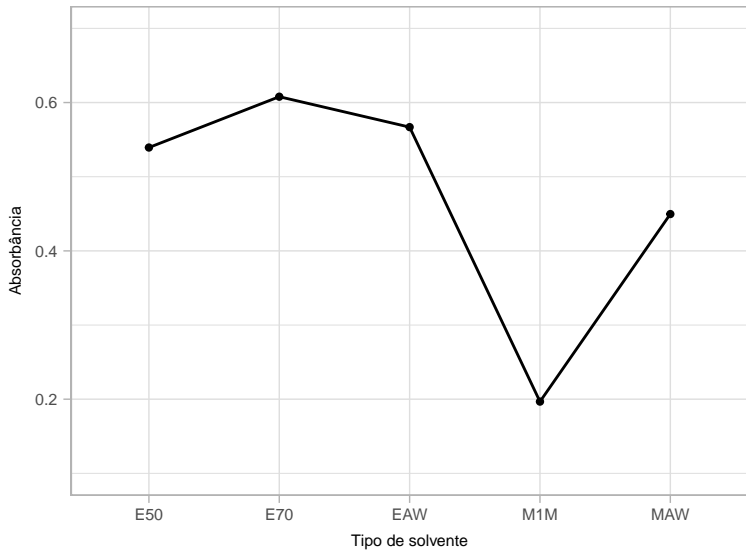
onde u.e. = unidades experimentais; com

- Erros (parte aleatória): $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$;
- μ, α_i não aleatório;
- $\mathcal{E}_{\xi_{ij}}(Y_{ij}) = \mu_i, \mathcal{V}_{\xi_{ij}}(Y_{ij}) = \sigma^2$;
- Parte sistemática: $\mu + \alpha_i$ que é a média populacional relacionada ao i -ésimo fator, $\alpha_1 = 0$;
- $Y_{ij} \stackrel{ind.}{\sim} N(\mu + \alpha_i, \sigma^2)$.

Análise descritiva

Não há sentido em construir box-plots ou histogramas (poucas observações por grupo).

Solvente	Medida descritiva					
	Média	DP	Var.	CV%	Mínimo	Máximo
E50	0,539	0,026	0,0007	4,937	0,510	0,562
E70	0,608	0,017	0,0003	2,744	0,583	0,629
EAW	0,567	0,015	0,0002	2,717	0,544	0,586
M1M	0,197	0,024	0,0006	12,107	0,165	0,225
MAW	0,450	0,037	0,0014	8,283	0,409	0,501



Teste de Bartlett para igualdade de variâncias

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ vs $H_1 : \sigma_i^2 \neq \sigma_j^2$ para pelo menos um $i \neq j$.

Estatística do teste:

$$Q_B = \frac{q}{c},$$

em que

$$q = (n - k) \ln S_p^2 - \sum_{i=1}^k (n_i - 1) \ln S_i^2, S_p^2 = QMR = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{n - k},$$

$$c = 1 + \frac{1}{3(k-1)} \left[\sum_{i=1}^k (n_i - 1)^{-1} - (n - k)^{-1} \right].$$

Sob H_0 , $Q_B \approx \chi_{(k-1)}^2$. Rejeita-se H_0 quando $P(Q_B > q_B | H_0) < \alpha$, onde q_B é o valor calculado e α é o nível de significância.

Teste de Levene para igualdade de variâncias

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ vs $H_1 : \sigma_i^2 \neq \sigma_j^2$ para pelo menos um $i \neq j$.

Estatística do teste:

$$Q_L = \frac{\sum_{i=1}^k n_i (\overline{Z_{i.}} - \overline{Z_{..}})^2 / (k - 1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - \overline{Z_{i.}})^2 / (n - k)},$$

em que

$$Z_{ij} = |Y_{ij} - \overline{Y_{i.}}|; \overline{Z_{i.}} = \frac{1}{n_i} \sum_{j=1}^{n_i} Z_{ij}; \overline{Z_{..}} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} Z_{ij}.$$

Sob H_0 , $Q_L \approx F_{(k-1, n-k)}$. Rejeita-se H_0 quando $P(Q_L > q_L | H_0) < \alpha$, onde q_L é o valor calculado e α é o nível de significância.

Testes para homocedasticidade

Teste de Bartlett : 3,772 (0,4378).

Teste de Levene : 0,696 (0,6033).

Hipótese de homocedasticidade parece razoável.

Tabela ANOVA

FV	SQ	GL	QM	Estatística F	pvalor
Solvente	0,541	4	0,135	212,81	< 0,0001
Resíduo	0,012	20	< 0,001		
Total	0.553	24			

Rejeita-se H_0 .

```
fit.model <- lm(mabsor~solvfac)
anova(fit.model)
```

Analysis of Variance Table

Response: mabsor

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
solvfac	4	0.54127	0.135317	212.81	4.378e-16 ***
Residuals	20	0.01272	0.000636		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ (E50)	0,539	0,011	[0,517; 0,561]	47,826	< 0,0001
α_2 (E70)	0,069	0,0160	[0,037 ; 0,010]	4,298	0,0003
α_3 (EAW)	0,028	0,0160	[-0,004 ; 0,059]	1,726	0,0998
α_4 (M1M)	-0,343	0,0160	[-0,374; -0,311]	-21,481	< 0,0001
α_5 (MAW)	-0,090	0,0160	[-0,121 ; -0,058]	-5,624	< 0,0001

Parâmetro α_3 não significativo. Isto sugere uma possível equivalência entre os solventes E50 e EAW.

Aplicação no exemplo 2

Lembrando os grupos : grupo 1(E50), grupo 2(E70), grupo 3(EAW), grupo 4(M1M), grupo 5(MAW).

Considere as hipóteses (H_0):

$$H_{01} : \begin{cases} \mu_1 - \mu_2 = 0, \text{ e} \\ \mu_1 - \mu_3 = 0 \end{cases}$$

$$H_{02} : \mu_1 = \mu_2.$$

$$H_{03} : \mu_1 = \mu_3.$$

$$H_{04} : \frac{\mu_1 + \mu_2 + \mu_3}{3} = \frac{\mu_4 + \mu_5}{2}.$$

$$H_{05} : \mu_3 = \mu_5.$$

Continuação: em termos das parametrização CR

Considere as hipóteses (H_0):

$$H_{01} : \begin{cases} \alpha_2 = 0, \text{ e} \\ \alpha_3 = 0 \end{cases}$$

$$H_{02} : \alpha_2 = 0.$$

$$H_{03} : \alpha_3 = 0.$$

$$H_{04} : 2\alpha_2 + 2\alpha_3 - 3\alpha_4 - 3\alpha_5 = 0.$$

$$H_{05} : \alpha_3 = \alpha_5.$$

Estatísticas (valores p)

Resultados:

$$H_{01} : 9,35 \quad (0,0014).$$

$$H_{02} : 18,47 \quad (< 0,0001).$$

$$H_{03} : 2,98 \quad (0,0998).$$

$$H_{04} : 581,90 \quad (< 0,0001).$$

$$H_{05} : 54,02 \quad (< 0,0001).$$

É importante tentar controlar o nível de significância global de todas as hipóteses testadas. Aconselha-se a utilizar $\alpha^* = \frac{\alpha}{k}$, em que α é o nível de significância adotado na tabela ANOVA (se for o caso) ou algum valor pré-fixado de interesse e k o número de hipóteses testadas.

Modelo reduzido (casela de referência)

$$Y_{ij} = \mu + \alpha_i + \xi_{ij}, \quad i = 1, 2, \dots, 5 \text{ (grupos)}, \quad j = 1, \dots, 5 \text{ (u.e.)};$$

onde u.e. = unidades experimentais; com

- Erros (parte aleatória): $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$;
- μ, α_i não aleatório;
- $\mathcal{E}_{\xi_{ij}}(Y_{ij}) = \mu_i, \mathcal{V}_{\xi_{ij}}(Y_{ij}) = \sigma^2$;
- $\mu + \alpha_i$: média populacional relacionada ao i-ésimo fator, $\alpha_1 = \alpha_3 = 0$.
- $Y_{ij} \stackrel{ind.}{\sim} N(\mu + \alpha_i, \sigma^2)$.

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ (E50/EAW)	0,553	0,008	[0,537;0,569]	66,310	< 0,0001
α_2 (E70)	0,055	0,0114	[0,026;0,083]	3,792	0,0011
α_4 (M1M)	-0,356	0,0114	[-0,385;-0,328]	-24,665	< 0,0001
α_5 (MAW)	-0,103	0,0114	[-0,132;-0,075]	-7,161	< 0,0001

Todos os incrementos α são significativos e todos parecem distintos entre si.

Estimativas finais das médias

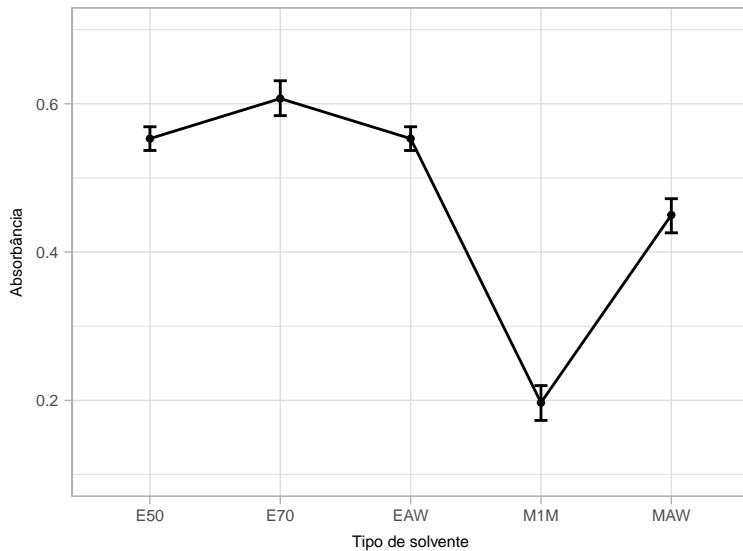
Solvente	Estimativa	EP	IC(95%)
E50/EAW	0,553	0,008	[0,537;0,569]
E70	0,607	0,012	[0,584;0,631]
M1M	0,197	0,012	[0,173;0,220]
MAW	0,450	0,012	[0,426;0,472]

Melhor solvente: E70.

Pior solvente: M1M.

Os solventes E50 e EAW são equivalentes.

Gráficos de perfis ajustados



Testes da Razão de Verossimilhança

Os testes desenvolvidos até agora foram derivados usando métodos informais baseados em características de somas de quadrados que têm distribuição de qui-quadrado e são independentes.

Esses mesmos testes podem ser obtidos de maneira mais formal, através de uma abordagem de razão de verossimilhanças.

Será apresentado os principais resultados do teste da razão de verossimilhanças no contexto simples de testar

$$H_0 : \beta = \mathbf{0} \text{ vs } H_1 : \beta \neq \mathbf{0}$$

Como vimos anteriormente, a função de verossimilhança $L(\beta, \sigma^2)$ é definida como a densidade conjunta dos Y 's. No nosso caso, supomos que $Y \sim N_n(X\beta, \sigma^2 I)$, então

$$L(\beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta) \right\}.$$

O método da razão da verossimilhança consiste em comparar o máximo valor de $L(\beta, \sigma^2)$ restrito por $H_0 : \beta = \mathbf{0}$, com o máximo valor de $L(\beta, \sigma^2)$ sob a hipótese $H_1 : \beta \neq \mathbf{0}$ que não é restrita, através da razão de verossimilhanças

$$LR = \frac{\max_{H_0} L(\beta, \sigma^2)}{\max_{H_1} L(\beta, \sigma^2)} = \frac{\max L(0, \sigma^2)}{\max L(\beta, \sigma^2)},$$

onde $0 \leq LR \leq 1$, uma vez que o máximo de L restrito a $\beta = \mathbf{0}$ não pode exceder o máximo irrestrito. Pequenos valores de LR favorecem H_1 , e grandes valores de LR favorecem H_0 .

Assim, nós rejeitamos H_0 se $LR \leq c$, onde c é escolhido de tal forma que $P(LR \leq c) = \alpha$ se H_0 é verdadeira.

Wald (1943) mostrou que, sob H_0 ,

$$-2 \ln(LR) \sim \chi^2(\nu),$$

para n grande, onde ν é o número de parâmetros estimados sob H_1 menos o número de parâmetros estimados sob H_0 .

No caso de $H_0 : \beta = \mathbf{0}$ vs $H_1 : \beta \neq \mathbf{0}$, nós temos $v = k + 2 - 1 = k + 1$, porque β e σ^2 são estimados sob H_1 , enquanto somente a variância σ^2 é estimada sob H_0 .

Teorema: Se $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, o teste da razão de verossimilhança para $H_0 : \boldsymbol{\beta} = \mathbf{0}$ vs $H_1 : \boldsymbol{\beta} \neq \mathbf{0}$ pode ser baseado em

$$F = \frac{\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{Y} / (k + 1)}{(\mathbf{Y}' - \mathbf{Y} \hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{Y}) / (n - k - 1)}.$$

E rejeitamos H_0 se $F > F_{\alpha, k+1, n-k-1}$.

- Notas de aula do Prof. Caio Azevedo.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley series in probability and statistics.