Lista 7 - MAE0328

Guilherme Navarro NºUSP:8943160 Leonardo Noronha NºUSP:9793436

Para ambos os exercícios abaixo, use os dados da Tabela 1. Apresente todas as suas resoluções teóricas e implementações computacionais de forma clara, comentando seu código onde julgar apropriado.

Exercício 1

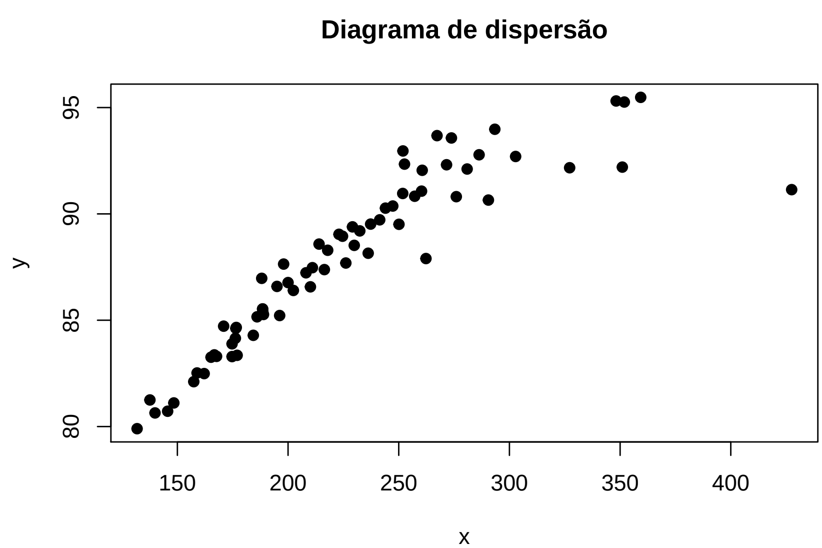
Considere o modelo não-linear de regressão dado por

$$y_i = \beta_0 x_i^{\beta_1} + e_i$$

em que $e_i \overset{iid}{\sim} N(0, \sigma^2),\ i \in \{1, ..., n\}$

(a) Apresente um gráfico de dispersão dos dados. Você julga adequada a pressuposição de um modelo linear homoscedástico para esses dados?

Resolução



(b) Apresente a esperança e variância de Y_i. A variância é homogênea? O modelo é linearizável?

Resolução

(c) Apresente a função de log-verossimilhança $l(\theta; y, X)$, em que $\theta = (\beta_0, \beta_1, \sigma^2)$, desse modelo e calcule a função escore e a matriz de informação de Fisher.

Resolução

(d) Com seus resultados do item (c), apresente uma implementação computacional do método escore de Fisher para obter a estimativa de máxima verossimilhança para θ. Usando sua implementação e considerando tolerância de erro de aproximação de 10^{-6}, obtenha as estimativas de máxima verossimilhança de β_0, β_1 e σ^2.

Resolução

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

  U <- function(b0,b1,s2){ # Função Escore
    U1 <- 1/s2*sum(x^b1*(y-b0*x^b1))
    U2 <- b0/s2*sum(x^b1*log(x)*(y-b0*x^b1))
    U3 <- -n/2*(1/s2)+sum(2/(2*s2)^2*((y-(b0*x^b1))^2))
    return (c(U1,U2,U3))
  }

  IF <- function(b0,b1,s2){ # Matriz informação de Fisher
    IF11 <- -(1/(2*s2)*sum((2*(x^b1*x^b1))))
    IF12 <- 1/(2*s2)*sum((2*(x^b1*log(x)*(y-(b0*x^b1))-x^b1*(b0*(x^b1*log(x))))))
    IF13 <- -(2/(2*s2)^2*sum((2*(x^b1*(y-(b0*x^b1))))))
    IF22 <- 1/(2*s2)*sum((2*(b0*(x^b1*log(x)*log(x))*(y-(b0*x^b1))-b0*(x^b1*log(x))*(b0*(x^b1*log(x))))
    IF23 <- -(2/(2*s2)^2*sum((2*(b0*(x^b1*log(x))*(y-(b0*x^b1))))))
    IF33 <- -(2*(2*(2*(2*s2)))/((2*s2)^2)^2*sum(((y-(b0*x^b1))^2)) -n/2*(1/s2^2))
    M <- matrix(c(IF11,IF12,IF13,IF12,IF22,IF23,IF13,IF23,IF33),3,3)
    return (M)
  }

  erro <- 10
  j <- 0
  b0 <- numeric()
  b1 <- numeric()
  s2 <- numeric()
  b0[1] <-theta0[1]
  b1[1] <-theta0[2]
  s2[1] <-theta0[3]

  while(erro > epsilon & j < it){
    j <- j+1
    Aux <- c(b0[j],b1[j],s2[j]) - solve(IF(b0[j],b1[j],s2[j]))%*%U(b0[j],b1[j],s2[j])
    b0[j+1] <- Aux[1]
    b1[j+1] <- Aux[2]
    s2[j+1] <- Aux[3]
    erro <- max(abs(c(b0[j+1]-b0[j],b1[j+1]-b1[j],s2[j+1]-s2[j])))
  }

  S <- list()
  S$erro <- erro
  S$Iteracoes <- j
  S$theta <- c(b0[j+1],b1[j+1],s2[j+1])
  S$IF <- IF(b0[j+1],b1[j+1],s2[j+1])
```

```
  S$U <- U(b0[j+1],b1[j+1],s2[j+1])
  return(S)
}
```

Oberservando o diagrama de disperção, tomamos as estimativas iniciais a apartir do ajuste via MMQ, as Estimativas Iniciais: $\widehat{\beta_0}^{(0)} = 1.63, \widehat{\beta_1}^{(0)} = 0.98, \widehat{\sigma^2}^{(0)} = 27.92$ Assim as estimativas de máxima vessimilhança utilizando o algoritmo de Escore-Fisher para β_0, β_1 e σ^2 são:

```
#Encontrando as estimativas de MV para theta
model00 <- lm(y~x)
theta0 <-  c(model00$coef,summary(model00)$sigma^2) # Estimativa inicial
fit <- EF(x,theta0)

# Estimativas:
fit$theta
```

(e) Usando a função *optim()* (ou alguma função de otimização numérica de sua preferência), estime novamente os parâmetros. Compare esses com os obtidos através do método escore de Fisher.

Resolução

Exercício 2

Considere agora uma modificação do modelo anterior, com

$$y_i = \beta_0 x_i^{\beta_1} + e_i$$

em que $e_i \overset{ind}{\sim} N(0, exp\{\gamma_0 + \gamma_1 x_i\}),\ i \in \{1, ..., n\}$.

(a) Entre essa e a especificação do exercício (1), qual você julga mais adequada? Justifique com base no gráfico de dispersão.

Resolução

(b) Apresente a função de log-verossimilhança desse modelo. Justifique o uso da função exponencial para definir a variância do modelo.

Resolução

(c) Encontre a estimativa de máxima verossimilhança para $\theta = (\beta_0, \beta_1, \gamma_0, \gamma_1)$ nesse novo modelo utilizando a função *optim()*. Compare-as com as obtidas no exercício (1).

Resolução

(d) Apresente um gráfico como os valores observados e as curvas de regressão estimadas dos dois modelos sugeridos.

Resolução