Lista 2 - MAE0328

Guilherme NºUSP: 8943160 e Leonardo NºUSP: 9793436

Exercício 5

Considere o modelo $Y = \beta_0 + \beta_1 + e_i$, em que $e_i \overset{iid}{\sim} Cauchy(0, 1),\ i \in \{1, ..., n\}$ i.e.

$$f(e_i) = \frac{1}{\pi(1 + e_i^2)} \mathbb{I}_R(e_i)$$

(a) Calcule o logaritmo da função de verossimilhança $L(\beta, y)$ e suas primeiras e segundas derivadas para o método iterativo de Newton-Raphson.

Resolução

Aplicando o teorema da transformação de variável: $Y = \beta_0 + \beta_1 x_i + e_i \Rightarrow Y^{-1} = y_i - (\beta_0 + \beta_1 x_i) \Rightarrow \frac{dY^{-1}}{dy_i} = 1$
Logo:

$$f(y_i - (\beta_0 - \beta_1 x_i)) = \frac{1}{\pi(1 + (y_i - (\beta_0 + \beta_1 x_i))^2)} \mathbb{I}_R(y_i)$$

Assim temos que se $e_i \overset{iid}{\sim} Cauchy(0, 1) \Rightarrow Y_i \sim Cauchy(\beta_0 + \beta_1 x_i, 1)$ Assim, sua verossimilhança é:

$$L(\beta, y) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \frac{1}{\pi(1 + (y_i - (\beta_0 + \beta_1 x_i))^2)} = (\frac{1}{\pi})^n \prod_{i=1}^{n} \frac{1}{1 + (y_i - \beta_0 - \beta_1 x_i)^2}$$

Aplicando o ln(logaritmo neperiano) em $L(\beta, y)$. Seja $l(\beta, y) = ln(L(\beta, y))$

$$l(\beta, y) = -n * ln(\pi) - \sum_{i=1}^{n} ln(1 + (y_i - \beta_0 - \beta_1 x_i)^2)$$

Calculando o vetor de suas primeiras derivadas:

$$\frac{\partial l(\beta, y)}{\partial \beta} = -n * ln(\pi) \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \sum_{i=1}^{n} \begin{pmatrix} \frac{-2(y_i - \beta_0 - \beta_1 x_i)}{1 + (y_i - \beta_0 - \beta_1 x_i)^2} \\ \frac{2x_i(y_i - \beta_0 - \beta_1 x_i)}{1 + (y_i - \beta_0 - \beta_1 x_i)^2} \end{pmatrix}$$

Agora, calculando o vetor de suas segundas derivadas:

$$\frac{\partial^2 l(\beta, y)}{\partial \beta^T \partial \beta} = -n * ln(\pi) \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} - \sum_{i=1}^{n} \begin{pmatrix} \frac{2((y_i - \beta_0 - \beta_1 x_i)^2 - 1)}{(1 + (y_i - \beta_0 - \beta_1 x_i)^2)^2} & \frac{2x_i((y_i - \beta_0 - \beta_1 x_i)^2 - 1)}{(1 + (y_i - \beta_0 - \beta_1 x_i)^2)^2} \\ \frac{2x_i((y_i - \beta_0 - \beta_1 x_i)^2 - 1)}{(1 + (y_i - \beta_0 - \beta_1 x_i)^2)^2} & \frac{2x_i^2((y_i - \beta_0 - \beta_1 x_i)^2 - 1)}{(1 + (y_i - \beta_0 - \beta_1 x_i)^2)^2} \end{pmatrix}$$

(b) Implemente no R o método de Newton-Raphson usando o exercício anterior. Considere como critério de parada do processo iterativo um erro máximo de 10^{-6}

Resolução

```r
# Newton-Raphson
x <- c(18.10,   10.71,  13.43,  11.04,  22.81,  18.71,  6.52 ,  15.78,
       19.59,   21.18,  15.82,  15.43,  20.61,  16.70,  18.52,  12.75,
       19.61,   14.80,  8.87 ,  21.82,  21.09,  26.47,  19.18,  23.31,
       15.9 ,   11.51,  7.09 ,  12.95,  21.99,  5.75 ,  15.88,  14.52,
       17.93,   5.19 ,  12.57,  19.78,  18.96,  14.24,  13.38,  19.38,
       2.47 ,   12.12,  5.46 ,  22.37,  22.51,  20.90,  9.92 ,  19.97,
       18.10,   20.50,  12.98,  21.18,  20.34,  30.71,  6.27 ,  15.47,
       18.08,   24.60,  7.26 ,  23.03,  23.67,  24.29,  11.70,  15.21,
       10.00,   10.93,  12.65,  15.70,  20.04,  21.97,  17.63,  10.72,
       19.71,   16.45,  15.04)

y <- c(23.38,   38.07,  29.79,  31.47,  18.44,  32.69,  36.14,  26.18,
      20.58,    18.69,  26.09,  9.27 ,  19.09,  24.20,  20.82,  29.33,
      19.94,    28.82,  35.16,  21.52,  17.11,  16.17,  22.28,  15.77,
      26.54,    33.40,  50.76,  30.41,  16.43,  35.83,  29.19,  28.16,
      23.34,    43.88,  31.64,  18.41,  21.59,  29.18,  31.34,  26.06,
      46.42,    32.52,  63.93,  15.96,  16.88,  19.55,  35.06,  18.90,
      24.13,    17.46,  30.12,  18.86,  9.10 ,  2.41 ,  39.80,  24.73,
      19.67,    13.06,  24.18,  14.01,  15.27,  13.23,  31.12,  27.41,
      70.36,    31.15,  32.10,  25.87,  20.60,  17.28,  24.25,  34.70,
      8.55 ,  24.48,    27.06)

MV = function(x,theta0,epsilon=10^(-6), it=1000){ # Dados, estimativa inicial, erro, iterações

  U <- function(t1,t2){
    U1 <- 2*sum((-t1-t2*x+y)/(((-t2*x-t1+y)^2)+1))
    U2 <- 2*sum((x*(-t2*x-t1+y))/(((-t2*x-t1+y)^2)+1))
    return (c(U1,U2))
  }

  H <- function(t1,t2){
    H11 <- 2*sum(((-t1-t2*x+y)^2-1)/((-t1-t2*x+y)^2+1)^2)
    H12 <- 2*sum(x*((-t1-t2*x+y)^2-1)/((-t1-t2*x+y)^2+1)^2)
    H22 <- 2*sum(x^2*((-t1-t2*x+y)^2-1)/((-t1-t2*x+y)^2+1)^2)
    M <- matrix(c(H11,H12,H12,H22),2,2)
    return (M)
  }

  erro <- 10
  j <- 0
  t1 <- numeric()
  t2 <- numeric()
  t1[1]<-theta0[1]
  t2[1]<-theta0[2]

  while(erro > epsilon & j < it){
    j <- j+1
```

```r
    Aux <- c(t1[j],t2[j]) - solve(H(t1[j],t2[j]))%*%U(t1[j],t2[j])
    t1[j+1] <- Aux[1]
    t2[j+1] <- Aux[2]
    erro <- max(abs(c(t1[j+1]-t1[j],t2[j+1]-t2[j])))
  }

  S <- list()
  S$erro <- erro
  S$Iteracoes <- j
  S$theta <- c(t1[j+1],t2[j+1])
  S$H <- H(t1[j+1],t2[j+1])
  S$U <- U(t1[j+1],t2[j+1])
  return(S)
}

#Encontrando as estimativas de MV para o exercício
theta0 <- c(52.45,-1.63) # Estimativa inicial
fit <- MV(x,theta0)

# Estimativas:
fit$theta
```

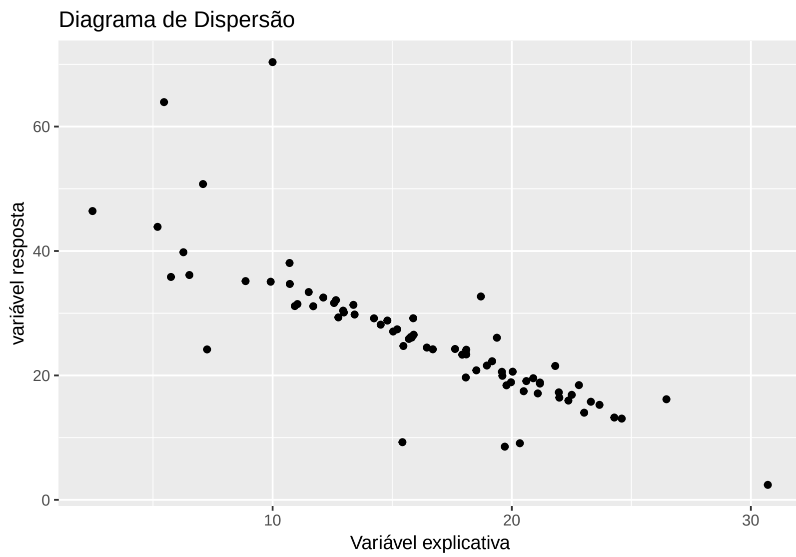
```
## [1] 49.96867 -1.49689
```

```r
# Verificando se os autovalores são negativos
eigen(fit$H) # Autovalores são estritamente negativos -> ponto de máximo
```

```
## eigen() decomposition
## $values
## [1]     -3.051852 -11971.936912
##
## $vectors
##             [,1]       [,2]
## [1,] -0.99848558 0.05501398
## [2,]  0.05501398 0.99848558
```

(c) Usando o conjunto de dados, faça um gráfico de dispersão e proponha estimativas inicias para β_0 e β_1. Encontre as estimativas de máxima verossimilhança do vetor de parâmetros β usando sua função do item (b).

Resolução

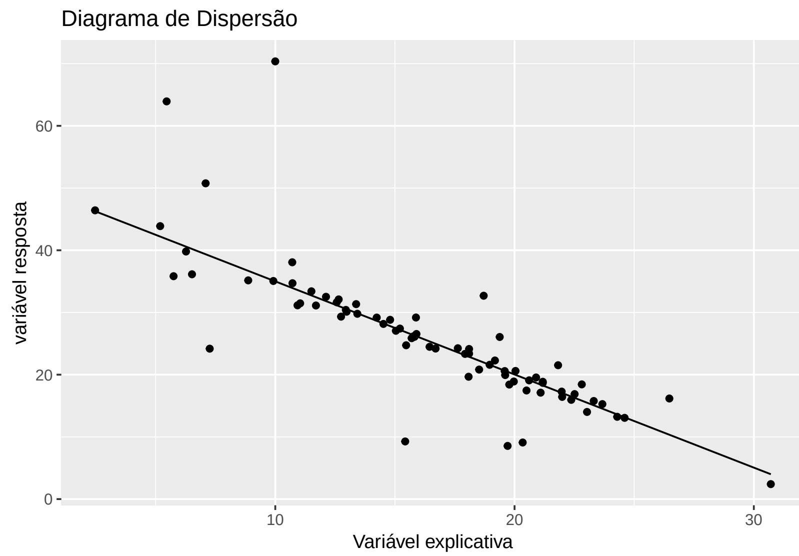


Como podemos notar uma tendencia linear no diagrama de disperção vamos propor as estimativas iniciais de mínimos quadrados onde $\widehat{\beta}_0^{MQ} = 52.454$ e $\widehat{\beta}_1^{MQ} = -1.629$. Assim pelo algoritmo do exercício anterior temos que as estimativas de máxima verossimilhança são $\widehat{\beta}_0^{MV} = 49.9688$ e $\widehat{\beta}_1^{MV} = -1.4969$

(d) Compare suas estimativas com a da função *optim()*

Resolução

No exercício (c) as estimativas de máxima verossimilhança encontradas $\widehat{\beta}_0^{MV} = 49.9688$ e $\widehat{\beta}_1^{MV} = -1.4969$ utilizando o algoritmo de Newton-Raphson com um erro de 10^{-6} ja as estimativas com a função *optim()* geraram as seguintes estimativas $\widehat{\beta}_0^{OP} = 49.9688$ e $\widehat{\beta}_1^{OP} = -1.4969$ onde foram utlizados a o $-ln(L(\beta, y))$ como função objetivo para máximização e os método BFGS, Nelder-Mead, L-BFGS-B e SANN notamos que as estimativas coicidiram com a de máxima verossimilhança, para verificação do ajuste, abaixo está o diagrama de disperção com a reta ajustada.



Exercício 6

O conjunto de dados contém os dados x : número de sinistros e y : pagamento total por todos os sinistros (em milhares de coroas suecas).

(a) Descreva o modelo linear simples que relacione o total pago, y_i, com número de sinistros, x_i. Apresente as suposições do modelo e interprete os parâmetros.

Resolução

O modelo linear simples utilizado é da forma $Y_i = \beta_0 + \beta_1 x_i + e_i$ supondo $\mathbb{E}(e_i) = 0$ e $Var(e_i) = \sigma^2$ e $e_1, .., e_n$ independentes.

Onde β_0 representa o valor esperado de Y_i (total pago de sinistros) quando $x_i = 0$(número de sinistros =0), assim como β_1 representa o valor de acréscimo ou decréscimo no valor esperado de Y_i(total pago de sinistros) quando x_i(número de sinistros) aumenta em uma unidade.

(b) Apresente as estimativas de mínimos quadrados de β_0 e β_1 e a estimativa não viciada para a variância σ^2

Resolução

Utilizando os estimadores de mínimos quadrados obtemos as seguintes estimativas:

$$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}} = \frac{Cov(x, y)}{Var(x)} = \frac{1861.604}{545.3134} = 3.4138$$

$$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1} * \bar{X} = 98.1873 - 3.4138 * 22.904 = 19.995$$

$$\widehat{\sigma}^2 = \frac{n}{n-2}\sigma^2_{MV} = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \widehat{\beta_0} - \widehat{\beta_1}x_i)^2 = \frac{78797.77}{63-2} = \frac{78797.77}{61} = 1291.7486$$

(c) Calcule a matriz de covariâncias estimada para $\widehat{\beta}$.

Resolução

A matriz de covariâncias pode ser calculada da seguinte forma:

$$Cov(\widehat{\beta}) = \mathbb{E}[(\widehat{\beta} - \mathbb{E}(\widehat{\beta}))(\widehat{\beta} - \mathbb{E}(\widehat{\beta}))^T] = Var(\widehat{\beta}) = Var((X^TX)^{-1}X^TY) = \sigma^2(X^TX)^{-1}$$

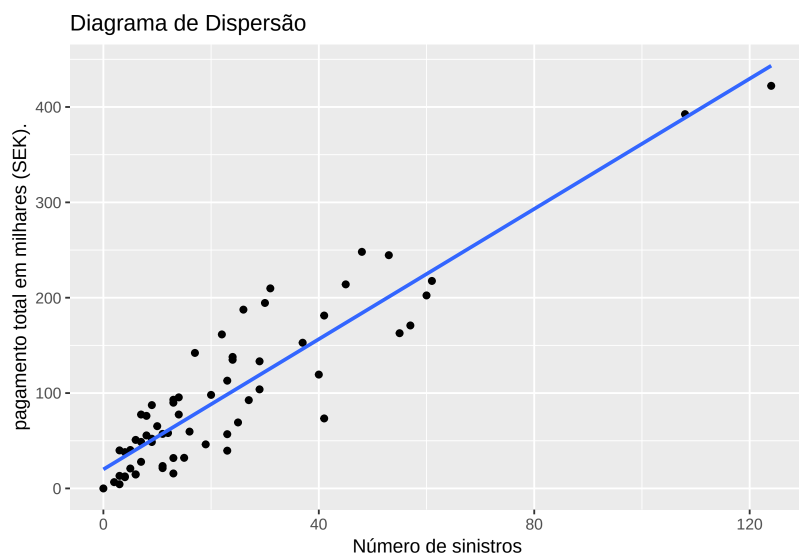
Sendo $\widehat{\sigma}^2 = 1291.74986$ e $(X^TX)^{-1} = \begin{pmatrix} 0.0314 & -0.0007 \\ -0.0007 & 0.00003 \end{pmatrix}$

Assim

$$Cov(\widehat{\beta}) = 1291.74986 \begin{pmatrix} 0.0314 & -0.0007 \\ -0.0007 & 0.00003 \end{pmatrix} = \begin{pmatrix} 40.524 & -0.875 \\ -0.875 & 0.382 \end{pmatrix}$$

(d) Faça o gráfico de dispersão e insira a reta estimada: $\widehat{\mu}_i = \widehat{\beta_0} + \widehat{\beta_1} x_i$ com $i = 1, ..., n$

Resolução



(d) Calcule o pagamento médio estimado para x = 100 sinistros.

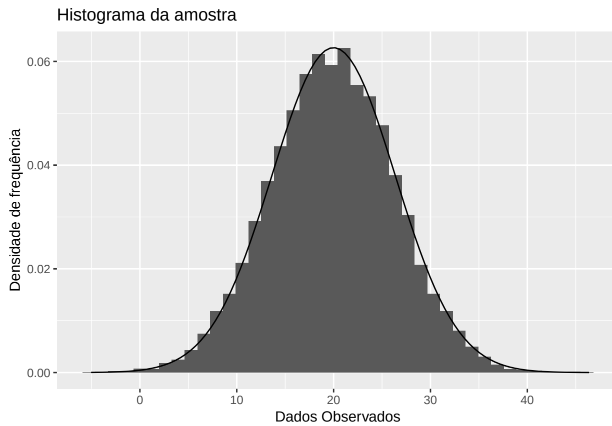
Resolução

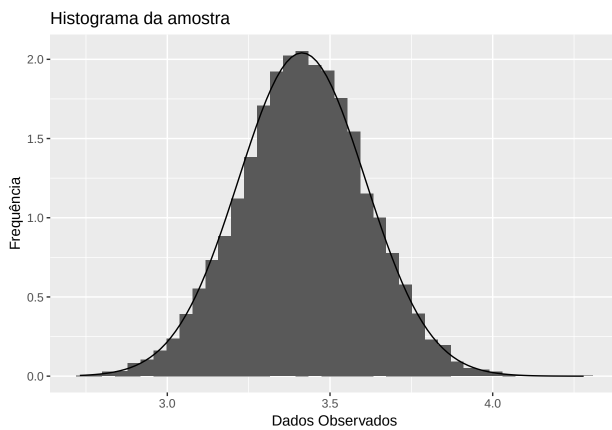
Como nossa equação do modelo é $\widehat{Y}_i = 19.99 + 3.41x_i \Rightarrow \mathbb{E}(Y_i|x_i = 100) = 19.995 + 3.4138 * 100 = 361.375$

Exercício 7

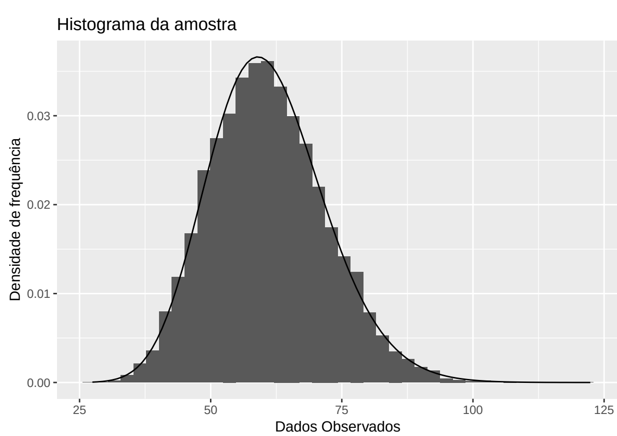
Considere $\widehat{\beta_0}$, $\widehat{\beta_1}$ e $\widehat{\sigma}^2$ obtidos no problema anterior.Simule $N = 1000$ conjuntos de dados, detamanho $n = 63$, segundo a equação $Y_i = \widehat{\beta_0} + \widehat{\beta_1} x_i + e_i$, em que x_i, $i = 1, ..., 63$, são dados no exercício anterior e os erros aleatórios devem ser gerados de acordo com $e_i \stackrel{iid}{\sim} N(0, \widehat{\sigma}^2)$ para $i = 1, ..., 63$. Para cada conjunto de dados simulado, estime novamente β_0, β_1 e σ^2 e armazene os resultados em uma matriz $N \times 3$. Finalmente, apresente os histogramas para os dados armazenados em cada uma das colunas dessa matriz e comente-os.

Resolução





Assim como mostra a teoria, o resultado simulado ilustra muito bem a normalidade dos dois estimadores seguindo $\beta_0 \sim N(\beta_0, \sigma^2(\frac{1}{n} + \frac{\bar{x}^2}{nS_{xx}}))$ e $\beta_1 \sim N(\beta_1, \frac{\sigma^2}{nS_{xx}})$. Podemos observar também que as estimativas de mínimos quadrados obtidos no exercício anterior são muito parecidos com as estimativas atuais, podemos verificar isso, pois as médias de $\widehat{\beta}_0$ e $\widehat{\beta}_1$ das 1000 estimativas simuladas são de 19.9169 e 3.414197 respectivamente, que coincidem com as estimativas encontradas no item b.



Por sua vez, $\widehat{\sigma}^2_{MV}$ é um estimador viciado, entretanto assintoticamente não viciado, com a correção $\widehat{\sigma}^2 = \frac{n}{n-2}\widehat{\sigma}^2_{MV}$ com essa correção temos que $\widehat{\sigma}^2 \sim \chi^2_{n-2}$ como $n = 63$ possui uma distribuição qui-quadrado com 61 graus de liberdade. Não é possível visualizar o valor estimado no exercício anterior pelo fato de que a distribuição Qui quadrado só depende de seus graus de liberdade.