

## **Algumas considerações práticas sobre a precisão dos estimadores de mínimos quadrados e máxima verossimilhança para os parâmetros de regressão linear normal**

*Some practical considerations about precision of estimators of minimum squares and likelihood for the parameters of linear regression normal*

Elisângela Aparecida da Silva Lizzi<sup>1</sup>, Angela Achcar<sup>2</sup>, Edson Zangiacomi Martinez<sup>3</sup> e Jorge Achcar<sup>4</sup>

<sup>1,2</sup> Doutoranda Departamento de Medicina Social-FMRP, Universidade de São Paulo  
elisangela.ass@hotmail.com; angelachcar@usp.br;

<sup>3,4</sup> Doutor, Prof. Departamento de Medicina Social-FMRP Universidade de São Paulo  
edson@fmrp.usp.br; achcar@fmrp.usp.br

### **Resumo**

*Neste artigo discutimos alguns aspectos práticos sobre o uso de diferentes métodos de estimação dos parâmetros do modelo de regressão linear e seus impactos nas inferências obtidas. Na prática, é comum usarmos softwares estatísticos para obter os estimadores de mínimos quadrados (EMQ) para os parâmetros de regressão e considerar um estimador não-viciado para a variância do erro baseado na soma de quadrados dos resíduos. Como alternativa, discutimos o uso do estimador de máxima verossimilhança (EMV) para a variância do erro (um estimador viciado, mas com erro quadrático menor) que pode levar a diferentes conclusões em termos de inferências. Dois exemplos numéricos são apresentados para ilustrar o estudo comparativo.*

**Palavras-chave:** Estimador de mínimos quadrados. Estimador de máxima verossimilhança. Modelo de regressão linear. Precisão.

### **Abstract**

*In this paper, we discuss some practical considerations on the use of different methods for the estimation of linear regression parameters and their impact on the obtained inferences. In practice, it is common to use some existing statistical software to obtain the least squares estimators (MSE) for the regression parameters and consider a non-biased estimator for the variance based on the residual squared sum of the residuals. Alternatively, we discuss the use of the maximum likelihood estimator (MLE) for the variance of the error (an unbiased estimator, but with smaller squared error) which may lead to different conclusions in terms of inferences. Two numerical examples are presented to illustrate the proposed method.*

**Keywords:** Least squares estimator. Maximum likelihood estimator. Linear regression models. Accuracy.

## 1 Introdução

Por que mais um artigo sobre regressão linear? Trata-se de um assunto muito explorado na literatura e bastante conhecido pela comunidade científica. A técnica de regressão linear é bastante utilizada na análise de dados quando tem-se duas ou mais variáveis relacionadas entre si (ver, por exemplo, MONTGOMERY E PECK,1992; MYERS,1990; NETER ET AL, 1985; WEISBERG, 2005).

Em geral, na análise de dados, utiliza-se um software estatístico disponível no mercado para a construção do modelo estatístico, visto que inicialmente é obtido os estimadores pontuais e intervalares, depois é feita uma diagnóstico da qualidade do modelo e por fim podemos fazer previsões com o modelo proposto.

Na prática, as precisões dos estimadores podem mudar as conclusões inferenciais usando métodos de estimação diferentes para a variância do erro associado ao modelo de regressão, especialmente na presença de pequenas amostras e quando o pesquisador considera um nível de significância pré-estabelecido (usualmente em 0,05) para a realização de um teste de hipóteses sobre um parâmetro de interesse para a tomada de decisões. Em muitos casos, um teste de hipóteses pode não levar à rejeição da hipótese de nulidade, quando usado método de mínimos quadrados para a estimação dos parâmetros de regressão e à rejeição da mesma hipótese de nulidade quando utilizamos o método de

estimação de máxima verossimilhança. Apesar dos estimadores pontuais serem iguais quando obtidos por meio de dois métodos de estimação, a precisão do estimador da variância do erro pode ser diferente, quando obtido por dois métodos de estimação.

Neste artigo, apresentamos uma pequena revisão dos métodos mais populares usados na estimação dos parâmetros de regressão linear e alguns exemplos ilustrando alguns conflitos existentes entre os resultados obtidos, como por exemplo, mudança no intervalo de confiança e rejeição ou não da hipótese de nulidade. Em especial, serão enfatizados modelos de regressão linear simples, apesar dos resultados serem facilmente estendidos para a situação de modelos de regressão linear múltipla.

No caso de uma regressão linear simples, para uma amostra de tamanho  $n$ , seja o modelo de regressão,

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (1)$$

em que  $i = 1, \dots, n$ ; a variável dependente é denotada por  $Y_i$ ; a variável independente associada a  $Y_i$  é denotada por  $x_i$  e  $\varepsilon$  é uma quantidade aleatória não-observada denotando o erro.

Em muitas situações, obtém-se as inferências de interesse considerando as respostas na escala original ou transformada, considerando-se os erros  $\varepsilon_i$  independentes com

uma distribuição normal com média zero e variância constante  $\sigma^2$ , isto é,  $\varepsilon_i \sim N(0, \sigma^2)$ . Em outras situações, assume-se que as quantidades aleatórias são não correlacionadas duas a duas sem assumir uma distribuição de probabilidade para as respostas  $Y_i$ , isto é, apenas assumindo que  $E(Y_i) = \beta_0 + \beta_1 x_i$  e  $\text{var}(Y_i) = \sigma^2$ .

Para a estimação dos parâmetros  $\beta_0, \beta_1$  e  $\sigma^2$ , a literatura apresenta várias alternativas dependendo da estrutura considerada para os dados: estimadores de mínimos quadrados, estimadores de máxima verossimilhança, estimadores não-viciados de variância mínima, estimadores Bayesianos, estimadores BLUE (estimadores lineares não-viciados ótimos) entre vários outros (ver por exemplo, SEBER E LEE, 2003; DRAPER E SMITH, 1981; BOLFARINE E SANDOVAL, 2001).

O presente artigo é organizado da seguinte forma: na seção 2, apresenta-se uma breve revisão dos métodos de estimação para os parâmetros de regressão; na seção 3, discute-se a precisão dos estimadores; na seção 4, mostra-se os intervalos de confiança e testes de hipóteses para o caso de modelos de regressão linear simples; na seção 5, há um breve sumário para modelos de regressão linear múltipla; na seção 6, duas ilustrações numéricas são introduzidas; e, finalmente na seção 7, apresenta-se algumas notas conclusivas.

## 2 Uma breve revisão dos métodos de estimação para os parâmetros de regressão

Inicialmente, assumo o modelo de regressão linear simples dado pela expressão (1).

### 2.1 Melhores estimadores lineares não-viciados (BLUE)

Assumo que  $Y_1, \dots, Y_n$  sejam variáveis aleatórias não correlacionadas duas a duas com médias  $E(Y_i) = \beta_0 + \beta_1 x_i$  e variância constante  $\text{var}(Y_i) = \sigma^2$ . Sem assumir qualquer distribuição de probabilidade paramétrica para os erros  $\varepsilon_i$ , o teorema de Gauss-Markov (SILVEY, 1970) estabelece que os estimadores de mínimos quadrados para  $\beta_0$  e  $\beta_1$  dados por

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \quad (2)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Em que  $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$ ,  $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ ,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  e  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  são os estimadores lineares não-viciados ótimos (BLUE).

### 2.2 Estimadores de mínimos quadrados (EMQ)

Os estimadores de mínimos quadrados para  $\beta_0$  e  $\beta_1$  em (1) são obtidos minimizando-se a soma de quadrados dos erros dada por

$$S = S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (3)$$

As primeiras derivadas parciais de  $S(\beta_0, \beta_1)$  com relação a  $\beta_0$  e  $\beta_1$  são dadas por

$$\frac{\partial S(\beta_0, \beta_1)}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) \quad (4)$$

$$\frac{\partial S(\beta_0, \beta_1)}{\partial \beta_1} = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i)$$

Resolvendo-se as equações  $\partial S(\beta_0, \beta_1)/\partial \beta_0 = 0$  e  $\partial S(\beta_0, \beta_1)/\partial \beta_1 = 0$ , encontramos os estimadores de mínimos quadrados (EMQ) para  $\beta_0$  e  $\beta_1$  dados por (2).

Em geral, um estimador da variância  $\sigma^2$  do erro (ver por exemplo, Seber e Lee, 2003; Draper e Smith, 1981) é dado por

$$\widehat{\sigma^2} = s^2 = \frac{SQR}{n - 2} \tag{5}$$

em que  $SQR = \sum_{i=1}^n \hat{\epsilon}_i^2$  é a soma de quadrados residual;  $\hat{\epsilon}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$  são os resíduos e  $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$  define a reta ajustada.

O estimador dado em (5) é um estimador não-viciado da variância  $\sigma^2$  do erro, isto é,  $E(s^2) = \sigma^2$ .

### 2.3 Estimadores de máxima verossimilhança (EMV)

Assumir que os erros  $\epsilon_i$  sejam independentes com uma distribuição normal com média zero e variância constante  $\sigma^2$ , isto é,  $\epsilon_i \sim N(0, \sigma^2)$ . Isso implica que as respostas  $Y_i$  seguem distribuições normais  $N(\beta_0 + \beta_1 x_i, \sigma^2)$  para  $i = 1, \dots, n$ .

Com essa suposição, a função de verossimilhança para  $\beta_0, \beta_1$  e  $\sigma$  é dada por

$$L(\beta_0, \beta_1, \sigma) = (2\pi\sigma^2)^{-n/2} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \right] \tag{6}$$

Os estimadores de máxima verossimilhança para  $\beta_0, \beta_1$  e  $\sigma$  são obtidos maximizando-se  $L(\beta_0, \beta_1, \sigma)$  ou equivalentemente, o logaritmo da função de verossimilhança  $l(\beta_0, \beta_1, \sigma) = \log[L(\beta_0, \beta_1, \sigma)]$  em relação a  $\beta_0, \beta_1$  e  $\sigma$ , dado por

$$l(\beta_0, \beta_1, \sigma) \propto -n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \tag{7}$$

As primeiras derivadas parciais de  $l(\beta_0, \beta_1, \sigma)$  com relação a  $\beta_0, \beta_1$  e  $\sigma$  são dadas por

$$\begin{aligned} \partial l / \partial \beta_0 &= 1/\sigma^2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) \\ \partial l / \partial \beta_1 &= 1/\sigma^2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i) \\ \partial l / \partial \sigma &= -n/\sigma + 1/\sigma^3 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \end{aligned} \tag{8}$$

Resolvendo-se o sistema de equações  $\partial l / \partial \beta_0 = 0, \partial l / \partial \beta_1 = 0$  e  $\partial l / \partial \sigma = 0$ , encontram-se os estimadores de máxima verossimilhança (EMV) para  $\beta_0$  e  $\beta_1$  dados por (2) e

$$\widetilde{\sigma^2} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \tag{9}$$

Observe que os EMV coincidem com os EMV para  $\beta_0$  e  $\beta_1$ , mas o EMV para  $\sigma^2$  é dado por  $\widetilde{\sigma^2} = SQR/n$  em vez do estimador  $\widehat{\sigma^2} = SQR/(n-2)$  dado em (5), isto é,  $\widetilde{\sigma^2}$  é um estimador viciado para  $\sigma^2$ , em que  $E(\widetilde{\sigma^2}) = (n-2)\sigma^2/n$  (ver por exemplo, MOOD E GRAYBILL 1974; OU BICKELL E DOCKSUN, 2001).

### 3 Precisão dos estimadores

Da suposição que as respostas  $Y_i$  têm distribuições normais  $N(\beta_0 + \beta_1 x_i, \sigma^2)$  para  $i = 1, \dots, n$ , temos os seguintes resultados:

A distribuição conjunta para  $Y_1, \dots, Y_n$  pertence à família exponencial com três parâmetros (ver por exemplo, MOOD, GRAYBILL E BOES, 1970; OU BICKELL E DOCKSON, 1980). Assim  $(\sum_{i=1}^n Y_i^2, \sum_{i=1}^n Y_i, \sum_{i=1}^n x_i Y_i)$  é um conjunto de estatísticas suficientes e completas.

A função geradora de momentos (fgm) para  $\widehat{\theta}_1 = (\widehat{\beta}_0 - \beta_0)/\sigma$ ,  $\widehat{\theta}_2 = (\widehat{\beta}_1 - \beta_1)/\sigma$  e  $\widehat{\theta}_3 = n\widehat{\sigma}^2/\sigma^2$  ( $\widehat{\sigma}^2$  dado em (9)) dada por,

$$m(t_1, t_2, t_3) = E\{\exp[t_1\widehat{\theta}_1 + t_2\widehat{\theta}_2 + t_3\widehat{\theta}_3]\} \quad (10)$$

pode ser decomposta na forma  $m(t_1, t_2, t_3) = m_1(t_1, t_2) m_2(t_3)$  onde  $m_1(t_1, t_2)$  é a fgm de uma distribuição normal bivariada com média  $(\beta_0, \beta_1)$  e matriz de variâncias-covariância dada por

$$\begin{pmatrix} \sigma^2 \sum_{i=1}^n x_i^2 / nS_{xx} & -\sigma^2 \bar{x} / S_{xx} \\ -\sigma^2 \bar{x} / S_{xx} & \sigma^2 / S_{xx} \end{pmatrix} \quad (11)$$

em que  $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$  e  $m_2(t_3)$  é a fgm de uma distribuição qui-quadrado com  $n-2$  graus de liberdade, isto é,

$$n\widehat{\sigma}^2/\sigma^2 = (1/\sigma^2) \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 \sim \chi_{n-2}^2 \quad (12)$$

Observe também que as variáveis aleatórias associadas com a fgm conjunta  $m_1(t_1, t_2)$

são independentes da variável aleatória associada com a fgm  $m_2(t_3)$ .

De (12), observa-se que  $E(n\widehat{\sigma}^2/\sigma^2) = n - 2$  e  $\text{var}(n\widehat{\sigma}^2/\sigma^2) = 2(n - 2)$ , isto é,  $E(\widehat{\sigma}^2) = (n-2)\sigma^2/n$  (um estimador viciado para  $\sigma^2$ ) e  $\text{var}(\widehat{\sigma}^2) = 2(n-2)\sigma^4/n^2$ .

Como  $\widehat{\sigma}^2 = SQR/n$ ,  $E(SQR) = (n - 2)\sigma^2$  e  $\text{var}(SQR) = 2(n - 2)\sigma^4$ . Dessa forma,  $\widehat{\sigma}^2 = SQR/(n-2)$  (dado em (5)) tem  $E(\widehat{\sigma}^2) = \sigma^2$  (um estimador não-viciado para  $\sigma^2$ ) e  $\text{var}(\widehat{\sigma}^2) = 2\sigma^4/(n-2)$ . Assim, observa-se que,  $\text{var}(\widehat{\sigma}^2) < \text{var}(\sigma^2)$ , pois,  $2(n-2)/n^2 = 2[(1-2/n)^2]/(n-2) < 2/(n-2)$ , para  $n > 2$ .

#### 3.1 Erros quadráticos médios (EQM) dos estimadores

O erro quadrático médio (EQM) de um estimador denotado por  $\widehat{\theta}$  de um parâmetro  $\theta$  é dado por

$$\text{EQM}(\widehat{\theta}) = E[(\widehat{\theta} - \theta)^2] = \text{var}(\widehat{\theta}) + [\text{vício}(\widehat{\theta}, \theta)]^2, \quad (13)$$

em que o vício  $(\widehat{\theta}, \theta) = E(\widehat{\theta}) - \theta$ .

Com  $\theta = \sigma^2$ , observa-se que  $\text{vício}(\widehat{\sigma}^2, \sigma^2) = 0$ , isto é,  $\text{EQM}(\widehat{\sigma}^2) = \text{var}(\widehat{\sigma}^2) = 2\sigma^4/(n - 2)$ . Da mesma forma, como  $\text{vício}(\widehat{\sigma}^2, \sigma^2) = -2\sigma^2/n$  e, tem-se que  $\text{EQM}(\widetilde{\sigma}^2) = \text{var}(\widetilde{\sigma}^2) + [-2\sigma^2/n]^2 = 2\sigma^4/n$ .

Desses resultados, observa-se que o estimador  $\widetilde{\sigma}^2$  (dado em (9)), é preferível ao estimador  $\widehat{\sigma}^2$  (dado em (5)), apesar de  $\widehat{\sigma}^2$  ser um estimador não-viciado de  $\sigma^2$ , pois tem menor vício. Isso pode levar a melhores precisões nas inferências sobre os parâmetros de regressão.

### 3.2 Resultados gerais para grandes amostras

Observe que também é possível obter as variâncias e covariâncias dos EMV para  $\beta_0, \beta_1$  e  $\sigma$ , a partir do resultado geral assintótico,

$$\hat{\theta} \sim N\{\theta, I^{-1}(\theta)\}, \tag{14}$$

em que  $\theta = (\theta_1, \dots, \theta_k)$  é um vetor de parâmetros de dimensão  $k$ ;  $\hat{\theta}$  é o vetor dos EMV de  $\theta_j$ ,  $j = 1, \dots, k$  e  $I(\theta)$  é a matriz de informação de Fisher,

$$I(\theta) = E\left(-\frac{\partial^2 l(\theta)}{\partial \theta_i \partial \theta_j}\right) \tag{15}$$

para  $i, j = 1, \dots, k$  e  $l(\theta)$  é o logaritmo da função de verossimilhança para  $\theta$ .

As segundas derivadas parciais do logaritmo da função de verossimilhança para  $\beta_0, \beta_1$  e  $\sigma$  (ver expressão (7)), são dadas t

$$\begin{aligned} \frac{\partial^2 l}{\partial \beta_0^2} &= -n / \sigma^2; \quad \frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} = \\ &= -n\bar{x} / \sigma^2; \\ \frac{\partial^2 l}{\partial \beta_0 \partial \sigma} &= -2 / \sigma^3 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i); \\ \frac{\partial^2 l}{\partial \beta_1^2} &= -\sum_{i=1}^n x_i^2 / \sigma^2; \\ \frac{\partial^2 l}{\partial \beta_1 \partial \sigma} &= -2 / \sigma^3 \sum_{i=1}^n x_i (y_i - \beta_0 - \\ &= \beta_1 x_i); \\ \frac{\partial^2 l}{\partial \sigma^2} &= n / \sigma^3 - 3 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 / \\ &= \sigma^4. \end{aligned} \tag{16}$$

De (16), obtém-se os valores esperados,  $E(-\partial^2 l / \partial \beta_0^2) = n / \sigma^2$ ;  $E(-\partial^2 l / \partial \beta_0 \partial \beta_1) = n\bar{x} / \sigma^2$ ;  $E(-\partial^2 l / \partial \beta_0 \partial \sigma) = 0$ ;  $E(-\partial^2 l / \partial \beta_1^2) = \sum_{i=1}^n (x_i)^2 / \sigma^2$ ;  $E(-\partial^2 l / \partial \beta_1 \partial \sigma) = 0$  e  $E(-\partial^2 l / \partial \sigma^2) = 2n / \sigma^2$ , os elementos da matriz de informação de Fisher

$I(\beta_0, \beta_1, \sigma)$  e assim a matriz de variâncias e covariâncias para os EMV para  $\beta_0, \beta_1$  e  $\sigma$ , dada por,

$$I^{-1}(\beta_0, \beta_1, \sigma) = \begin{pmatrix} \sigma^2 \sum_{i=1}^n x_i^2 / nS_{xx} & -\sigma^2 \bar{x} / S_{xx} & 0 \\ -\sigma^2 \bar{x} / S_{xx} & \sigma^2 / S_{xx} & 0 \\ 0 & 0 & \sigma^2 / 2n \end{pmatrix} \tag{17}$$

De (17), tem-se:

$$\begin{aligned} \text{var}(\hat{\beta}_0) &= \sigma^2 \sum_{i=1}^n x_i^2 / nS_{xx}; \\ \text{var}(\hat{\beta}_1) &= \sigma^2 / S_{xx}; \\ \text{var}(\hat{\sigma}) &= \sigma^2 / 2n; \\ \text{cov}(\hat{\beta}_0, \hat{\beta}_1) &= -\sigma^2 \bar{x} / S_{xx}, \end{aligned} \tag{18}$$

$$\text{em que } S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 .$$

## 4 Intervalos de confiança e testes de hipóteses (regressão linear simples)

Dos resultados dados na seção 3, obtêm-se os intervalos de confiança para os parâmetros de interesse.

### 4.1 Intervalo de confiança para $\beta_0$

Considerar os seguintes resultados:

$$Z = \frac{(\hat{\beta}_0 - \beta_0) \sqrt{nS_{xx}}}{\sigma \sqrt{\sum_{i=1}^n x_i^2}} \sim N(0,1) \text{ (distribuição}$$

normal padronizada);

$$U = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2 \text{ (distribuição qui-quadrado com } n-2 \text{ graus de liberdade);}$$

Z e U são independentes;

$$T = \frac{(\hat{\beta}_0 - \beta_0) \sqrt{nS_{xx}}}{\hat{\sigma} \sqrt{\sum_{i=1}^n x_i^2}} \sim t_{n-2} \text{ (distribuição t de}$$

Student com n-2 graus de liberdade).

Assim, um intervalo de confiança 100 (1- $\alpha$ )% para  $\beta_0$  é dado por,

$$\left( \hat{\beta}_0 - t_{(n-2; \frac{\alpha}{2})} \hat{\sigma} \sqrt{\frac{\sum_{i=1}^n x_i^2}{nS_{xx}}}; \hat{\beta}_0 + t_{(n-2; \frac{\alpha}{2})} \hat{\sigma} \sqrt{\frac{\sum_{i=1}^n x_i^2}{nS_{xx}}} \right) \quad (19)$$

em que  $\hat{\sigma}$  é dado por (5) ou por (9);  $t_{n-2; \alpha/2}$  é determinado de uma tabela da distribuição t de Student com  $n - 2$  graus de liberdade, considerando  $P(T > t_{n-2; \alpha/2}) = \alpha/2$ .

## 4.2 Intervalo de confiança para $\beta_1$

Considerar os seguintes resultados:

- $Z = \frac{(\hat{\beta}_1 - \beta_1) \sqrt{S_{xx}}}{\sigma} \sim N(0,1)$ ;
- $U = \frac{(n-2) \hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2$ ;
- Z e U são independentes;
- $T = \frac{(\hat{\beta}_1 - \beta_1) \sqrt{S_{xx}}}{\hat{\sigma}} \sim t_{n-2}$ ;

Assim, um intervalo de confiança 100(1- $\alpha$ ) % para  $\beta_1$  é dado por,

$$\left( \hat{\beta}_1 - t_{(n-2; \frac{\alpha}{2})} \frac{\hat{\sigma}}{S_{xx}}; \hat{\beta}_1 + t_{(n-2; \frac{\alpha}{2})} \frac{\hat{\sigma}}{S_{xx}} \right) \quad (20)$$

## 4.3 Intervalo de confiança para $\mu(x^*) = \beta_0 + \beta_1 x^*$ , em que $\hat{\mu}(x^*) = \hat{\beta}_0 + \hat{\beta}_1 x^*$

Considerar os seguintes resultados:

$$Z = \frac{\mu(x^*) - \hat{\mu}(x^*)}{\sqrt{\text{var}[\hat{\mu}(x^*)]}} \sim N(0,1) \text{ onde } \text{var}[\hat{\mu}(x^*)] = \sigma^2 \{1/n + (\bar{x} - x^*)^2 / S_{xx}\}$$

$$U = \frac{(n-2) \hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2;$$

Z e U são independentes;

$$T = \frac{(\hat{\beta}_0 + \hat{\beta}_1 x^*) - (\beta_0 + \beta_1 x^*)}{DP(\hat{\beta}_0 + \hat{\beta}_1 x^*)} \sim t_{n-2}; \text{ em que}$$

$$DP(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(\bar{x} - x^*)^2}{S_{xx}}}$$

Assim, um intervalo de confiança 100(1- $\alpha$ )% para  $\beta_0 + \beta_1 x^*$  é dado por,

$$\left( \hat{\beta}_0 + \hat{\beta}_1 x^* - t_{(n-2; \frac{\alpha}{2})} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(\bar{x} - x^*)^2}{S_{xx}}}; \hat{\beta}_0 + \hat{\beta}_1 x^* + t_{(n-2; \frac{\alpha}{2})} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(\bar{x} - x^*)^2}{S_{xx}}} \right) \quad (21)$$

Observe também que  $\sigma$  pode ser estimado por (5) ou (9).

Considerar o teste de hipóteses  $H_0 : \beta_1 = 0$  versus  $H_1 : \beta_1 \neq 0$ . A estatística do teste é dada por,

$$T = \frac{\hat{\beta}_1 \sqrt{S_{xx}}}{\hat{\sigma}} \sim t_{n-2} \quad (22)$$

em que  $\sigma$  pode ser estimado por (5) ou (9); dessa forma podemos ter valores diferentes para o valor-p, especialmente para amostras pequenas, considerando cada estimador de  $\sigma^2$ .

### 5 Algumas considerações sobre regressão linear múltipla

A expressão do modelo linear geral de regressão é dada por

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i. \tag{23}$$

Sejam

$$\mathbf{Y}_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{X}_{n \times p} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1,p-1} \\ 1 & X_{21} & \dots & X_{2,p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{n,p-1} \end{bmatrix}, \quad \boldsymbol{\beta}_{p \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} \quad e$$

Em termos matriciais, o modelo de regressão linear geral é dado por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{24}$$

onde  $\boldsymbol{\varepsilon}$  é um vetor de variáveis aleatórias independentes e normalmente distribuídas com esperança (média),  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$  e matriz de variância-covariância dada por

$$\sigma^2(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2 \mathbf{I}$$

Assim, o vetor das observações  $\mathbf{Y}$  tem esperança e variância dadas por

$$E_{n \times 1}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \quad e \quad \sigma_{n \times n}^2 = \sigma^2 \mathbf{I}. \tag{25}$$

Os estimadores de mínimos quadrados são dados por

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \tag{26}$$

E também,  $\text{var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ . Um estimador para a variância do erro é dado por

$$\widehat{\sigma^2} = s^2 = \frac{SQR}{n - p}, \tag{27}$$

em que SQR é a soma de quadrados dos resíduos,

$$\mathbf{e} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\mathbf{b} \tag{28}$$

Considerando estimadores de máxima verossimilhança, o EMV para  $\boldsymbol{\beta}$  é dado por (26) e substituímos  $n - p$  por  $n$  em (27).

#### 5.1 Intervalo de confiança para a predição

Pode-se obter o intervalo de confiança para a resposta média em um ponto particular da amostra. Seja o vetor  $\mathbf{x}^* = (1, x_1^*, x_2^*, \dots, x_{p-1}^*)$  um valor fixado para as variáveis independentes.

A resposta média para este ponto é,

$$E(\mathbf{Y} \mid \mathbf{x}^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_{p-1} x_{p-1}^* = \mathbf{x}^* \boldsymbol{\beta} \tag{29}$$

O estimador da média neste ponto é  $E(\widehat{\mathbf{Y}} \mid \mathbf{x}^*) = \mathbf{x}^* \widehat{\boldsymbol{\beta}}$ . A variância de  $\mathbf{x}^* \widehat{\boldsymbol{\beta}}$  é dada por  $\sigma^2 \mathbf{x}^{*'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^*$ . Assim temos que

$$Z = \frac{\mathbf{x}^* \widehat{\boldsymbol{\beta}} - \mathbf{x}^* \boldsymbol{\beta}}{\sigma \sqrt{\mathbf{x}^{*'} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}^*}} \sim N(0; 1).$$

$$U = \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p} \text{ onde } Z \text{ e } U \text{ são}$$

independentes

Assim,

$$T = \frac{x^*\hat{\beta} - x^*\beta}{\hat{\sigma}\sqrt{x^{*'}(X'X)x^*}} \sim t_{n-p}. \tag{30}$$

Portanto, o intervalo com confiança  $100(1 - \alpha)\%$  para a resposta média  $E(Y|x^*)$  no ponto  $x^*$  é dado por

$$\left( x^*\hat{\beta} - t_{\left(\frac{\alpha}{2}; n-p\right)}\hat{\sigma}\sqrt{x^{*'}(X'X)x^*}; x^*\hat{\beta} + t_{\left(\frac{\alpha}{2}; n-p\right)}\hat{\sigma}\sqrt{x^{*'}(X'X)x^*} \right) \tag{31}$$

## 6 Ilustrações numéricas

### 6.1 Um exemplo com regressão linear simples

Considere os dados relacionados a um experimento industrial (Tabela 1) onde  $x$  denota uma covariável relacionada com a resposta  $y$ , ( $n = 12$  observações). Na Figura 1, tem-se o gráfico de  $y$  versus  $x$ . Observa-se uma boa relação linear entre as variáveis.

Usando o software MINITAB® versão 16, inicialmente considera-se o ajuste da regressão linear (1) usando método de mínimos quadrados e o estimador (5) para a variância do erro, encontrando a reta ajustada dada por,

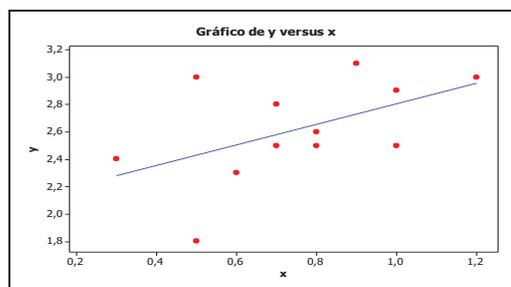
$$\hat{y} = 2,06 + 0,746 x \tag{32}$$

Na Tabela 2, tem-se os EMQ dos parâmetros de regressão; os erros-padrões dos estimadores; os valores observados para a

estatística  $t$  de Student e o valor- $p$  associado ao teste de hipóteses  $H_0:\beta = 0$  versus  $H_1:\beta \neq 0$ .

**Tabela 1** - Dados de regressão

Observação	$x$	$y$	Observação	$x$	$y$
1	0,5	1,8	7	0,7	2,8
2	0,3	2,4	8	0,8	2,5
3	0,5	3,0	9	0,9	3,1
4	0,7	2,5	10	1,0	2,9
5	0,6	2,3	11	1,0	2,5
6	0,8	2,6	12	1,2	3,0



**Figura 1** - Gráfico de dispersão  $y$  versus  $x$

**Tabela 2** - Estimadores de mínimos quadrados (EMQ) dos parâmetros

Preditor	EMQ	EP	T	$p$ -valor
Constante	2,068	0,3105	6,62	<0,001
$x$	0,7465	0,3938	1,90	0,087

(EP são erros-padrões das estimativas;  $T$  é o valor de referência para a distribuição  $t$  de Student).

O desvio-padrão do erro é estimado (ver (5)) por  $S = 0,331818$ . Observa-se dos gráficos dos resíduos dados na Figura 2 que as suposições do modelo são verificadas (normalidade e variância constante).

A partir desses resultados, e fixando-se um nível de significância igual a 5%, não rejeita-

se a hipótese  $H_0: \beta_1 = 0$ , pois o valor- $p$  (0,087) é maior do que 0,05.

Usando o método de máxima verossimilhança, e o software MINITAB, encontramos reta ajustada dada por

$$\hat{y} = 2,06 + 0,746 x \quad (33)$$

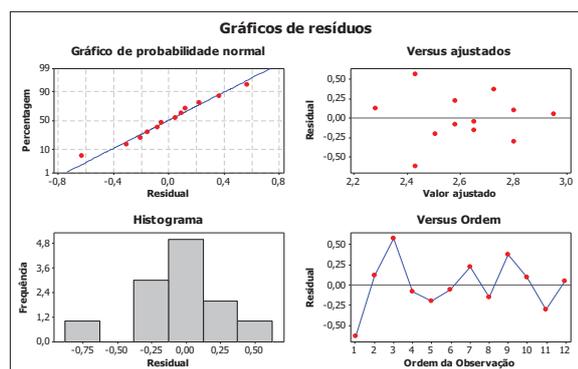


Figura 2 - Gráficos de resíduos (regressão linear simples)

Na Tabela 3, tem-se os EMV dos parâmetros de regressão; os erros-padrões dos estimadores e o valor- $p$  associado ao teste de hipóteses  $H_0: \beta = 0$  versus  $H_1: \beta \neq 0$ .

Tabela 3 - Estimadores de máxima verossimilhança (EMV) dos parâmetros

Preditor	EMV	EP	$p$ -valor
Constante	2,0568	0,2834	<0,001
x	0,7465	0,3595	0,038

Nota: EP são erros-padrões das estimativas

Neste caso, o desvio-padrão do erro é estimado (ver (5) com  $n - 2$  substituído por  $n = 12$ ) por  $S = 0,3029$ . Neste caso, observa-se melhor precisão para os estimadores obtidos quando comparados com os estimadores obtidos pelo método de mínimos quadrados. Também se observa que deve-se rejeitar a hipótese  $H_0$ :

$\beta_1 = 0$ , pois o valor- $p$  (0,038) é menor do que 0,05 (nível de significância fixado pelo pesquisador).

Adicionalmente, é possível determinar intervalos de confiança para os valores de previsão  $\mu(x^*) = \beta_0 + \beta_1 x^*$  (ver seção 4.3) assumindo diferentes valores para  $x^*$  e usando as duas estimativas para o desvio-padrão do erro (dados por (5) ou (9) relacionados respectivamente aos métodos de mínimos quadrados e máxima verossimilhança). Na Tabela 4 temos esses intervalos de confiança. Observa-se maior precisão (intervalos de confiança 95% mais curtos) considerando-se as estimativas de máxima verossimilhança para o desvio-padrão dos erros.

Tabela 4 - Intervalos de confiança 95% para

$$\mu(x^*) = \beta_0 + \beta_1 x^*$$

$x^*$	IC 95% (EMQ)	IC 95% (EMV)
0,25	(1,758516 ; 2,734484)	(1,801043 ; 2,691957)
0,35	(1,910248 ; 2,731952)	(1,946054 ; 2,696146)
0,45	(2,056767 ; 2,734633)	(2,086305 ; 2,705095)
0,55	(2,193968 ; 2,746632)	(2,218051 ; 2,722549)
0,65	(2,314142 ; 2,775658)	(2,334253 ; 2,755547)
0,75	(2,406085 ; 2,832915)	(2,424684 ; 2,814316)
0,85	(2,463342 ; 2,924858)	(2,483453 ; 2,904747)
0,95	(2,492368 ; 3,045032)	(2,516451 ; 3,020949)
1,05	(2,504367 ; 3,182233)	(2,533905 ; 3,152695)

Assim observa-se resultados diferentes usando os dois métodos de estimação. Esse resultado pode ser de grande interesse prático.

## 6.2 Um exemplo com regressão linear múltipla

Considere dados de uma indústria alimentícia do Estado de São Paulo. A indústria alimentícia estudada localiza-se no interior do Estado de São Paulo e atua no processamento de frutas tropicais para fabricação de sucos concentrados. Vários fatores são observados na linha de produção afetando o desempenho do equipamento da indústria. A medição de paradas e perdas de tempo foi monitorada na última etapa – envase – considerada gargalo da linha. A empresa trabalha com três turnos de produção. Os tempos de paradas devido a falhas dos equipamentos foram classificadas em sete tipos: máquinas M1 (descarga de frutas); M2 (moagem); M3 (despolpamento); M4 (concentração); M5 (envase); M6 (resíduos) e M7 (utilidades), que podem ser afetadas por alguns fatores como safra (1 para período 1010/2011) e 2 para período 2011/2012); tipo de equipamento (E para máquina elétrica); M para máquina mecânica) e turno (1 (primeiro), 2 (segundo), 3 (terceiro) (Dados apresentados no quadro 1).

Para uma análise estatística dos dados, foi considerado um modelo de regressão linear múltipla assumindo erros independentes com distribuição normal  $N(0, \sigma^2)$  com variância  $\sigma^2$  constante. As variáveis respostas são dadas pelos logaritmos dos tempos de paradas e as variáveis explanatórias independentes são dadas por:

- uma variável “dummy” ou categórica para tipo de máquina (mecânica=1, 0=outra parte);

- uma variável “dummy” ou categórica indicadora para safra( safra1=1, 0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 1 (M1=1, 0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 2(M2=1,0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 3(M3=1,0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 4(M4=1,0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 5(M5=1,0=outra parte);
- uma variável “dummy” ou categórica indicadora para máquina 6(M6=1,0=outra parte);
- uma variável “dummy” ou categórica indicadora para turno 1 (turno1=1, 0=outra parte);
- uma variável “dummy” ou categórica indicadora para turno 2 (turno 2=1, 0=outra parte)

O modelo de regressão linear múltipla ajustado por mínimos quadrados (uso do software MINITAB® versão 16), é dado por:

$$\widehat{\log(\text{tempo parada})} = 4,48 - 0,249 (\text{indicador safra1=1}) - 0,831 (\text{indicador máquina M1=1; outro} = 0) - 1,57 (\text{indicador máquina M2=1; outro} = 0) - 1,01 (\text{indicador máquina M3=1; outro} = 0) -$$

Quadro 1 - Dados para o exemplo 2 (Pandolfi,2013)

obs safra tipo turno máq tempo	obs safra tipo turno máq tempo	obs safra tipo turno máq tempo
1 1 1 1 1 100	46 1 0 3 5 607	91 1 1 2 6 67
2 1 0 1 1 26	47 1 0 3 5 10	92 1 1 2 6 154
3 1 1 2 1 28	48 1 0 2 5 6	93 1 0 2 7 200
4 1 1 2 1 12	49 1 0 1 5 15	94 0 1 1 1 27
5 1 1 3 1 20	50 1 0 1 5 46	95 0 1 1 2 36
6 1 1 1 1 4	51 1 0 1 5 17	96 0 0 3 3 8
7 1 1 3 1 75	52 1 0 3 5 14	97 0 1 1 4 120
8 1 1 1 1 30	53 1 0 3 5 20	98 0 1 1 4 4
9 1 1 1 1 17	54 1 0 2 5 30	99 0 1 3 4 50
10 1 1 2 2 6	55 1 0 3 5 85	100 0 1 2 4 56
11 1 1 2 2 23	56 1 0 1 5 35	101 0 1 2 4 26
12 1 1 3 2 7	57 1 0 2 5 34	102 0 1 1 4 32
13 1 1 3 2 108	58 1 0 3 5 4	103 0 1 1 5 15
14 1 1 2 2 6	59 1 0 2 5 10	104 0 0 3 5 279
15 1 1 1 2 14	60 1 0 2 5 5	105 0 0 2 5 13
16 1 1 1 2 7	61 1 0 1 5 7	106 0 0 1 5 36
17 1 1 2 2 10	62 1 0 1 5 42	107 0 1 1 5 18
18 1 1 2 2 9	63 1 1 1 5 8	108 0 0 1 5 10
19 1 1 2 2 8	64 1 1 1 5 40	109 0 0 2 5 50
20 1 1 3 3 64	65 1 1 2 5 7	110 0 0 3 5 37
21 1 1 1 3 5	66 1 1 2 5 42	111 0 0 3 5 7
22 1 1 3 3 32	67 1 0 3 5 19	112 0 0 3 5 8
23 1 0 3 3 36	68 1 0 3 5 5	113 0 0 2 5 48
24 1 0 1 3 60	69 1 1 3 6 16	114 0 0 3 5 8
25 1 0 2 3 10	70 1 1 2 6 15	115 0 0 1 5 5
26 1 0 1 3 28	71 1 1 3 6 15	116 0 0 1 5 49
27 1 0 1 3 75	72 1 1 1 6 24	117 0 0 2 5 24
28 1 1 3 3 24	73 1 1 2 6 15	118 0 0 2 5 30
29 1 1 2 3 20	74 1 1 2 6 184	119 0 0 1 5 16
30 1 1 2 3 22	75 1 1 1 6 28	120 0 1 3 5 110
31 1 1 3 3 34	76 1 1 3 6 41	121 0 1 3 5 33
32 1 1 3 3 18	77 1 1 3 6 53	122 0 1 2 5 20
33 1 0 3 3 29	78 1 1 3 6 18	123 0 1 1 5 12
34 1 0 1 3 5	79 1 1 1 6 11	124 0 1 2 5 30
35 1 0 1 4 2	80 1 1 1 6 30	125 0 1 1 5 20
36 1 1 2 4 61	81 1 1 2 6 31	126 0 1 3 6 138
37 1 1 2 4 27	82 1 1 1 6 26	127 0 0 1 7 15
38 1 1 1 4 50	83 1 1 2 6 22	
39 1 0 3 4 109	84 1 1 2 6 86	
40 1 1 3 4 28	85 1 1 3 6 37	
41 1 1 2 4 18	86 1 1 1 6 5	
42 1 1 1 4 8	87 1 1 1 6 31	
43 1 1 2 4 6	88 1 1 3 6 17	
44 1 1 1 4 5	89 1 1 3 6 21	
45 1 1 3 5 70	90 1 1 3 6 12	

1,05(indicador máquina M4=1; outro = 0)-  
 1,08(indicador máquina M5=1; outro = 0) -  
 0,776(indicador máquina M6=1; outro = 0) -  
 0,497(indicador turno 1=1; outro = 0) -  
 0,207(indicador turno 2=1; outro = 0) + 0,118(tipo  
 máquina mecânica =1; outro = 0)

Na Tabela 5, têm-se as estimativas obtidas, os erros-padrões das estimativas de mínimos quadrados, os valores de referência da distribuição t de Student e os valores-p associados a cada parâmetro de regressão.

**Tabela 5** - EMQ para os coeficientes de regressão para o modelo de regressão múltipla considerando os logaritmos dos tempos de paradas

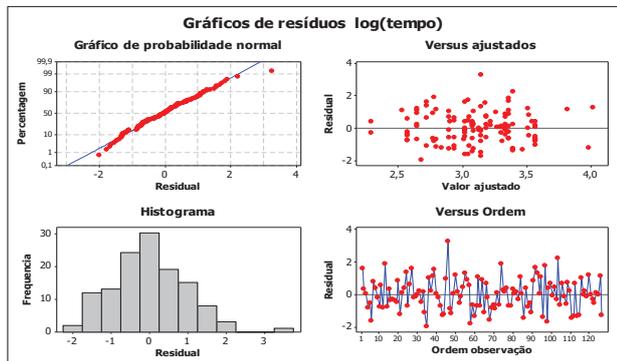
preditor	EMQ	EP	T	p-valor
Constante	4,4795	0,7341	6,1	<0,001
safra1=1	-0,24	0,225	-1,11	0,271
máquinaM1=1	-0,83	0,810	-1,03	0,307
máquinaM2=1	-1,56	0,809	-1,94	0,055
máquinaM3=1	-1,01	0,767	-1,32	0,190
máquinaM4=1	-1,05	0,775	-1,36	0,177
máquinaM5=1	-1,08	0,720	-1,51	0,135
máquinaM6=1	-0,77	0,783	-0,99	0,324
turno 1=1	-0,49	0,217	-2,28	0,024
turno 2=1	-0,20	0,226	-0,91	0,364
tipo máquina				
mecânica=1	0,117	0,249	0,47	0,637

Nota: EMQ são estimadores de mínimos quadrados; EP são erros-padrões das estimativas;T é o valor de referencia para a distribuição t de Student

A partir dos resultados da Tabela 5, observa-se que o fator turno leva a uma diferença significativa entre os tempos de paradas (escala logarítmica), pois o valor-p associado ao turno 1, quando comparado com o turno de referência 3 é igual a 0,024 (significativo em um nível de significância igual a 5%) isto é, o turno 1 tem tempos menores de paradas das máquinas pois o coeficiente de regressão estimado e associado à variável “dummy”, indicadora de turno 1 tem sinal negativo. As outras covariáveis não indicam efeitos significativos nas respostas (valor-p > 0,05). O desvio-padrão do erro é estimado (ver (5)) por S = 0,988619.

Na Figura 3, têm-se os gráficos dos resíduos do modelo de regressão ajustado. Observa-se que as suposições para a validade das inferências são verificadas (normalidade dos resíduos e variância constante dos erros).

Na Tabela 6, têm-se os EMV para os parâmetros de regressão, os erros-padrões dos EMV e os valores-p associados a cada parâmetro de regressão. A partir dos resultados da Tabela 6, observa-se que o fator turno 1 e máquina M2 levam a diferenças significativas entre os tempos de paradas (escala logarítmica), pois o valor-p associado ao turno 1, quando comparado com os outros turnos é igual a 0,017 (significativo em um nível de significância igual à 5%) e máquina M2 tem valor-p igual à 0,043. As outras covariáveis não indicam efeitos significativos nas respostas (valor-p > 0,05).



**Figura 3** - Gráficos de resíduos (regressão linear múltipla)

Aqui o desvio-padrão do erro é estimado (ver (27) com  $n - 11 = 116$  substituído por  $n = 127$ ) por  $S = 0,944835$ . Neste caso, observa-se melhor precisão para os estimadores obtidos quando comparados com os estimadores obtidos pelo método de mínimos quadrados.

Intervalos de confiança para os valores de previsão  $\beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_{p-1} x_{p-1}^* = \mathbf{x}^* \boldsymbol{\beta}$  (Ver seção 5.1) assumindo diferentes valores para  $\mathbf{x}^*$  e usando as duas estimativas para o desvio-padrão do erro (relacionados respectivamente aos métodos de mínimos quadrados e máxima verossimilhança) são dados por (31). Observar que como  $n > 100$ , podemos construir intervalos de confiança assumindo a aproximação normal para a distribuição  $t$  de Student. Na Tabela 7 temos intervalos de confiança para a previsão assumindo várias combinações das variáveis independentes. Observa-se maior precisão (intervalos de confiança mais curtos) considerando-se as estimativas de máxima verossimilhança para o desvio-padrão dos erros em todos os casos.

**Tabela 6** - EMV para os coeficientes de regressão para o modelo de regressão múltipla considerando os logaritmos dos tempos de paradas

Preditor	EMV	EP	p-valor
Constante	4,479	0,701	<0,001
safra1=1	-0,249	0,2154	0,247
máquina M1=1	-0,830	0,7743	0,283
máquina M2=1	-1,568	0,7733	0,043
máquina M3=1	-1,010	0,7336	0,168
máquina M4=1	-1,054	0,7415	0,155
máquina M5=1	-1,086	0,6886	0,115
máquina M6=1	-0,776	0,7490	0,300
turno 1=1	-0,496	0,2081	0,017
turno 2=1	-0,206	0,2166	0,340
Tipo máquina			
mecânica=1	0,1179	0,2382	0,621

Nota: EMV são estimadores de mínimos quadrados; EP são erros-padrões das estimativas

**Tabela 7** - Intervalos de confiança 95% para predições em várias combinações das variáveis independentes

safra	1	1	1	1	1	1	1	0	0
M1	1	1	1	0	0	0	0	0	0
M2	0	0	0	0	0	0	0	1	0
M3	0	0	0	0	0	0	0	0	0
M4	0	0	0	0	0	0	0	0	0
M5	0	0	0	1	1	0	0	0	1
M6	0	0	0	0	0	1	1	0	0
tipo	1	1	1	0	0	1	1	1	1
Turno1	1	0	1	0	1	0	1	1	1
Turno2	0	1	0	0	0	0	0	0	0
LI-EMQ	2,385	2,623	2,385	2,702	2,203	3,122	2,605	1,786	2,479
LS-EMQ	3,656	3,998	3,656	3,590	3,095	4,022	3,545	3,278	3,553
LI-EMV	2,413	2,653	2,413	2,722	2,222	3,142	2,626	1,819	2,503
LS-EMV	3,628	3,968	3,628	3,570	3,075	4,002	3,524	3,245	3,529

## 7 Algumas conclusões finais

O uso de diferentes métodos de estimação para os parâmetros de regressão linear com as suposições usuais (normalidade dos resíduos, variância constante) pode levar a diferentes conclusões em termos de inferências obtidas. Na prática é comum usarmos softwares estatísticos para obter os estimadores de mínimos quadrados (EMQ) para os parâmetros de regressão e usar um estimador não viciado para a variância do erro baseado na soma de quadrados dos resíduos. Com a suposição de normalidade dos erros, os EMQ e EMV para os parâmetros de regressão coincidem, mas diferentes métodos de estimação da variância do erro podem levar à diferentes conclusões. Essa metodologia é praticamente padrão nas aplicações. Como alternativa, o uso do estimador de máxima verossimilhança (EMV) para a variância do erro (um estimador viciado, mas com erro quadrático menor), pode levar a diferentes conclusões em termos de inferências.

Em dois exemplos numéricos mostrou-se que o uso de diferentes métodos, muitas vezes pode mudar as conclusões obtidas por um experimento, dependendo do tamanho amostral. O tamanho amostral pequeno pode levar a erros padrões diferentes usando cada um dos dois métodos apresentados, o que pode acarretar diferentes conclusões inferenciais. Vale enfatizar, que entre os métodos de estimação, não há método que seja superior ao outro.

## Referências

- A.M MOOD, F.A. GRAYBILL, D.C. Boes, Introduction to the Theory of Statistics. Front Cover. McGraw-Hill, 1974.
- BICKEL P.J.; DOKSUM, K.A. Mathematical statistic: basic ideas and selected topics, vol 1, 2<sup>nd</sup> edn. Prentice Hall, Upper Saddle River, 2001
- BOLFARINE,H.; SANDOVAL,H.E. Introdução à Inferência Estatística, ed. Sociedade Brasileira de Matemática, 2001
- DRAPER, N.R.; SMITH, H. Applied regression analysis. Wiley series in probability and mathematical statistics, 1981.
- M.A.C. PANDOLFI. Implantação de estratégias tecnológicas e organizacionais em um ambiente em mudança:estudo de caso em uma empresa processadora de frutas tropicais. 2013. Dissertação (Mestrado em Engenharia de produção) - Centro Universitário de Araraquara,UNIARA,Araraquara,SP
- MOOD, F.A.; GRAYBILL, D.C. Introduction to the Theory of Statistics. McGraw-Hill, 1974.
- MONTGOMERY,D.C., PECK E.A. Introduction to Linear Regression Analysis, 2<sup>nd</sup> ed., John Wiley and Sons, Inc., New York, 1992.

MYERS,R.H. Classical and Modern Regression with Applications, 2th Ed., Duxbury Press, California USA, p.488,1990

NETER,J.; WASSERMAN, W.; KUTNER, M. H. Applied linear statistical models: regression, analysis of variance, and experimental designs, 2nd ed., Homewood, Ill. : R.D. Irwin,1985.

PANDOLFI, M.A.C. Implantação de estratégias tecnológicas e organizacionais em um ambiente em mudança: estudo de caso em uma empresa processadora de frutas tropicais, dissertação de mestrado profissional, Centro Universitário de Araraquara - UNIARA , 2013.

SEBER, G.A.F.; LEE, A.J. Linear regression analysis, second edition. Wiley series in probability and mathematical statistics, 2003.

WEISBERG,S Applied Linear Regression, 3ª Edição, John Wiley and Sons, Inc., New York, 2005