Prova Inferência

Ivan Cunha Vieira

2025-01-29

Questão 2.

1.

Espaço paramétrico: $\Theta = (0, \infty)$.

Suporte de X: $A(x) = (0, \theta)$.

2.

Para mostrar que $\hat{\theta}_1$ é não viciado para θ, devemos utilizar a natureza da distribuição de X, tal que $X \sim U(0, \theta)$ e, portanto, $E[X] = \theta$ e $Var[X] = \frac{\theta^2}{12}$. Seguindo para $\hat{\theta}_1$, temos:

$\hat{\theta}_1 = 2\bar{X} = \sum_{i=1}^{n} \frac{2X_i}{n} \to E[\hat{\theta}_1] = \frac{2}{n} E[\sum_{i=1}^{n} X_i] = \frac{2n}{n} E[X] = E[X] = \theta$.

Assim, $\hat{\theta}_1$ é não viciado, pois: $B(\hat{\theta}_1) = E[\hat{\theta}_1] - \theta = 0$

Para mostrar que $\hat{\theta}_2$ é não viciado para θ, devemos utilizar de probabilidade, tal que $P(X_{(n)} \leq x) = P(X_1 \leq x, ..., X_n \leq x) = P(X \leq x)^n$. Como $X \sim U(0, \theta) \to P(X \leq x)^n = (\dfrac{x}{\theta})^n$.

Dessa forma, encontramos a função de distribuição acumulada, que nos leva à função de densidade de probabilidade, tal que f.d.p. $= \dfrac{nx^{n-1}}{\theta^n}$. Assim, $E[X_{(n)}] = \frac{n}{\theta^n} \int_0^\theta x.x^{n-1} dx = \dfrac{n\theta^{n+1}}{\theta^n(n+1)} = \dfrac{n\theta}{n+1}$.

Agora, $\hat{\theta}_2 = \frac{n+1}{n} X_{(n)} \to E[\hat{\theta}_2] = \frac{n+1}{n} E[X_{(n)}] = \theta$.

Com isso, encontramos que $\hat{\theta}_2$ é não viciado, pois: $B(\hat{\theta}_2) = E[\hat{\theta}_2] - \theta = 0$

3.

Calculando os EQMs de $\hat{\theta}_1$ e $\hat{\theta}_2$:

$Var(\hat{\theta}_1) = Var(\frac{2\sum_{i=1}^{n} Xi}{n}) = \frac{4n}{n^2} Var(X) = \dfrac{4\theta^2}{12n} = \dfrac{\theta^2}{3n}$.

Como $\hat{\theta}_1$ é não viciado, $EQM(\hat{\theta}_1) = Var(\hat{\theta}_1) = \dfrac{\theta^2}{3n}$.

$Var(\hat{\theta}_2) = \dfrac{(n+1)^2}{n^2} Var(X_{(n)})$ $E[\hat{\theta}_2^2] = \frac{n}{\theta^n} \int_0^\theta x^2.x^{n-1} dx = \dfrac{n\theta^{n+2}}{\theta^n(n+2)} = \dfrac{n\theta^2}{n+2}$.

Agora, $Var(X_{(n)}) = \dfrac{n\theta^2}{n+2} - (\dfrac{n\theta}{n+1})^2 = \dfrac{n\theta^2[(n+1)^2 - n(n+2)]}{(n+2)(n+1)^2} = \dfrac{n\theta^2}{(n+2)(n+1)^2}$.

Dessa forma, $Var(\hat{\theta}_2) = \dfrac{(n+1)^2}{n^2} Var(X_{(n)}) = \dfrac{\theta^2}{n(n+2)}$.

Como $\hat{\theta}_2$ é não viciado, $EQM(\hat{\theta}_2) = Var(\hat{\theta}_2) = \dfrac{\theta^2}{n(n+2)}$.

E, portanto, encontramos que para $n > 1 \to EQM(\hat{\theta}_2) < EQM(\hat{\theta}_1)$, o que faz de $\hat{\theta}_2$ mais eficiente.

4.

Um estimador consistente possui as seguintes característica: $lim_{n\to\infty} E[\theta] = \theta$ e $lim_{n\to\infty} Var[\theta] = 0$. Para os estimadores do problema, temos que:

$lim_{n\to\infty} E[\hat{\theta}_1] = lim_{n\to\infty}\, \theta = \theta$ e $lim_{n\to\infty} Var[\hat{\theta}_1] = lim_{n\to\infty} \dfrac{\theta^2}{3n} = 0$.

$lim_{n\to\infty} E[\hat{\theta}_2] = lim_{n\to\infty}\, \theta = \theta$ e $lim_{n\to\infty} Var[\hat{\theta}_2] = lim_{n\to\infty} \dfrac{\theta^2}{n+2} = 0$.

Dessa forma, observamos que ambos os estimadores são consistentes, pois atendem às características estipuladas.

Questão 5.

1.

$X_i \sim Ber(\theta) \to E[X_i] = \theta\,; Var(X_i) = \theta(1-\theta)$. Para $\hat{\theta}_n$, temos que $E[\hat{\theta}_n] = E[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n}\sum_1^n E[X] = \frac{n}{n}\theta = \theta$.

Com isso, encontramos que $\hat{\theta}_n$ é não viciado, pois: $B(\hat{\theta}_n) = E[\hat{\theta}_n] - \theta = 0$.

2.

$lim_{n\to\infty} E[\hat{\theta}_n] = lim_{n\to\infty}\, \theta = \theta$.

$Var(\hat{\theta}_n) = Var(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{n}{n^2} Var(X) = \dfrac{\theta(1-\theta)}{n} \to lim_{n\to\infty} Var[\hat{\theta}_n] = lim_{n\to\infty} \dfrac{\theta(1-\theta)}{n} = 0$.

Portanto, observamos que o estimador $\hat{\theta}_n$ é consistente, pois atende às características estipuladas.

3.

Sendo X_n uma sequência de v.a's i.i.d. com distribuição $Ber(\theta)$ e $S_n = \sum_1^n X_i$, de forma que $E[S_n] = n\theta$ e $Var(S_n) = n\theta(1-\theta)$ temos que, pelo teorema de Moivre-Laplace, $\dfrac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}} \sim N(0,1)$, ou, paralelamente, $S_n \sim N(n\theta, n\theta(1-\theta))$.

Agora, para $\hat{\theta}_n$, com parâmetros $E[\hat{\theta}_n] = \theta$ e $Var(\hat{\theta}_n) = \dfrac{\theta(1-\theta)}{n}$, encontramos o seguinte:

$$\frac{\hat{\theta}_n - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} = \frac{\frac{1}{n}\sum_{i=1}^n X_i - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} = \frac{\frac{1}{n}S_n - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \to \frac{\frac{1}{n}S_n - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}}.\frac{n}{n} = \frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}} \sim N(0,1).$$

Isso nos leva à $\dfrac{\hat{\theta}_n - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \sim N(0,1) \to \hat{\theta}_n \sim N(\theta, \frac{\theta(1-\theta)}{n})$

4.

Input:

- Probabilidade de sucesso: $\theta = 0.5$;
- Número de amostras: $num\_amostras = 10000$;
- Vetor para conter as médias obtidas: seq_medias.

Código:

- Gerar amostra de números aleatórios com distribuição Bernoulli($\theta = 0.5$);
- Armazenar cada uma das médias acumuladas dos valores gerados de 1 à $num\_amostras$;
- Plotar as médias amostrais obtidas em função do tamanho da amostra.

Output:

- Gráfico apresentando as médias obtidas (tanto as acumuladas quanto a teórica);
- Valor esperado e valor simulado da média acumulada.

5.

Input:

- Probabilidade de sucesso: $\theta = 0.5$;
- Número de amostras: $num\_amostras = 10000$;
- Número de simulações a gerar: $simulacoes = 10000$.

Código:

- Obter a média e desvio padrão teóricos para $\hat{\theta}_n \sim N(\theta, \frac{\theta(1-\theta)}{n})$;
- Gerar amostras de números aleatórios com distribuição Bernoulli($\theta = 0.5$);
- Armazenar as médias e desvios obtidos das simulações realizadas para $\theta_n \sim N(\theta, \frac{\theta(1-\theta)}{n})$;
- Plotar as médias amostrais obtidas.

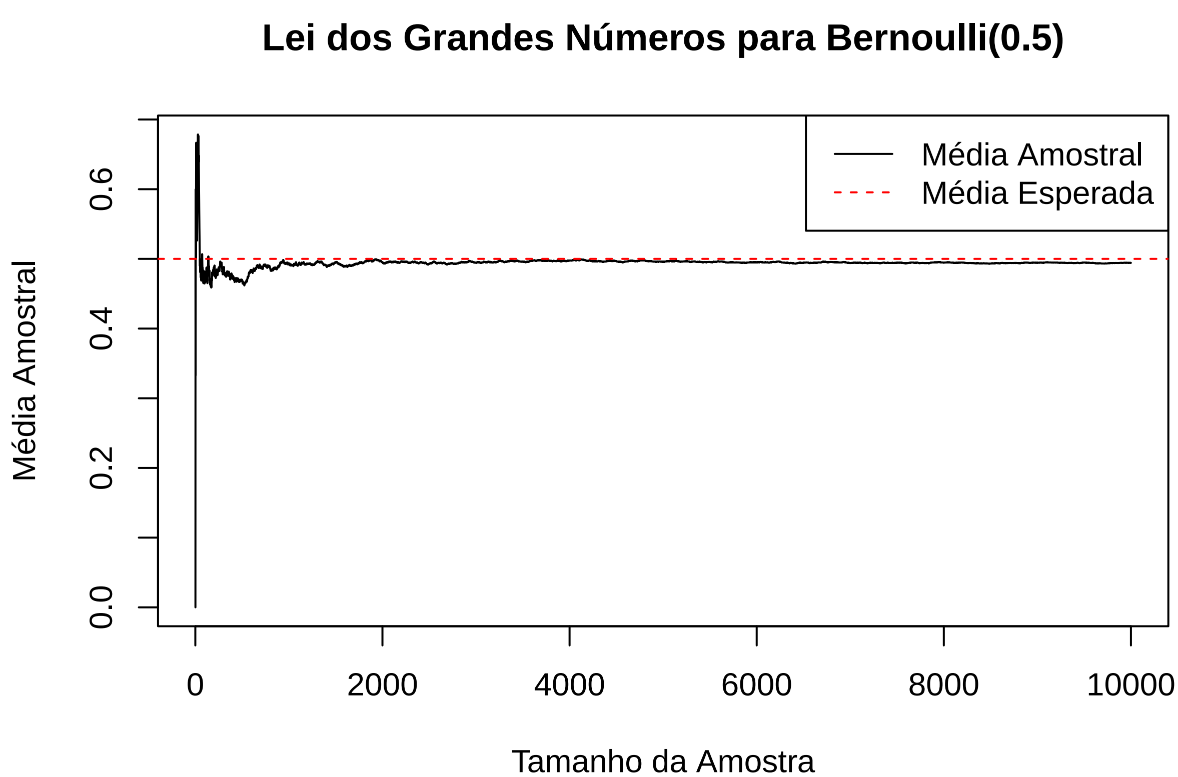
Output:

- Histograma apresentando as médias obtidas (tanto as simuladas quanto a teórica);
- Valores esperados e simulados da média e desvio padrão para $\hat{\theta}_n \sim N(\theta, \frac{\theta(1-\theta)}{n})$.

6.

```r
# Parâmetros
theta <- 0.5

# Número de amostras a gerar
num_amostras <- 10000

# Vetor que irá conter as médias obtidas
```

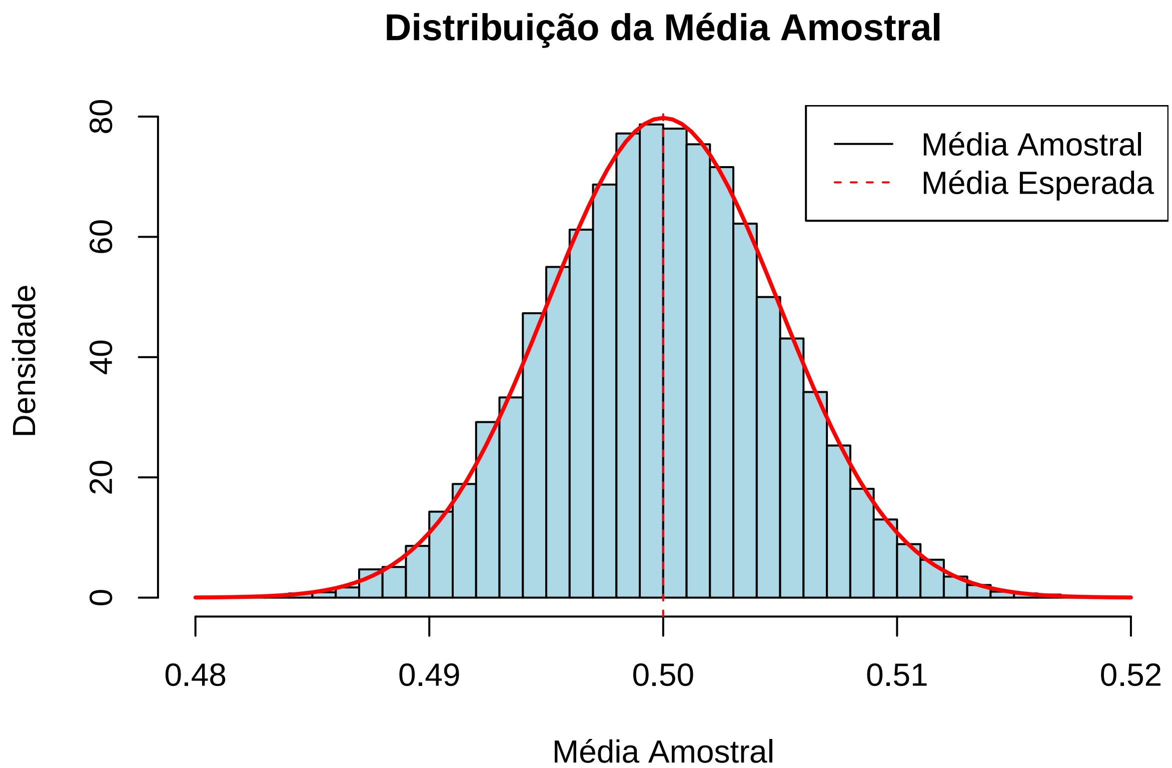
```r
seq_medias <- NULL

# Armazenar os números aleatórios gerados
set.seed(123)
y <- rbinom(num_amostras, size = 1, prob = theta)

for (i in 1:num_amostras){
seq_medias[i] <- mean(y[1:i])
}

# Plotar as médias amostrais em função do tamanho da amostra
plot(1:num_amostras, seq_medias, type = "l",
main = "Lei dos Grandes Números para Bernoulli(0.5)",
xlab = "Tamanho da Amostra", ylab = "Média Amostral")

abline(h = theta, col = "red", lty = 2)

legend("topright", legend = c("Média Amostral", "Média Esperada"),
col = c("black", "red"), lty = c(1, 2))
```




```
## [1] "Média esperada = 0.5"

## [1] "Média obtida = 0.4943"
```

7.

```
# Parâmetros
theta <- 0.5

# Número de amostras a gerar
num_amostras <- 10000

# Número de simulações
simulacoes <- 10000

# Valores teóricos para a Normal
media_esperada <- theta
desvio_esperado <- sqrt(theta * (1 - theta) / num_amostras)

tcl <- function(num_amostras, theta, simulacoes) {
set.seed(123)
seq_medias <- NULL
for (i in 1:simulacoes) {
amostra <- rbinom(num_amostras, size = 1, prob = theta)
seq_medias[i] <- mean(amostra)}
return(seq_medias)}
medias_simuladas <- tcl(num_amostras, theta, simulacoes)

hist(medias_simuladas, breaks = 30, probability = TRUE,
main = "Distribuição da Média Amostral",
xlab = "Média Amostral", col = "lightblue", border = "black",
ylab = "Densidade")

curve(dnorm(x, mean = media_esperada, sd = desvio_esperado),
col = "red", lwd = 2, add = TRUE)

abline(v = theta, col = "red", lty = 2)

legend("topright", legend = c("Média Amostral", "Média Esperada"),
col = c("black", "red"), lty = c(1, 2))
```




```
## [1] "Média esperada = 0.5"

## [1] "Média obtida = 0.49998524"

## [1] "Desvio esperado = 0.005"

## [1] "Desvio obtido = 0.00500558625255556"
```