Teste de hipótese em modelos normais lineares: ANOVA

Prof. Caio Azevedo

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:

$$H_0 : \beta_1 = 0 \; vs \; H_1 : \beta_1 \neq 0, i = 1, ..., p$$

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

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

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:

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

- 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:

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

- Logo $Y_i \overset{ind.}{\sim} N(\beta_0 + \sum_{j=1}^{p-1} \beta_j x_{ji}, \sigma^2)$.
- O modelo acima defina uma média (condicional aos valores de $x_{ji}, j = 1, ..., p-1; i = 1, ..., n$) para cada observação Y_i.
- Defina $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{1i} + \widehat{\beta}_2 x_{2i} + ... + \widehat{\beta}_{p-1} x_{(p-1)i}$ (valor predito pelo modelo).
- O resíduo é definido por $R_i = \widehat{\xi}_i = Y_i - \widehat{Y}_i$.

- Nosso objetivo é considerar um modelo que explique adequadamente a variabilidade dos dados, ou seja, um modelo para o qual os resíduos sejam "pequenos".
- Pode-se provar que, a soma de quadrados total $SQT = \sum_{i=1}^{n}(Y_i - \overline{Y})^2$, pode ser decomposta como:

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

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

- 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 distrisbuição 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: $\mathbf{Y} = \mathbf{X}\beta + \xi$, pode-se demonstrar que:
    - SQT $= \mathbf{Y}'\left(\mathbf{I} - n^{-1}\mathbf{J}\right)\mathbf{Y}$, em que $\mathbf{J} = \mathbf{1}\mathbf{1}'$.
    - SQM $= \mathbf{Y}'\left(\mathbf{H} - n^{-1}\mathbf{J}\right)\mathbf{Y}$, em que $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.
    - SQR $= \mathbf{Y}'\left(\mathbf{I} - \mathbf{H}\right)\mathbf{Y}$.
- 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 $\widehat{\mathbf{Y}}$ 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 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:

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

- Vamos demonstrar a validade do resultado dado pela equação (1).
- Por simplicidade, admita que $\boldsymbol{\mu} = \mathbf{0}$.

- Assim, temos que :

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

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

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

- A qual corresponde à fgm de uma distribuição χ^2_p.

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 distrbuição 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 qui-quadrado não central

- Sejam $X_1, X_2, ..., X_n \overset{ind}{\sim} N(\mu_i, \sigma_i^2)$.
- 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 $Y \sim \chi^2_{(n,\delta)}$, 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 $W_{n+2i} \sim \chi^2_{(n+2i)}$

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

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 $F \sim \chi^2_{(n,m,\delta)}$, cuja fdp é dada por

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

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

Principais teoremas de formas quadráticas normais

- 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.

- 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.

Cont.

- 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.

- 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.

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

Cont.

- Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathbf{A}$ 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 $\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{0}$.
- Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathbf{A}$ 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.

Voltando ao problema

- Lembremos que:
  - SQM $= \mathbf{Y}'\left(\mathbf{H} - n^{-1}\mathbf{J}\right)\mathbf{Y}$, em que $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.
  - SQR $= \mathbf{Y}'\left(\mathbf{I} - \mathbf{H}\right)\mathbf{Y}$.
  - $\mathbf{A} = \mathbf{H} - n^{-1}\mathbf{J}$ e $\mathbf{B} = \mathbf{I} - \mathbf{H}$
  - $\mathbf{Y} \sim N_n(\mathbf{X}\beta, \sigma^2\mathbf{I})$.
- Defina $W = SQM/\sigma^2 = \mathbf{Y}'\left(\dfrac{\mathbf{H} - n^{-1}\mathbf{J}}{\sigma^2}\right)\mathbf{Y}$ e $V = SQR/\sigma^2 = \mathbf{Y}'\left(\dfrac{\mathbf{I} - \mathbf{H}}{\sigma^2}\right)\mathbf{Y}$.

Voltando ao problema

- 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 $W \perp V$.
- Pode-se provar que $\frac{\mathbf{A}}{\sigma^2}(\sigma^2\mathbf{I}) = \mathbf{A}$ e $\frac{\mathbf{B}}{\sigma^2}(\sigma^2\mathbf{I}) = \mathbf{B}$. Como $\mathbf{A}$ e $\mathbf{B}$ são matrizes idempotentes, logo $\mathbf{AI}$ e $\mathbf{BI}$, também o são.
- Pode-se provar ainda 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 $H_1, F \sim F_{\left[(p-1),(n-p),\delta=\frac{1}{\sigma^2}(\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}-n^{-1}\boldsymbol{\beta}'\mathbf{X}'\mathbf{J}\mathbf{X}\boldsymbol{\beta})\right]}$.
- 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 (com 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 $H_0 : \beta_1 = \beta_2 = ... = \beta_{(p-1)} = 0$ vs H_1 : Há pelo menos uma diferença.

FV	SQ	GL	QM	Estatística F	pvalor
Modelo	$SQM = \mathbf{Y}'\mathbf{A}\mathbf{Y}$	p-1	$QMM = \frac{SQM}{p-1}$	$F = \frac{QMM}{QMR}$	$\min(P(F > f\|H_0), P(F < f\|H_0)$
Resíduo	$SQR = \mathbf{Y}'\mathbf{B}\mathbf{Y}$	n-p	$QMR = \frac{SQR}{n-p}$		
Total	SQT	n-1			

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

Comentários

- 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" $\mathbf{C}\boldsymbol{\beta}$ (o qual veremos mais adiante).
- Os testes do tipo $\mathbf{C}\boldsymbol{\beta}$ são úteis (como veremos) para testar outras hipóteses de interesse como, por exemplo, descobrir quais componentes do vetor $\boldsymbol{\beta}$ são diferentes de 0.

Exemplo do Solvente generalizando para k grupos

- Neste caso, as hipóteses de interesse são $H_0 : \mu_1 = \mu_2 = ...\mu_k \Leftrightarrow \alpha_2 = \alpha_3 = ... = \alpha_k = 0$ vs H_1 : há pelo menos uma diferença.
- Soma de quadrados total $\overline{y} = \frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n_i} y_{ij}, n = n_1 + n_2$ (provar a expressão abaixo)

  $SQT = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \overline{y})^2 =$

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

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

Somas de quadrados do resíduo

- Erro $\boldsymbol{\xi} = \mathbf{Y} - \mathbf{X}\beta$. Valor predito $\mathbf{Y} = \mathbf{X}\widehat{\beta}$.
- Resíduo (valor predito para o erro)
  $\widehat{\boldsymbol{\xi}} = \mathbf{Y} - \mathbf{X}\widehat{\beta} = \mathbf{Y} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}'\right)^{-1}\mathbf{X}\mathbf{Y} = \left[\mathbf{I} - \mathbf{H}'\right]\mathbf{Y}$.
  $\mathbf{H} = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}$ (matriz "hat" ou matriz de projeção).
- SQR $= \widehat{\boldsymbol{\xi}}'\widehat{\boldsymbol{\xi}} = \mathbf{Y}'\left[\mathbf{I} - \mathbf{H}\right]\left[\mathbf{I} - \mathbf{H}\right]\mathbf{Y} = \mathbf{Y}'\mathbf{B}\mathbf{Y}$.
- A matriz $\mathbf{B} = \mathbf{I} - \mathbf{H}$ é simétrica, idempotente de posto=traço = n-k (p=k).
- 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 $\mathbf{Y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\xi}$ (modelo a ser testado).
- Note que $\boldsymbol{\beta}^* = \mu$ e $\mathbf{X}^* = \mathbf{1}_n = \mathbf{1}$, $n = \sum_{i=1}^{k} n_i$.
- $SQR^* = \mathbf{Y}' \left[\mathbf{I} - \mathbf{1}\left(\mathbf{1}'\mathbf{1}\right)^{-1}\mathbf{1}'\right]\mathbf{Y} = \mathbf{Y}'\left[\mathbf{I} - n^{-1}\mathbf{J}\right]\mathbf{Y}$, $\mathbf{J} = \mathbf{1}\mathbf{1}'$.
- Objetivo na comparação de médias: medir $SQF = SQR^* - SQR$ (soma de quadrados dos fatores ou do modelo) (exercício).

Cont.

- Portanto $SQF = \mathbf{Y}'\left[\mathbf{I} - n^{-1}\mathbf{J}\right]\mathbf{Y} - \mathbf{Y}'[\mathbf{I} - \mathbf{H}]\mathbf{Y} = \mathbf{Y}'\left[\mathbf{H} - n^{-1}\mathbf{J}\right]\mathbf{Y}$.
- Fato: independentemente da parametrização (modelo de médias, casela de referência e desvios com restrição), o valor predito é dado por

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

Cont.

- Ou seja:

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

Cont.

- Portanto: $n^{-1}\mathbf{HJ} = n^{-1}\mathbf{J}$ (exercício). Seja $\mathbf{A} = \left[\mathbf{H} - n^{-1}\mathbf{J}\right]$
- 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, $Y_{ij} \sim N(\mu, \sigma^2)$ (modelo reduzido).
- Assim, sob H_0, $SQF/\sigma^2 \sim \chi^2_{(k-1)}$, $SQR/\sigma^2 \sim \chi^2_{(n-k)}$ e $SQF \perp SQR$.
- Logo, sob H_0, $F = \frac{SQF/(k-1)}{SQR/(n-k)} \sim F_{(k-1,n-k)}$.
- Sob H_1, tem-se resultado semelhante ao caso anterior (modelos de regressão).

Tabela de ANOVA (matricial)

- Para testar a igualdade simultânea das médias

FV	SQ	GL	QM	Estatística F	pvalor
Fatores	$SQF = \mathbf{Y}'\mathbf{A}\mathbf{Y}$	k-1	$QMF = \frac{SQF}{k-1}$	$F = \frac{QMF}{QMR}$	$P(F > f \mid H_0)$
Resíduo	$SQR = \mathbf{Y}'\mathbf{B}\mathbf{Y}$	n-k	$QMR = \frac{SQR}{n-k}$		
Total	SQT	n-1			

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