



UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE CIÊNCIAS  
BACHARELADO EM ESTATÍSTICA

ANTÔNIO ARTHUR SILVA DE LIMA

RELATÓRIO  
INFERÊNCIA ESTATÍSTICA I

FORTALEZA  
2023

ANTÔNIO ARTHUR SILVA DE LIMA

## RELATÓRIO DE INFERÊNCIA ESTATÍSTICA I

Relatório apresentado ao curso de Bacharelado em Estatística, do Centro de Ciências da Universidade Federal do Ceará, como parte integrante da nota na disciplina de Inferência Estatística I, do semestre 2023.1.

Orientador: Prof. Dr. João Maurício Araújo Mota

FORTALEZA

2023

# Sumário

<b>1</b>	<b>Introdução</b>	<b>6</b>
<b>2</b>	<b>Metodologia</b>	<b>6</b>
<b>3</b>	<b>Questões</b>	<b>7</b>
<b>4</b>	<b>Conclusões</b>	<b>13</b>
<b>5</b>	<b>Códigos</b>	<b>14</b>

## Lista de Figuras

1	Gráfico 1a), com área da probabilidade hachurada. . . . .	7
2	Gráfico 1b), com área da probabilidade hachurada. . . . .	7
3	Histograma de uma amostra de tamanho $n = 50$ , questão 1. . . . .	7
4	Gráfico 2a), com área da probabilidade hachurada. . . . .	8
5	Gráfico 2b), com área da probabilidade hachurada. . . . .	8
6	Histograma de uma amostra de tamanho $n = 50$ , questão 2. . . . .	8
7	Gráfico 3a), com área da probabilidade hachurada. . . . .	9
8	Gráfico 3b), com área da probabilidade hachurada. . . . .	9
9	Histograma de uma amostra de tamanho $n = 50$ , questão 3 . . . . .	10
10	Gráfico 4a), com área da probabilidade preenchida. . . . .	11
11	Gráfico 4a), da normal padrão, com área da probabilidade preenchida. . . .	11
12	Gráfico 4b), com área da probabilidade preenchida. . . . .	11
13	Gráfico 4b), da normal padrão, com área da probabilidade preenchida. . . .	11
14	Histograma de uma amostra de tamanho $n = 50$ , questão 4 . . . . .	12

## Lista de Tabelas

1	Comparação entre dados populacionais e amostrais - questão 1. . . . .	8
2	Comparação entre dados populacionais e amostrais - questão 2. . . . .	9
3	Comparação entre dados populacionais e amostrais - questão 3. . . . .	10
4	Comparação entre dados populacionais e amostrais - questão 4. . . . .	12

# 1 Introdução

Com o intuito de aprimorar o uso do R junto às aulas de Inferência e suas aplicações, o presente relatório, proposto pelo Professor Dr. João Maurício, busca responder algumas questões — as quais compunham a segunda chamada da 6ª verificação de aprendizagem — envolvendo o uso de tabelas e suas distribuições, por meio do software estatístico R.

Aqui, os gráficos foram introduzidos diretamente como respostas das questões, a fim de ilustrar os resultados decorrentes das probabilidades calculadas, e, em uma seção dedicada ao final do relatório, são apresentados os códigos feitos para a resolução de cada questão.

Como material de apoio, foram utilizadas algumas notas de aulas disponibilizadas pelo Professor, bem como apostilas voltadas para o aprendizado da linguagem R.

# 2 Metodologia

Para a realização deste relatório foi utilizado o software R e suas funções nativas para criar gráficos. Além disso, a biblioteca auxiliar *latex2exp* também foi utilizada, a fim de estilizar os títulos e as legendas dos gráficos com o intuito de formatar as expressões estatísticas, assim como no formato dos documentos L<sup>A</sup>T<sub>E</sub>X.

Para cada questão, foram calculadas as probabilidades pedidas, a fim de verificar os pontos corretos a serem plotados. Para as amostradas aleatórias geradas, foi feita uma breve descrição das mesmas, obtendo valores como quartis, média, e valores de máximo e mínimo, além da sua variância, os quais foram encontrados por meio dos comandos apresentados na seção **Códigos**, de acordo com cada questão.

### 3 Questões

1. Seja  $X \sim \chi^2(15)$ .

a) Calcule  $p_a = P(X \leq 11,721)$ .

b) Qual a mediana de  $X$ ?

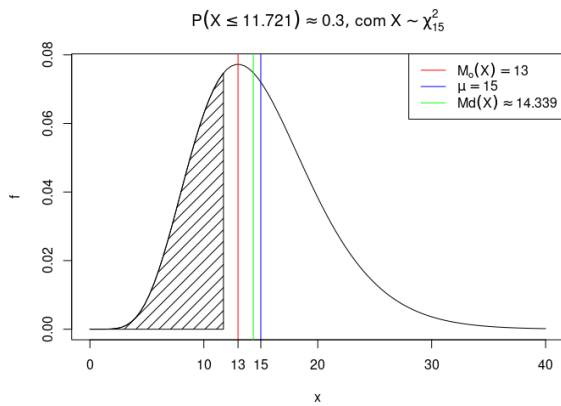


Figura 1: Gráfico 1a), com área da probabilidade hachurada.

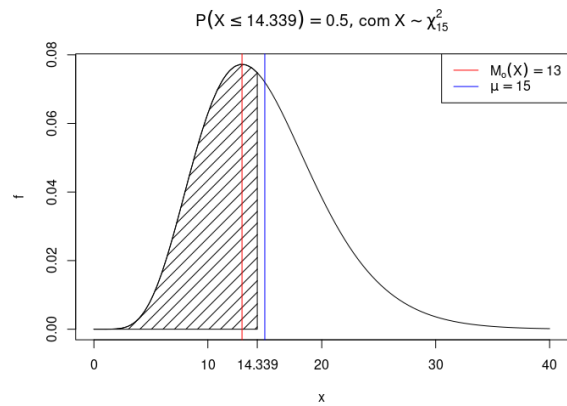


Figura 2: Gráfico 1b), com área da probabilidade hachurada.

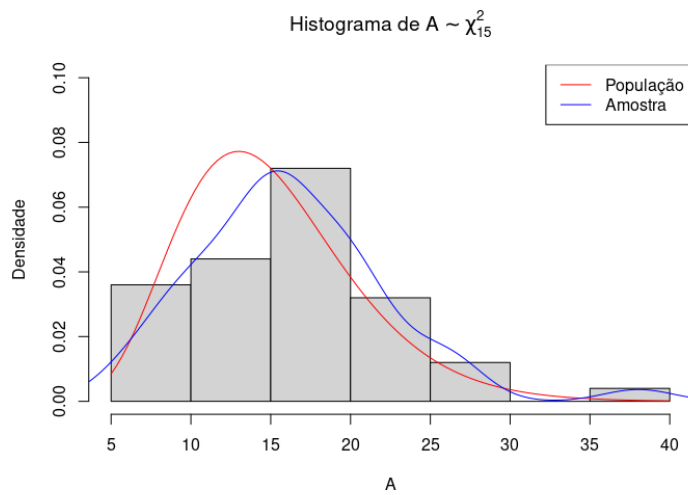


Figura 3: Histograma de uma amostra de tamanho  $n = 50$ , questão 1.

Sabemos que de uma  $\chi^2_{15}$ ,  $\mu = 15$ ,  $M_o(X) = 15 - 2 = 13$ , e vemos com o uso do R que  $Med(X) \approx 14,339$ . Já na amostra aleatória gerada, temos que  $\bar{x} = 16,235$ ,  $Med(A) = 15,598$  e a moda  $M_o(A) \approx 15$ . De maneira mais clara, a Tabela 1 mostra tal comparação.

Comparação de medidas resumo		
	População	Amostra
Média	15	16,235
Moda	13	$\approx 15$
Mediana	14,339	15,598

Tabela 1: Comparação entre dados populacionais e amostrais - questão 1.

2. Seja  $X \sim t(21)$ .

a) Calcule  $p_a = P(X \leq 0,859)$

b) Qual o sexto decil de  $X$ ?

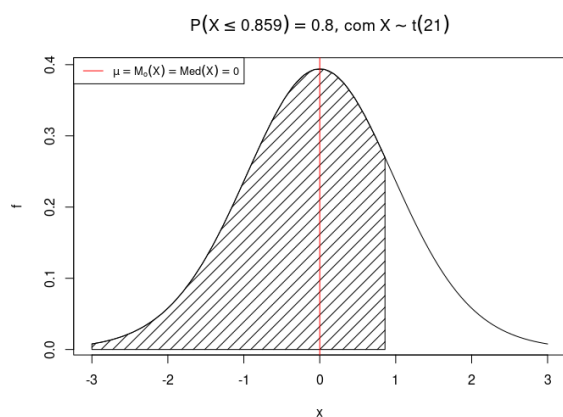


Figura 4: Gráfico 2a), com área da probabilidade hachurada.

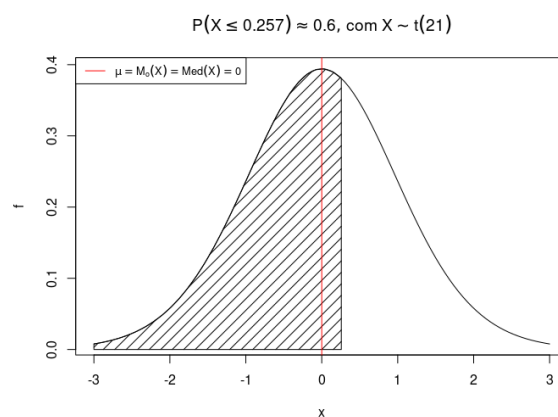


Figura 5: Gráfico 2b), com área da probabilidade hachurada.

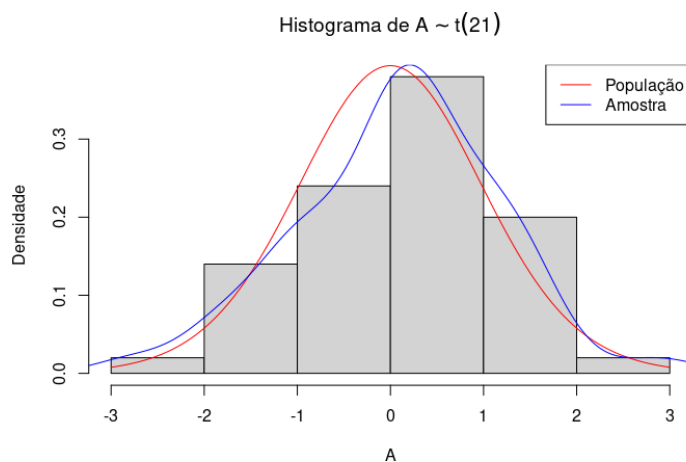


Figura 6: Histograma de uma amostra de tamanho  $n = 50$ , questão 2.



Sabemos que de uma  $t(21)$ ,  $\mu = M_o(X) = Med(X) = 0$ . Podemos ver usando o R, que da amostra gerada,  $\bar{x} = 0,0817$ ,  $Med(A) = 0,162$ , e que  $M_o(A) \approx 0,2$ . De maneira mais clara, a Tabela 2 mostra tal comparação.

Comparação de medidas resumo		
	População	Amostra
Média	0	0,0817
Moda	0	$\approx 0,2$
Mediana	0	0,162

Tabela 2: Comparação entre dados populacionais e amostrais - questão 2.

3. Seja  $X \sim F(15, 15)$

- Calcule  $p_a = P(X \leq 2,40)$
- Qual o quinto percentil de  $X$ ?

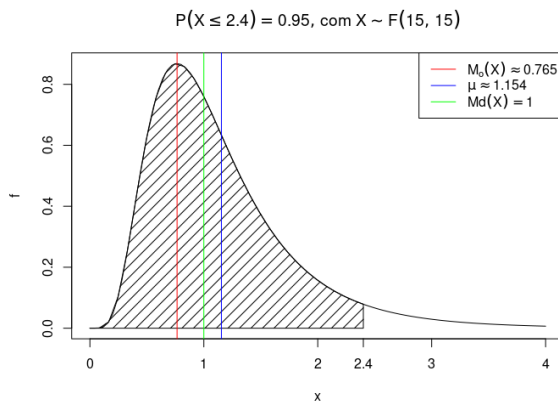


Figura 7: Gráfico 3a), com área da probabilidade hachurada.

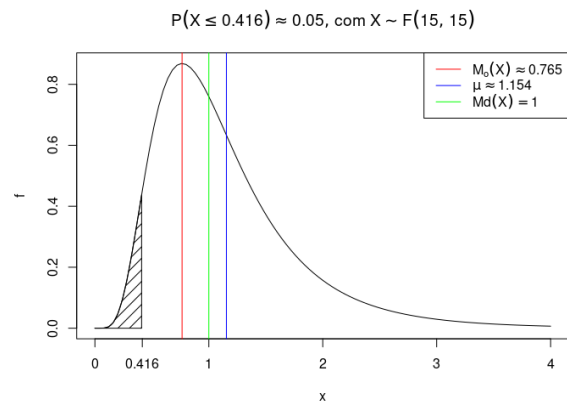


Figura 8: Gráfico 3b), com área da probabilidade hachurada.

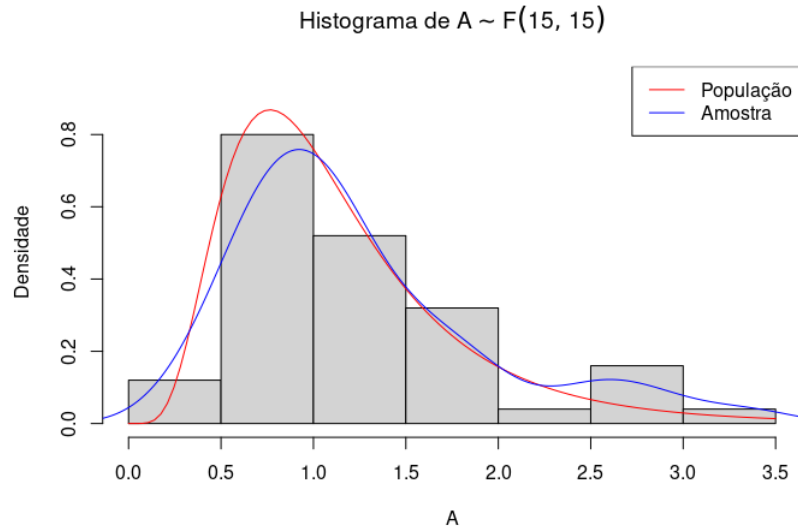


Figura 9: Histograma de uma amostra de tamanho  $n = 50$ , questão 3

Sabemos que de uma  $F(15, 15)$ ,  $\mu = \frac{15}{13}$ ,  $M_o(X) = \frac{13}{15} \times \frac{15}{17}$ , e pelo R,  $Med(X) = 1$ . Já nossa amostra, possui  $\bar{x} = 1,247$ ,  $M_o(X) \approx 0,9$ , e  $Med(X) = 1,07$ . De maneira mais clara, a Tabela 3 mostra tal comparação.

Comparação de medidas resumo		
	População	Amostra
Média	1,154	1,247
Moda	0,765	$\approx 0,9$
Mediana	1	1,07

Tabela 3: Comparação entre dados populacionais e amostrais - questão 3.

4. Seja  $X \sim N(40, 25)$

- Calcule  $p_a = P(|X - 40| < 12)$
- Qual o percentil de ordem 5 de  $X$ ?

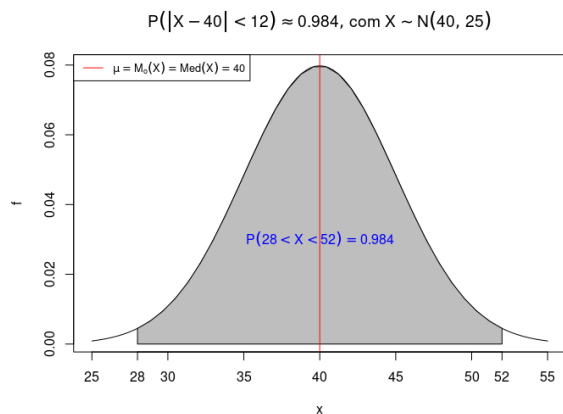


Figura 10: Gráfico 4a), com área da probabilidade preenchida.

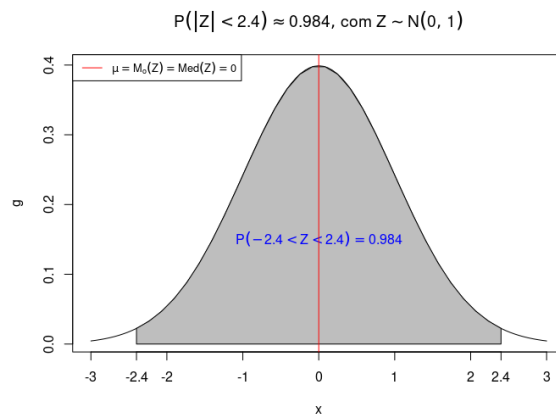


Figura 11: Gráfico 4a), da normal padrão, com área da probabilidade preenchida.

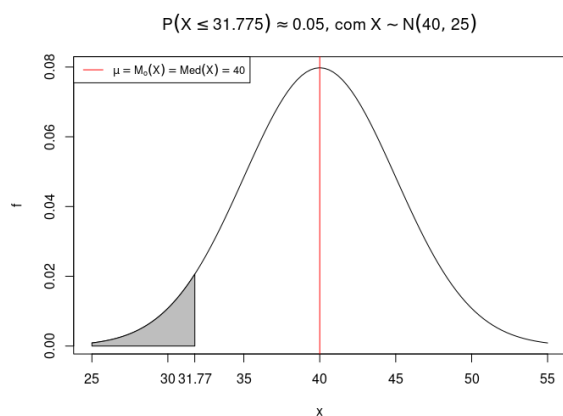


Figura 12: Gráfico 4b), com área da probabilidade preenchida.

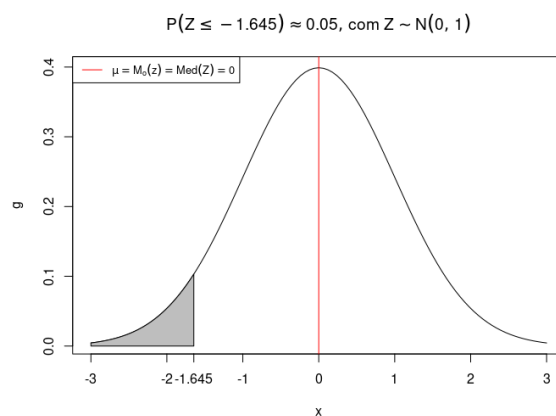


Figura 13: Gráfico 4b), da normal padrão, com área da probabilidade preenchida.

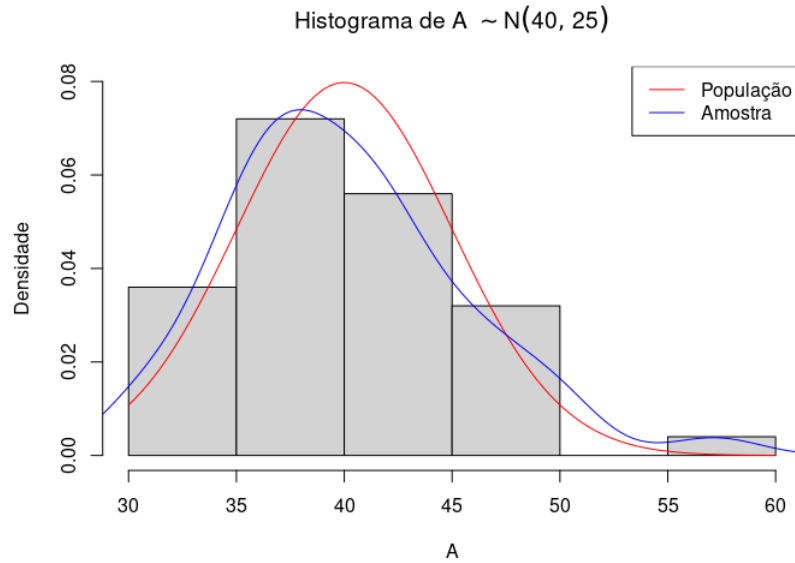


Figura 14: Histograma de uma amostra de tamanho  $n = 50$ , questão 4

Sabemos que de uma  $N(40, 25)$ ,  $\mu = M_o(X) = Med(X) = 40$ . Assim como na normal padrão,  $\mu = M_o(Z) = Med(Z) = 0$ . Na nossa amostra, encontramos  $\bar{x} = 39,97$ ,  $M_o(X) \approx 38$  e  $Med(X) = 39,28$ . De maneira mais clara, a Tabela 4 mostra tal comparação.

Comparação de medidas resumo		
	População	Amostra
Média	40	39,97
Moda	40	$\approx 38$
Mediana	40	39,28

Tabela 4: Comparação entre dados populacionais e amostrais - questão 4.

## 4 Conclusões

De maneira geral, pôde-se observar que as amostras aleatórias, que foram brevemente descritas com o uso do comando `summary()` nativo do R, apresentaram medidas muito próximas às das suas respectivas distribuições populacionais, sendo a mais “discrepante” dentre elas a amostra gerada para a população  $\chi^2_{15}$ , na 1ª questão.

Também vemos que as distribuições amostrais se aproximam bastante da distribuição real populacional de cada uma, o que tem por consequência as semelhanças entre as medidas descritivas como Média e Mediana, apresentadas nos gráficos e nas tabelas comparativas.

Finalmente, o relatório apresentou muitos desafios, que culminaram em uma experiência frutífera no que tange tanto ao aprendizado do software R, bem como na fixação do conteúdo da disciplina, pois trabalhar com a Inferência em uma abordagem um pouco mais prática, usando a tecnologia, desperta maior curiosidade e vontade de aprender.

## 5 Códigos

### Questão 1

```
set.seed(508492)
library(latex2exp)

# gerando uma A.A de tamanho n=50
A = rchisq(50,15)
A = round(sort(A),3);A

# breve descrição da amostra
mean(A); var(A); summary(A)

# estimando a moda de A
estimate_mode = function(x) {
  d = density(x)
  d$x[which.max(d$y)]
}

estimate_mode(A)

# a)  $P(X < 11.721)$ 
round(pchisq(11.721,15),3)
# b)  $P(X < x) = 0.5$ 
round(qchisq(0.5,15),3)

# podemos ver que a moda é 13
dchisq(13,15)<dchisq(12,15)
dchisq(13,15)<dchisq(15,15)

# função densidade de uma Chi-Quadrado com 15 g.l
f = function(x) dchisq(x,15)

# gráfico do item a)
plot(f,0,40,main=TeX("$P(X \leq 11.721) \approx 0.3$, com\\, $X \sim$,
  \(\chi^2_{15}\) $"))
ex=c(0,seq(0,11.721,l=30),11.721)
ey=c(0,f(seq(0,11.721,l=30)),0)
```

```

polygon(ex,ey,dens=10)
abline(v=13,col='red')
abline(v=15,col='blue')
abline(v=14.339,col='green')
axis(1,13)
axis(1,15)
legend(x="topright",legend=c(TeX("$M_{o}(X)=13$"),TeX("$\\mu =
↪ 15$"),TeX("$Md(X) \\approx
↪ 14.339$")),col=c('red','blue','green'),lty=1)

# gráfico do item b)
plot(f,0,40,main=TeX("$P(X \\leq 14.339) = 0.5, com\\, X \\sim\\,
↪ \\chi^2_{15} $"))
ex2 = c(0,seq(0,14.339,l=30),14.339)
ey2 = c(0,f(seq(0,14.339,l=30)),0)
polygon(ex2,ey2,dens=10)
abline(v=13,col='red')
abline(v=15,col='blue')
axis(1,14.339)
legend(x='topright',legend=c(TeX("$M_{o}(X)=13$"),TeX("$\\mu =
↪ 15$")),col=c('red','blue'),lty=1)

# vamos plotar a distribuição da nossa A.A
hist(A,prob=T,ylim=c(0,0.1),main=TeX("$Histograma\\, de\\, A \\sim\\,
↪ \\chi^2_{15}$"),ylab="Densidade")
curve(f,add=T,col='red')
lines(density(A),col='blue')
legend(x="topright",legend=c("População",
↪ "Amostra"),col=c('red','blue'),lty=1)

```

## Questão 2

```

set.seed(508492)
library(latex2exp)

# gerando uma A.A de tamanho n=50
A = rt(50,21)
A = round(sort(A),3);A

```

```

# breve descrição da amostra
mean(A); var(A); summary(A)

# estimando a moda de A
estimate_mode = function(x) {
  d = density(x)
  d$x[which.max(d$y)]
}

estimate_mode(A)

# a)  $P(X \leq 0.859)$ 
round(pt(0.859,21),3)
# b)  $P(X < x) = 0.6$ 
round(qt(0.6,21),3)

# função densidade de uma t de Student com 21 g.l
f = function(x) dt(x,21)

# gráfico do item a)
plot(f,-3,3,main=TeX("$P(X \\leq 0.859) = 0.8, com\\, X \\sim\\, \\rightarrow t(21)$"))
ex = c(-3,seq(-3,0.859,l=30),0.859)
ey = c(0,f(seq(-3,0.859,l=30)),0)
polygon(ex,ey,dens=10)
abline(v=0,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(X)=Med(X)=0$"),lty=1,col='red',
  ↪ cex=0.8)

# gráfico do item b)
plot(f,-3,3,main=TeX("$P(X \\leq 0.257) \\approx 0.6, com\\, X \\sim\\, \\rightarrow t(21)$"))
ex2 = c(-3,seq(-3,0.257,l=30),0.257)
ey2 = c(0,f(seq(-3,0.257,l=30)),0)
polygon(ex2,ey2,dens=10)
abline(v=0,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(X)=Med(X)=0$"),lty=1,col='red',
  ↪ cex=0.8)

```



```

# vamos plotar a distribuição da nossa A.A
hist(A,prob=T,main=TeX("$Histograma\\, de\\, A \\sim\\,
↪ t(21)$"),ylab="Densidade")
curve(f,add=T,col='red')
lines(density(A),col='blue')
legend(x="topright",legend = c("População",
↪ "Amostra"),col=c('red','blue'),lty=1)

```

### Questão 3

```

set.seed(508492)
library(latex2exp)

# gerando uma A.A de tamanho n=50
A = rf(50,15,15)
A = round(sort(A),3); A

# breve descrição da amostra
mean(A); var(A); summary(A)

# estimando a moda de A
estimate_mode = function(x) {
  d = density(x)
  d$x[which.max(d$y)]
}

estimate_mode(A)

# a)  $P(X \leq 2.4)$ 
round(pf(2.4,15,15),3)
# b)  $P(X \leq x) = 0.05$ 
round(qf(0.05,15,15),3)
round(1/qf(0.95,15,15),3)

# função densidade de uma F com 15 e 15 g.l
f = function(x) df(x,15,15)

# gráfico do item a)

```

```

plot(f,0,4,main=TeX("$P(X \\leq 2.4) = 0.95, com\\, X \\sim\\, \\rightarrow F(15,15)$"))
ex = c(0,seq(0,2.4,l=30),2.4)
ey = c(0,f(seq(0,2.4,l=30)),0)
polygon(ex,ey,dens=10)
axis(1,2.4)
abline(v=1,col='green')
abline(v=0.765,col='red')
abline(v=1.154,col='blue')
legend(x="topright",legend=c(TeX("$M_{o}(X) \\approx 0.765$"),TeX("$\\mu \\rightarrow \\approx 1.154$"),TeX("$Md(X) = 1")),col=c('red','blue','green'),lty=1)

# gráfico do item b)
plot(f,0,4,main=TeX("$P(X \\leq 0.416) \\approx 0.05, com\\, X \\sim\\, \\rightarrow F(15,15)$"))
ex2 = c(0,seq(0,0.41,l=30),0.41)
ey2 = c(0,f(seq(0,0.41,l=30)),0)
polygon(ex2,ey2,dens=10)
axis(1,0.416)
abline(v=1,col='green')
abline(v=0.765,col='red')
abline(v=1.154,col='blue')
legend(x="topright",legend=c(TeX("$M_{o}(X) \\approx 0.765$"),TeX("$\\mu \\rightarrow \\approx 1.154$"),TeX("$Md(X) = 1")),col=c('red','blue','green'),lty=1)

# histograma e distribuição amostral
hist(A,prob=T,ylim=c(0,0.95),main=TeX("$Histograma\\, de\\, A \\sim\\, \\rightarrow F(15,15)$"),ylab="Densidade")
curve(f,add=T,col='red')
lines(density(A),col='blue')
legend(x="topright",legend = c("População", "Amostra"),col=c('red','blue'),lty=1)

```

Questão 4

```

set.seed(508492)
library(latex2exp)

```

```

# gerando uma A.A de tamanho n=50
A = rnorm(50,40,5)
A = round(sort(A),3); A

# breve descrição da amostra
mean(A); var(A); summary(A)

# estimando a moda de A
estimate_mode = function(x) {
  d = density(x)
  d$x[which.max(d$y)]
}

estimate_mode(A)

# a)  $P(|X-40| < 12)$ 
round(pnorm(40+12,40,5) - pnorm(40-12,40,5),3) # direto
round(2*pnorm(12/5)-1,3) # padronizado
# b)  $P(X \leq x) = 0.05$ 
round(qnorm(0.05,40,5),3)
round(40-1.645*5,3)

# função densidade de uma  $N(40,25)$ 
f = function(x) dnorm(x,40,5)

# gráfico do item a), usando  $N(40,25)$ 
plot(f,40-3*5,40+3*5,main=TeX("$P(|X-40| < 12) \approx 0.984$, com\\, X
↪ \\sim\\, N(40,25)$"))
ex = c(40-12,seq(40-12,40+12,l=30),40+12)
ey = c(0,f(seq(40-12,40+12,l=30)),0)
polygon(ex,ey,col='grey')
text(40,0.03,TeX("$P(28 < X < 52) = 0.984$"),col="blue")
axis(1,c(28,52))
abline(v=40,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(X)=Med(X)=40$"),lty=1,col='red',
↪ cex=0.8)

```

```

# gráfico dpo item a), usando  $N(0,1)$ 
g = function(x) dnorm(x)
plot(g,-3,3,main=TeX("$P(|Z| < 2.4) \\approx 0.984$, com\\, Z \\sim\\,
↪  $N(0,1)$ "))
ex2 = c(-2.4,seq(-2.4,2.4,l=30),2.4)
ey2 = c(0,g(seq(-2.4,2.4,l=30)),0)
polygon(ex2,ey2,col='grey')
text(0,0.15,TeX("$P(-2.4 < Z < 2.4) = 0.984$"),col='blue')
axis(1,c(-2.4,2.4))
abline(v=0,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(Z)=Med(Z)=0$"),lty=1,col='red',
↪ cex=0.8)

# gráfico do item b), usando  $N(40,25)$ 
plot(f,40-3*5,40+3*5,main=TeX("$P(X \\leq 31.775) \\approx 0.05$, com\\,
↪  $X \\sim\\, N(40,25)$ "))
ex3 = c(40-3*5,seq(25,31.775,l=30),31.776)
ey3 = c(0,f(seq(25,31.775,l=30)),0)
polygon(ex3,ey3,col='grey')
axis(1,31.77)
abline(v=40,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(X)=Med(X)=40$"),lty=1,col='red',
↪ cex=0.8)

# gráfico do item b), usando  $N(0,1)$ 
plot(g,-3,3,main=TeX("$P(Z \\leq -1.645) \\approx 0.05$, com\\, Z
↪  $\\sim\\, N(0,1)$ "))
ex4 = c(-3,seq(-3,-1.645,l=30),-1.645)
ey4 = c(0,g(seq(-3,-1.645,l=30)),0)
polygon(ex4,ey4,col='grey')
axis(1,-1.645)
abline(v=0,col='red')
legend("topleft",legend=TeX("$\\mu=M_{o}(Z)=Med(Z)=0$"),lty=1,col='red',
↪ cex=0.8)

# histograma e distribuição amostral
hist(A,prob=T,main=TeX("$Histograma\\, de\\, A\\, \\sim\\,
↪  $N(40,25)$ "),ylim=c(0,0.08),ylab="Densidade")

```

```
curve(f,add=T,col='red')
lines(density(A),col='blue')
legend(x="topright",legend = c("População",
↪  "Amostra"),col=c('red','blue'),lty=1)
```