Distribuição de Formas Quadráticas

Gilberto A. Paula

Departamento de Estatística
IME-USP, Brasil
giapaula@ime.usp.br

1º Semestre 2023

Sumário

Objetivos

Neste material serão apresentados alguns resultados relacionados com formas quadráticas e aplicações nas somas de quadrados SQReg e SQRes:

- Resultados de Formas Quadráticas
- Distribuição Somas de Quadrados
- Estatística F
- Acréscimo na SQRes
- Salário de Executivos
- Referências

Sumário

Definição

Seja $\mathbf{Y} \sim \mathrm{N}_n(\boldsymbol{\mu}, \mathbf{V})$ e considere a forma quadrática

$$U = \mathbf{Y}^\top \mathbf{A} \mathbf{Y},$$

em que $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$, $\mathbf{V}$ é uma matriz $n \times n$ positiva definida e $\mathbf{A}$ é uma matriz $n \times n$ de constantes. Tipicamente em regressão linear múltipla tem-se que $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ e $\mathbf{V} = \sigma^2 \mathbf{I}_n$.

Resultado 1

Se **AV** ou **VA** é idempotente de posto r, então

$$U \sim \chi^2_r(\lambda),$$

em que $\lambda = \boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu}$. Em particular, se $\mathbf{V} = \sigma^2 \mathbf{I}_n$ e **A** é idempotente de posto r então

$$\frac{U}{\sigma^2} \sim \chi^2_r(\lambda),$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu}}{\sigma^2}$.

Resultado 2

Seja $\mathbf{B}$ outra matriz $n \times n$ de constantes e considere a forma quadrática

$$W = \mathbf{Y}^\top \mathbf{B} \mathbf{Y},$$

em que $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$. Se $\mathbf{BVA} = \mathbf{0}$, então as duas formas quadráticas U e W são independentes. Em particular, se $\mathbf{V} = \sigma^2 \mathbf{I}_n$ então as formas quadráticas U e W são independentes se $\mathbf{BA} = \mathbf{0}$.

Propriedades

Supor $\mathbf{Y} \sim \mathrm{N}_n(\boldsymbol{\mu}, \mathbf{V})$, $\mathbf{A}$ uma matriz $n \times n$ de constantes e $\mathbf{a}$ um vetor $n \times 1$ de constantes. Seguem algumas propriedades:

- $\mathrm{E}(\mathbf{a}^\top \mathbf{Y}) = \mathbf{a}^\top \boldsymbol{\mu}$.
- $\mathrm{E}(\mathbf{AY}) = \mathbf{A}\boldsymbol{\mu}$.
- $\mathrm{Var}(\mathbf{a}^\top \mathbf{Y}) = \mathbf{a}^\top \mathbf{Va}$.
- $\mathrm{Var}(\mathbf{AY}) = \mathbf{AVA}^\top$.
- $\mathrm{E}(\mathbf{Y}^\top \mathbf{AY}) = \text{traço}(\mathbf{AV}) + \boldsymbol{\mu}^\top \mathbf{A}\boldsymbol{\mu}$.

Em particular, se $\mathbf{V} = \sigma^2 \mathbf{I}_n$ então tem-se que $\mathrm{Var}(\mathbf{AY}) = \sigma^2 \mathbf{AA}^\top$ e $\mathrm{E}(\mathbf{Y}^\top \mathbf{AY}) = \sigma^2 \text{traço}(\mathbf{A}) + \boldsymbol{\mu}^\top \mathbf{A}\boldsymbol{\mu}$.

Sumário

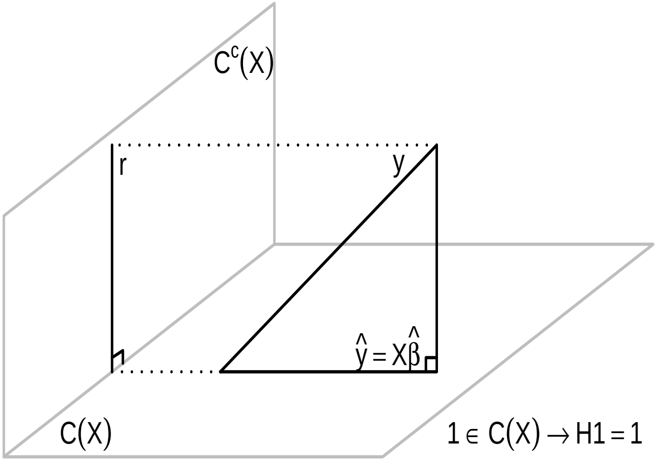
Distribuição SQReg

Tem-se que SQReg pode ser expressa nas formas

$$
\begin{aligned}
\text{SQReg} &= \sum_{i=1}^{n} (\widehat{Y}_i - \bar{Y})^2 \\
&= \{\widehat{\mathbf{Y}} - \mathbf{1}\bar{Y}\}^\top \{\widehat{\mathbf{Y}} - \mathbf{1}\bar{Y}\} \\
&= \{\mathbf{HY} - \mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\mathbf{Y}\}^\top \{\mathbf{HY} - \mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\mathbf{Y}\} \\
&= \mathbf{Y}^\top \{\mathbf{H} - \mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\}\mathbf{Y},
\end{aligned}
$$

usando o resultado $\mathbf{H1} = \mathbf{1}$, em que $\mathbf{1}$ denota vetor $n \times 1$ de 1's.

Ilustração Solução de Mínimos Quadrados



Distribuição SQReg

Portanto, segue que

$$\mathrm{SQReg} = \mathbf{Y}^\top \mathbf{A}\mathbf{Y},$$

em que $\mathbf{A} = \{\mathbf{H} - \mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\}$. É possível mostrar que $\mathbf{AA} = \mathbf{A}$ (idempotente) e que traço($\mathbf{A}$) $=$ traço($\mathbf{H}$) $-$ traço$\{\mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\}$ = $(p-1)$[a]. Portanto, usando Resultado 1 segue que

$$\frac{\mathrm{SQReg}}{\sigma^2} \sim \chi^2_{(p-1)}(\lambda),$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top\mathbf{A}\boldsymbol{\mu}}{\sigma^2}$. Ainda obtém-se $\mathrm{E}\left(\frac{\mathrm{SQReg}}{\sigma^2}\right) = (p-1) + \frac{\boldsymbol{\mu}^\top\mathbf{A}\boldsymbol{\mu}}{\sigma^2}$.

[a] Se $\mathbf{A}$ é idempotente então traço($\mathbf{A}$) = posto($\mathbf{A}$)

Distribuição SQRes

Tem-se que SQRes pode ser expressa nas formas

$$
\begin{aligned}
\text{SQRes} &= \sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 \\
&= \{\mathbf{Y} - \widehat{\mathbf{Y}}\}^\top \{\mathbf{Y} - \widehat{\mathbf{Y}}\} \\
&= \{\mathbf{Y} - \mathbf{HY}\}^\top \{\mathbf{Y} - \mathbf{HY}\} \\
&= \mathbf{Y}^\top \{\mathbf{I}_n - \mathbf{H}\}^\top \{\mathbf{I}_n - \mathbf{H}\}\mathbf{Y} \\
&= \mathbf{Y}^\top \{\mathbf{I}_n - \mathbf{H}\}\{\mathbf{I}_n - \mathbf{H}\}\mathbf{Y} \\
&= \mathbf{Y}^\top \{\mathbf{I}_n - \mathbf{H}\}\mathbf{Y}.
\end{aligned}
$$

Distribuição SQRes

Portanto, segue que

$$\text{SQRes} = \mathbf{Y}^\top \mathbf{B} \mathbf{Y},$$

em que $\mathbf{B} = (\mathbf{I}_n - \mathbf{H})$. É possível mostrar que $\mathbf{BB} = \mathbf{B}$ (idempotente) e que traço($\mathbf{B}$) $=$ traço($\mathbf{I}_n$) $-$ traço($\mathbf{H}$) = $(n-p)$. Portanto, usando Resultado 1 segue que

$$\frac{\text{SQRes}}{\sigma^2} \sim \chi^2_{(n-p)}(\lambda),$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{B} \boldsymbol{\mu}}{\sigma^2} = \frac{\boldsymbol{\mu}^\top (\mathbf{I}_n - \mathbf{H}) \boldsymbol{\mu}}{\sigma^2} = 0$. Logo, $\mathrm{E}\left(\frac{\text{SQRes}}{\sigma^2}\right) = (n-p)$.

Sumário

Definição

Supor que $U \sim \chi_r^2(\lambda)$ e $W \sim \chi_s^2$ e que U e W são independentes. Então segue que

$$F = \frac{U/r}{W/s} \sim F_{r,s}(\lambda),$$

em que $F_{r,s}(\lambda)$ denota distribuição F com r e s graus de liberdade e parâmetro de não centralidade λ.

Aplicação

Denote $U = \text{SQReg} = \mathbf{Y}^\top \mathbf{AY}$ e $W = \text{SQRes} = \mathbf{Y}^\top \mathbf{BY}$. Então

$$\frac{U}{\sigma^2} \sim \chi^2_{(p-1)}(\lambda) \quad \text{e} \quad \frac{W}{\sigma^2} \sim \chi^2_{(n-p)},$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu}}{\sigma^2}$, $\mathbf{A} = \{\mathbf{H} - \mathbf{1}(\mathbf{1}^\top \mathbf{1})^{-1} \mathbf{1}^\top\}$ e $\mathbf{B} = (\mathbf{I}_n - \mathbf{H})$. Usando o resultado $\mathbf{H1} = \mathbf{1}$ segue que $\mathbf{AB} = \mathbf{0}$.

Aplicação

Desde que $\mathbf{AB} = \mathbf{0}$, então segue pelo Resultado 2 que U e W são independentes. Portanto

$$F = \frac{U/(p-1)}{W/(n-p)} \sim F_{(p-1),(n-p)}(\lambda),$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu}}{\sigma^2}$.

Parâmetro de Não Centralidade

Denote $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, em que $\mathbf{X} = [\mathbf{1}, \mathbf{X}_R]$ e $\boldsymbol{\beta} = [\beta_1, \boldsymbol{\beta}_R^\top]^\top$. Logo, o parâmetro de nâo centralidade λ fica dado por

$$\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu}}{\sigma^2} = \frac{1}{\sigma^2}[\beta_1, \boldsymbol{\beta}_R^\top]^\top \begin{bmatrix} \mathbf{1}^\top \\ \mathbf{X}_R^\top \end{bmatrix} \mathbf{A}[\mathbf{1}, \mathbf{X}_R] \begin{bmatrix} \beta_1 \\ \boldsymbol{\beta}_R \end{bmatrix},$$

em que $\mathbf{A} = \{\mathbf{H} - \mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\}$. Note que $\mathbf{X}_R$ é uma matriz $n \times (p-1)$ que corresponde à matriz modelo dos coeficientes da regressão $\boldsymbol{\beta}_R$.

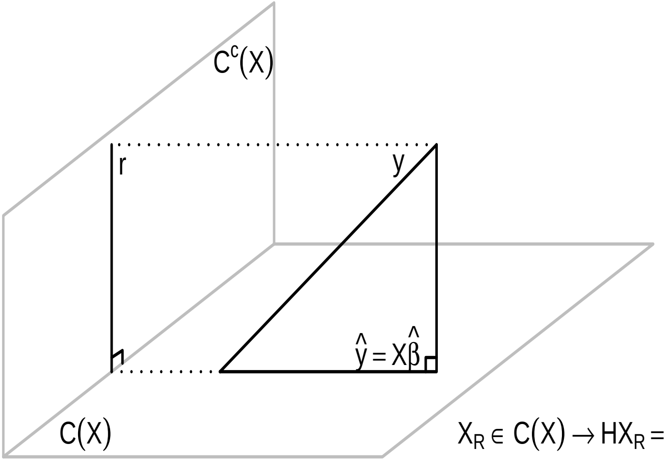
Parâmetro de Não Centralidade

Usando os resultados $\mathbf{H1} = \mathbf{1}$ e $\mathbf{HX}_R = \mathbf{X}_R$, pode-se mostrar após algumas manipulações algébricas que

$$\lambda = \frac{1}{\sigma^2}\boldsymbol{\beta}_R^\top[\mathbf{X}_R^\top\mathbf{X}_R - \mathbf{X}_R^\top\mathbf{1}(\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top\mathbf{X}_R]\boldsymbol{\beta}_R.$$

Portanto, se $\boldsymbol{\beta}_R = \mathbf{0}$ então $\lambda = 0$.

Ilustração Solução de Mínimos Quadrados



$X_R \in C(X) \rightarrow HX_R = X_R$

Aplicação

Supor que o interesse seja testar $H_0 : \boldsymbol{\beta}_R = \mathbf{0}$ contra $H_1 : \boldsymbol{\beta}_R \neq \mathbf{0}$. Portanto, segue que

$$F = \frac{\text{SQReg}/(p-1)}{\text{SQRes}/(n-p)} \overset{H_0}{\sim} F_{(p-1),(n-p)}.$$

Rejeita-se H_0 se $F > F_{(1-\alpha),(p-1),(n-p)}$, em que $F_{(1-\alpha),(p-1),(n-p)}$ denota o quantil $(1-\alpha)$ da distribuição F com $(p-1)$ e $(n-p)$ graus de liberdade.

Sumário

Modelo Particionado

Considere o modelo linear múltiplo particionado

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon},$$

em que $\mathbf{X}_1$ e $\mathbf{X}_2$ são matrizes $n \times p_1$ e $n \times p_2$, respectivamente, enquanto $\boldsymbol{\beta}_1$ e $\boldsymbol{\beta}_2$ são vetores de dimensão $p_1 \times 1$ e $p_2 \times 1$, respectivamente.

Modelo Particionado

Sob o modelo $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \epsilon$ a soma de quadrados de resíduos fica dada por

$$\mathbf{Y}^\top(\mathbf{I}_n - \mathbf{H})\mathbf{Y},$$

em que $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$ com $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$.

Similarmente, sob o modelo $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \epsilon$ a soma de quadrados de resíduos fica dada por

$$\mathbf{Y}^\top(\mathbf{I}_n - \mathbf{H}_1)\mathbf{Y},$$

em que $\mathbf{H}_1 = \mathbf{X}_1(\mathbf{X}_1^\top\mathbf{X}_1)^{-1}\mathbf{X}_1^\top$.

Modelo Particionado

Portanto, o acréscimo na soma de quadrados de resíduos devido à restrição $\boldsymbol{\beta}_2 = \mathbf{0}$ fica dado por

$$\mathrm{ASQ}(\boldsymbol{\beta}_2 = \mathbf{0}) = \mathbf{Y}^\top(\mathbf{I}_n - \mathbf{H}_1)\mathbf{Y} - \mathbf{Y}^\top(\mathbf{I}_n - \mathbf{H})\mathbf{Y} = \mathbf{Y}^\top\mathbf{C}\mathbf{Y},$$

em que $\mathbf{C} = (\mathbf{H} - \mathbf{H}_1)$. É possível mostrar que $\mathbf{CC} = \mathbf{C}$ (idempotente) e que traço($\mathbf{C}$) $=$ traço($\mathbf{H}$) $-$ traço($\mathbf{H}_1$) = $(p_1 + p_2 - p_1) = p_2$. Portanto, usando Resultado 1 segue que

$$\frac{\mathrm{ASQ}(\boldsymbol{\beta}_2 = \mathbf{0})}{\sigma^2} \sim \chi^2_{p_2}(\lambda),$$

em que $\lambda = \frac{\boldsymbol{\mu}^\top\mathbf{C}\boldsymbol{\mu}}{\sigma^2}$.

Modelo Particionado

O acréscimo na soma de quadrados de resíduos

$$\mathrm{ASQ}(\beta_2 = \mathbf{0}) = \mathbf{Y}^\top \mathbf{C}\mathbf{Y}$$

pode ser interpretado como o acréscimo na SQReg devido à inclusão do parâmetro β_2. É possível mostrar que $\mathbf{CB} = \mathbf{0}$. Logo, pelo Resultado 2 tem-se a independência entre $U = \mathbf{Y}^\top \mathbf{C}\mathbf{Y}$ e $W = \mathbf{Y}^\top \mathbf{B}\mathbf{Y}$.

Parâmetro de Não Centralidade

Pode-se mostrar após algumas manipulações algébricas que

$$\lambda = \frac{\boldsymbol{\mu}^\top \mathbf{C} \boldsymbol{\mu}}{\sigma^2} = \frac{1}{\sigma^2} \boldsymbol{\beta}_2^\top [\mathbf{X}_2^\top \mathbf{X}_2 - \mathbf{X}_2^\top \mathbf{H}_1 \mathbf{X}_2] \boldsymbol{\beta}_2.$$

Portanto, se $\boldsymbol{\beta}_2 = \mathbf{0}$ então $\lambda = 0$.

Estatística F

Supor que o interesse seja testar $\mathrm{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ contra $\mathrm{H}_1 : \boldsymbol{\beta}_2 \neq \mathbf{0}$. Portanto, segue que

$$F = \frac{\mathrm{ASQ}(\boldsymbol{\beta}_2 = \mathbf{0})/p_2}{\mathrm{SQRes}/(n-p)} \overset{\mathrm{H}_0}{\sim} F_{p_2,(n-p)}.$$

Rejeita-se H_0 se $\mathrm{F} > \mathrm{F}_{(1-\alpha),p_2,(n-p)}$, em que $\mathrm{F}_{(1-\alpha),p_2,(n-p)}$ denota o quantil $(1-\alpha)$ da distribuição F com p_2 e $(n-p)$ graus de liberdade.

Relação Estatísticas t e F

Supor que o interesse seja testar $\text{H}_0 : \beta_j = 0$ contra $\text{H}_1 : \beta_j \neq 0$ para algum $j = 2, \ldots, p$. A estatísticas t e F ficam dadas por

$$\text{t} = \frac{\widehat{\beta}_j}{\widehat{\text{EP}}(\widehat{\beta}_j)} \overset{\text{H}_0}{\sim} \text{t}_{(n-p)} \quad \text{e} \quad F = \frac{\text{ASQ}(\beta_j = 0)}{\text{SQRes}/(n-p)} \overset{\text{H}_0}{\sim} F_{1,(n-p)}.$$

É possível mostrar que $F = t^2$.

Sumário

Acréscimo na SQRes

Considere novamente o modelo linear múltiplo

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

em que $\mathbf{X}$ denota a matriz modelo de dimensão $n \times p$, $\boldsymbol{\beta}$ é um vetor de dimensão $p \times 1$ de coeficientes da regressão e $\boldsymbol{\epsilon} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$. Supor as restrições lineares $\mathbf{R}\boldsymbol{\beta} = \mathbf{0}$, em que $\mathbf{R}$ é uma matriz de dimensão $r \times p$ e posto linha completo $r \leq p$.

Acréscimo na SQRes

É possível mostrar que o acréscimo na soma de quadrados de resíduos no modelo linear múltiplo, devido às restrições lineares $\mathbf{R}\beta = \mathbf{0}$, pode ser expresso na forma

$$\text{ASQ}(\mathbf{R}\beta = \mathbf{0}) = \widehat{\beta}^\top \mathbf{R}^\top \{\mathbf{R}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{R}^\top\}^{-1} \mathbf{R}\widehat{\beta},$$

em que $\widehat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$.

Estatística F

Supor que o interesse agora seja testar $\mathrm{H}_0 : \mathbf{R}\beta = \mathbf{0}$ contra $\mathrm{H}_1 : \mathbf{R}\beta \neq \mathbf{0}$. A estatística F fica dada por

$$F = \frac{\mathrm{ASQ}(\mathbf{R}\beta = \mathbf{0})/r}{\mathrm{SQRes}/(n-p)} \overset{\mathrm{H}_0}{\sim} F_{r,(n-p)}.$$

Rejeita-se H_0 se $\mathrm{F} > \mathrm{F}_{(1-\alpha),r,(n-p)}$, em que $\mathrm{F}_{(1-\alpha),r,(n-p)}$ denota o quantil $(1-\alpha)$ da distribuição F com r e $(n-p)$ graus de liberdade.

Exemplos

Considere o modelo de regressão linear

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i,$$

em que $\epsilon_i \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma^2)$, para $i = 1, \ldots, n$. Para testar a hipótese linear $\mathrm{H}_0 : \beta_2 = \beta_4$ a matriz $\mathbf{R}$ fica dada por $\mathbf{R} = [0\ 1\ 0 - 1]$.
Se o interesse for testar a hipótese conjunta $\mathrm{H}_0 : \beta_2 = 0$ e $\beta_3 = \beta_4$, então tem-se que

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

Sumário

Modelo

No exemplo salário de executivos foi sugerido inicialmente o seguinte modelo linear:

$$y_i = \alpha + \beta_1 \times \text{gênero}_i + \beta_2 \times \text{experiência}_i + \beta_3 \times \text{posição}_i + \epsilon_i,$$

para $i = 1, \ldots, 220$, em que y_i denota o salário do i-ésimo executivo da amostra com $\epsilon_i \stackrel{\text{iid}}{\sim} \text{N}(0, \sigma^2)$.

Estimativas

Na tabela abaixo são apresentados os valores das Estatísticas t e F para cada efeito.

Efeito	Estimativa	valor-t	valor-F	valor-P
Constante	115,262	82,25	-	0,00
Posição	6,710	21,46	460,72	0,00
Experiência	-0,472	-4,17	17,38	0,00
GêneroM	-2,201	-2,04	4,15	0,04

Posição e experiência são altamente significativos, enquanto gênero é significativo ao nível de 5%.

Tabela ANOVA

A Tabela ANOVA faz a decomposição da SQReg segundo a ordem em que os efeitos são colocados. Abaixo segue uma das saídas possíveis.

Efeito	S.Quadrados	G.L.	Q. Médio	valor-F	valor-P
Posição	23029,4	1	23029,4	502,81	0,00
Experiência	1218,6	1	1218,6	26,61	0,00
GêneroM	190,1	1	190,1	4,15	0.04
Resíduos	9893,0	216	45,8		
Total	34331,1	219			

Nessa ordem de entrada dos efeitos tem-se resultados similares aos da tabela anterior.

Inclusão de Interação

Na tabela abaixo são apresentados os valores da Estatística F para cada interação dados os efeitos principais.

Interação	valor-F	valor-P
gênero*experiência	1,615	0,20
gênero*posicão	0,001	0,97
experiência*posição	7,594	0,00

Portanto, foi incluída no modelo apenas a interação experiência*posição.

Sumário

Referências

- Foster, D. P., Stine, R. A. e Waterman, R. P. (1998). *Business Analysis using Regression*. New York: Springer.
- Montgomery, D. C.; Peck, E. A. e Vining, G. G. (2021, Apêndice C). *Introduction to Linear Regression Analysis, 6th Edition*. Hoboken: Wiley.