

## Simulação em um MLG Poisson

Será apresentado um estudo de simulação de algumas propriedades do estimador de máxima verossimilhança (EMV) aplicado a um modelo linear generalizado com resposta Poisson.

```
## Exemplo 6, p. 21
## Demétrio, C. G. B. (2002), "Modelos Lineares Generalizados em
## Experimentação Agronômica", ESALQ

y <- c(13, 14, 17, 22, 9, 14, 6, 14, 4, 4, 3, 5, 3, 2, 1, 3,
      2, 1, 3, 2, 2)
dil <- rep(c(0.3162, 0.1778, 0.1, 0.0562, 0.0316),
         times = c(4, 4, 4, 4, 5))

n <- length(y) # Tamanho da amostra

## Modelo Poisson com ligação canônica
m1 <- glm(y ~ dil, family = poisson)
summary(m1)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.7580      0.1818   4.170 3.05e-05 ***
dil           6.7848      0.7519   9.024 < 2e-16 ***
---
Residual deviance: 19.399 on 19 degrees of freedom
```

A estimativa do parâmetro de dispersão está próxima de 1, lembrando que no modelo Poisson temos  $\phi = 1$ . Outros aspectos da adequação do modelo não foram verificados.

No estudo de simulação serão avaliados o viés e o erro padrão dos EMV de  $\beta_1$  e  $\beta_2$ . Com base nas estimativas acima, selecionamos 0,8 e 6,8 como verdadeiros valores, respectivamente. Nas simulações os valores da covariável (dil) são mantidos fixos, sendo gerados  $M = 3000$  réplicas do vetor de respostas.

```
# Verdadeiros valores e preditor linear
nomes <- c("beta1", "beta2")
b10 <- 0.8
b20 <- 6.8
b0 <- c(b10, b20)
X <- model.matrix(m1)
eta0 <- exp(X %*% b0) # Preditor linear
```

No bloco abaixo são geradas as réplicas da variável resposta. Para cada vetor de réplicas (ysim) obtemos estimativas dos parâmetros (b1 e b2) e também estimativas dos erros padrão (epb1 e epb2).

Nota 1. Procure entender todas as linhas do bloco abaixo.

```

# Dados simulados e estimação
b1 <- b2 <- epb1 <- epb2 <- c()
for (m in 1:M) {
  ysim <- rpois(n, lambda = eta0)
  msim <- summary(glm(ysim ~ dil, family = poisson))
  b1[m] <- msim$coefficients[, "Estimate"][[1]]
  b2[m] <- msim$coefficients[, "Estimate"][[2]]
  epb1[m] <- msim$coefficients[, "Std. Error"][[1]]
  epb2[m] <- msim$coefficients[, "Std. Error"][[2]]
}

```

Em seguida podemos observar os verdadeiros valores dos parâmetros e as médias das EMV.

```

# Verdadeiros valores e médias das estimativas
res1 <- cbind(b0, c(mean(b1), mean(b2)))
rownames(res1) <- nomes
colnames(res1) <- c("Verdadeiro", "Estimativa")
cat("\n Tamanho da amostra:", n)
print(res1, digits = 3)

```

```

Tamanho da amostra: 21
      Verdadeiro Estimativa
beta1      0.8      0.79
beta2      6.8      6.81

```

Nota 2. Comente os resultados acima.

A partir das  $M$  ( $= 3000$ , neste caso) estimativas de  $\beta_1$  e  $\beta_2$  podemos obter o desvio padrão das estimativas (chamado de empírico), lembrando que o desvio padrão amostral é um estimador consistente do desvio padrão populacional (finito e maior do que 0). Comparamos com os erros padrão assintóticos baseados obtidos da teoria de máxima verossimilhança.

```

# Desvios padrão e médias dos erros padrão
res2 <- cbind(c(sd(b1), sd(b2)), c(mean(epb1), mean(epb2)))
rownames(res2) <- nomes
colnames(res2) <- c("Desvio padrão", "Erro padrão")
print(res2, digits = 3)

```

```

      Desvio padrão Erro padrão
beta1      0.179      0.179
beta2      0.735      0.740

```

Nota 3. Comente os resultados acima.

Nota 4. Como foram obtidos os erros padrão assintóticos acima?

Nota 5. Como realizar estudos de simulação com outros tamanhos amostrais?

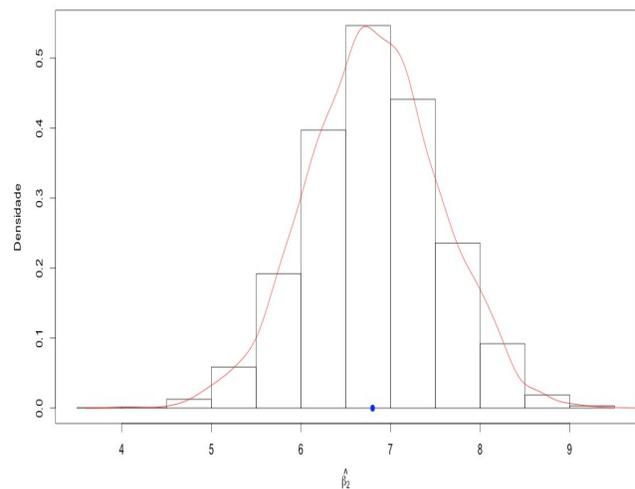
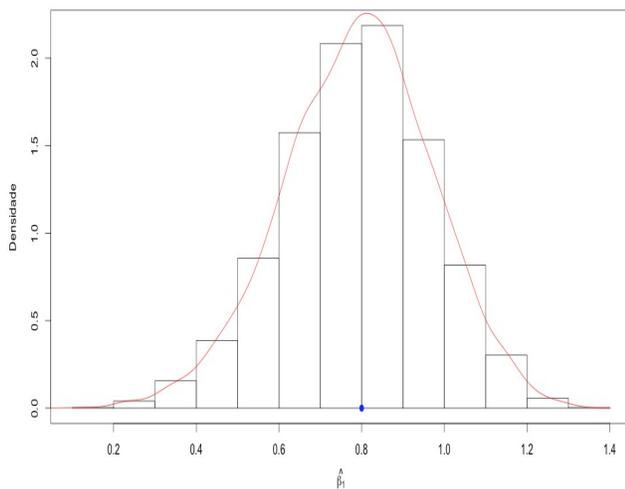
Nota 6. Procure desenvolver um estudo de simulação com os testes da razão de verossimilhanças, Wald e escore para testar  $H_0: \beta_1 = \beta_{10}$  contra  $H_1: \beta_1 \neq \beta_{10}$ .

```

# Gráficos
hist(b1, main = "", freq = FALSE, ylab = "Densidade", xlab =
expression(hat(beta[1])))
lines(density(b1), col = "red")
points(b10, 0, pch = 19, col = "blue")
box()

hist(b2, main = "", freq = FALSE, ylab = "Densidade", xlab =
expression(hat(beta[2])))
lines(density(b2), col = "red")
points(b20, 0, pch = 19, col = "blue")
box()

```



Nota 7. Obtenha as distribuições assintóticas dos EMV de  $\beta_1$  e  $\beta_2$  e superponha seus gráficos às figuras acima. Surpresa?

Nota 8. Procure desenvolver estudos de simulação envolvendo métodos de estimação e de testes de hipóteses vistos em outras disciplinas.