UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
BACHARELADO EM ESTATÍSTICA

ROMULO BARROS DE FREITAS

RELATÓRIO DA DISCIPLINA DE INFERÊNCIA ESTATÍSTICA I UTILIZANDO O SOFTWARE R

FORTALEZA
2023

ROMULO BARROS DE FREITAS

RELATÓRIO DA DISCIPLINA DE INFERÊNCIA ESTATÍSTICA I UTILIZANDO O SOFTWARE R

Relatório apresentado ao curso de Bacharelado em Estatística do Centro de Ciências da Universidade Federal do Ceará, como parte dos requisitos para a aprovação na disciplina de Inferência Estatística I no semestre de 2023.1.
Orientador: Prof. Dr. João Maurício Araújo Mota.

FORTALEZA
2023

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

As distribuições de probabilidade desempenham um papel fundamental na análise estatística, permitindo modelar e compreender o comportamento aleatório de fenômenos diversos. Elas fornecem uma estrutura teórica que nos permite fazer inferências sobre os dados observados e tomar decisões fundamentadas. Este relatório aborda as principais distribuições contínuas de probabilidade utilizadas na análise estatística: Qui-quadrado, t-Student, F-Snedecor e Normal, também conhecida como distribuição Gaussiana. Cada distribuição possui características únicas e aplicações específicas.

2 Objetivos

Este relatório tem o propósito de apresentar algumas análises e comparações estatísticas entre a população e uma amostra das principais distribuições contínuas utilizando o software estatístico R. As distribuições estudadas foram: Qui-quadrado com 15 graus de liberdade $X \sim \chi(15)$, t de Student com 21 graus de liberdade $X \sim t(21)$, F de Fisher-Snedecor com m = 15 graus de liberdade e n = 15 graus de liberadade $X \sim F(15, 15)$ e uma distribuição Normal com $\mu = 40$ e $\sigma^2 = 25$ $X \sim N(40, 25)$.

3 Metodologia

Para a realização do trabalho foram geradas amostras de tamanho 50 das distribuições de probabilidade citadas anteriormente usando como semente o número de matrícula 521353. A amostra foi gerada atráves do software gratuito conhecido como R, disponível para download no endereço www.rproject.org. Em seguida, foram calculadas as medidas populacionais e as medidas amostrais para termos uma comparação teórica e prática de cada variável. Por fim, foram gerados alguns gráficos, como histogramas, e qqplots, juntamente com o teste de Shapiro-Wilk para verificar a possível normalidade de cada amostra. É importante salientar que todas as análises foram feitas utilizando o RStudio, interface do software R.

4 Análises

Questão 01

Seja $X \sim \chi(15)$. Responda ao que se pede:
a. Calcule pa $= P(X \leq 11,721)$.

```
pa = pchisq(11.721, 15, lower.tail = TRUE);pa;round(pa, 3)
[1] 0.2999874
[1] 0.3
```

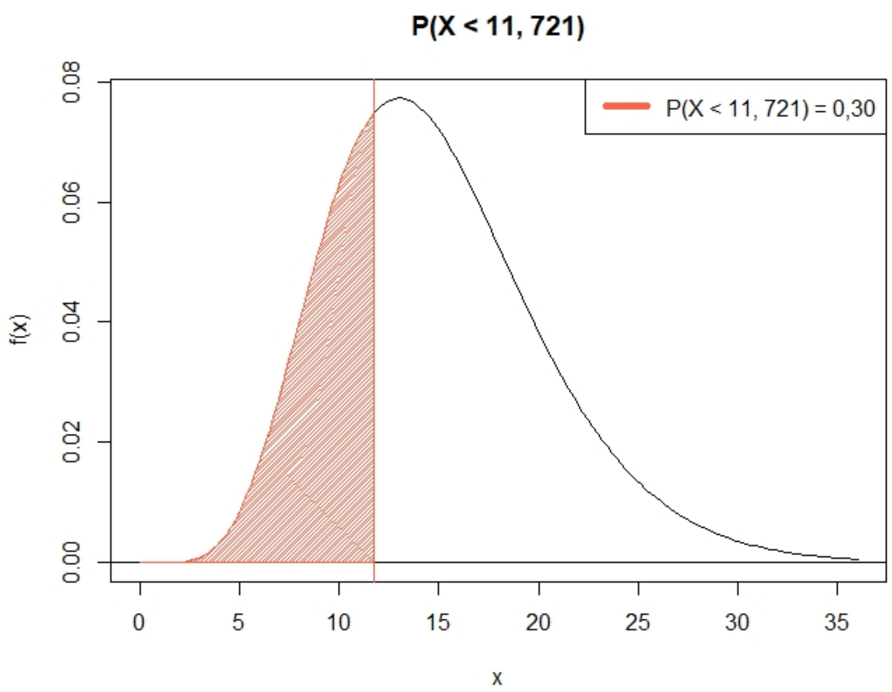

Figura 4.1: pa $= P(X \leq 11,721)$

No gráfico acima podemos visualizar a área hachurada correspondente à probabilidade $P(X \leq 11,721)$. A área equivale a 0,30.

b. Qual a Mediana de X?

```
Md = qchisq(0.50, 15);Md;round(Md, 3) # Mediana de X
[1] 14.33886
[1] 14.339
```

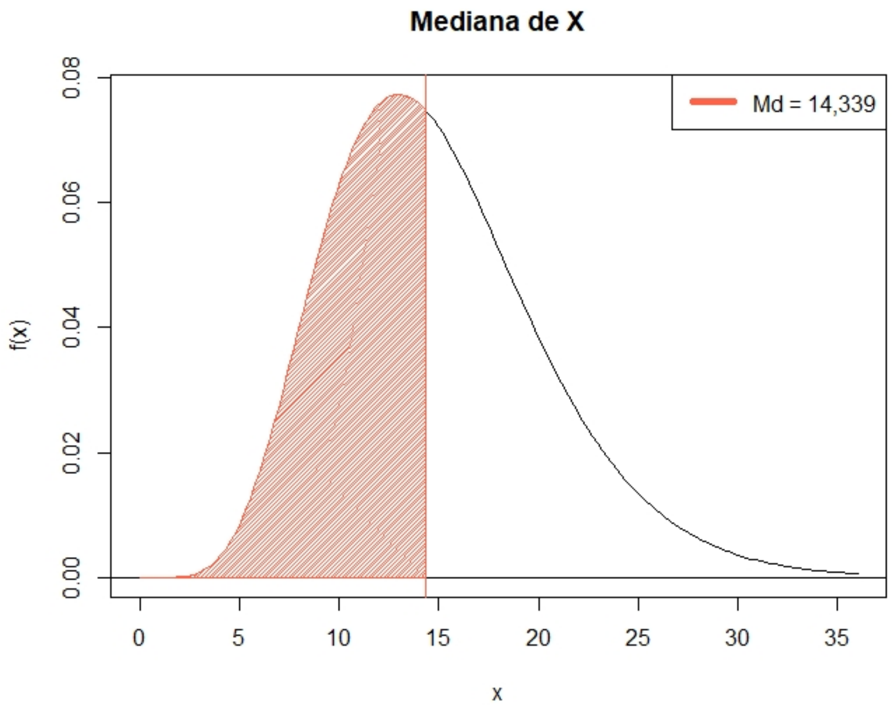

Figura 4.2: Mediana de X

No gráfico acima, a linha vermelha na vertical corresponde a 14,339, valor referente à mediana de $X \sim \chi(15)$. Ou seja, 14,339 é o valor que separa nossos dados meio a meio. Ela é muito utilizada em distribuições assimétricas como é caso da Qui-Quadrado.

4.1 Análise de Medidas Populacionais e de Medidas Amostrais de $X \sim \chi(15)$

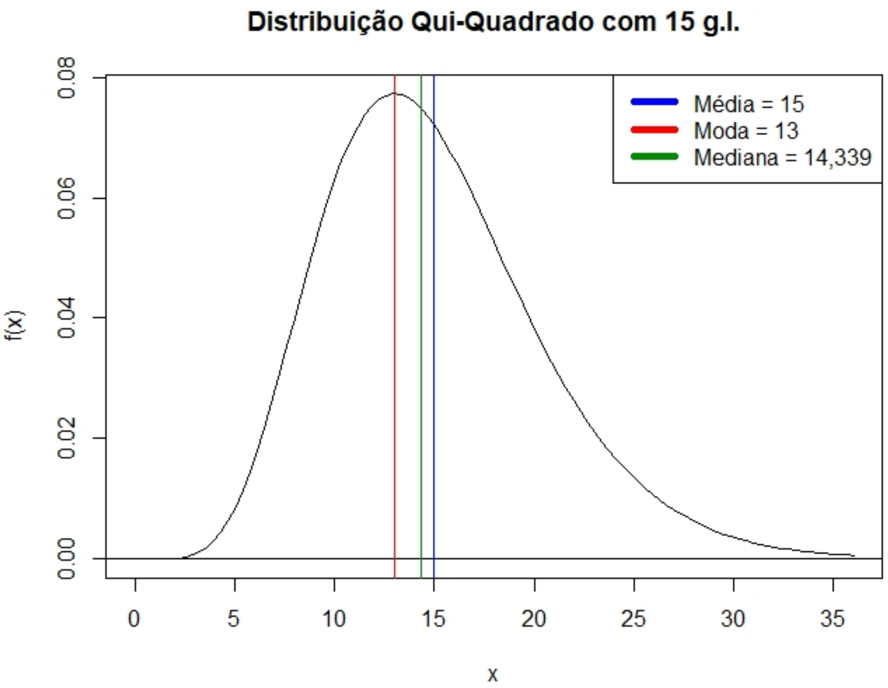

Figura 4.3: Medidas populacionais de $X \sim \chi(15)$

No gráfico acima está destacado algumas medidas de tendência central (Média, Moda e Mediana) de uma variável qui-quadrado com 15 graus de liberdade. A média é dado pelo total de graus de liberdade da variável, logo, $\mu = k = 15$, a moda é dada por $Mo = max(k-2, 0) = max(15-2, 0) = 13$, já a mediana, calculada no item b da questão 1 é Md = 14,339. Além disso, a variância $\sigma^2 = 2k = 30$.

Após ser gerada uma amostra de tamanho 50 usando o RStudio, implementando uma semente com o número de matrícula 521353, foram calculadas algumas medidas descritivas da amostra de análise. A seguir, temos as estatísticas pontuais de uma amostra de tamanho n = 50 de uma variável com distribuição Qui-Quadrado com 15 graus de liberdade, onde DP significa Desvio Padrão da amostra.

Estatísticas Pontuais da Amostra						
Mínimo	**1º Quartil**	**Mediana**	**Média**	**3º Quartil**	**Máximo**	**DP**
5,655	9,554	12,699	13,692	15,197	33,811	5,610

Tabela 4.1: Estatísticas Pontuais de uma Amostra n = 50 da Variável $X \sim \chi(15)$

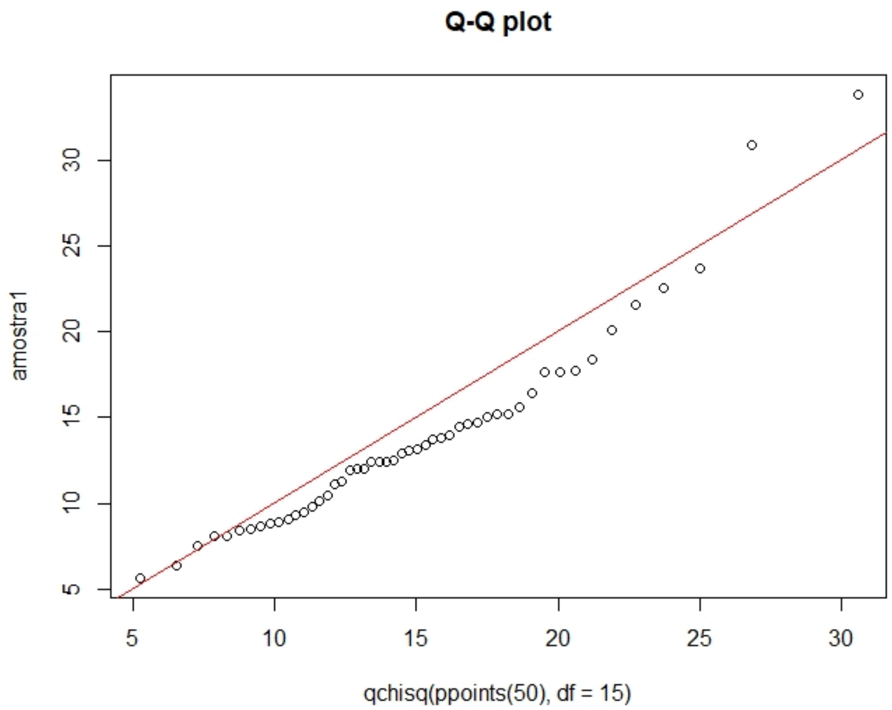

Figura 4.4: QQ Plot de uma Amostra n = 50 da Variável $X \sim \chi(15)$

Na figura 4.4 temos a representação de um qqplot de uma Amostra n = 50 da Variável $X \sim \chi(15)$. Esse tipo de gráfico serve para termos uma noção visual se os dados da nossa amostra seguem ou não uma distribuição normal. Ao visualizarmos o gráfico percebemos que os valores da nossa amostra tendem a se concentrar em torno da reta, o que nos leva suspeitar da hipóteses de nossos valores observados seguem uma distribuição normal. Entretanto, isso é apenas uma hipótese visual, não podemos afirmar tal hipótese apenas graficamente. Portanto, foi realizado um teste denominado Shapiro-Wilk para a confirmação da hipótese. É importante pontuar que, existem diversos outros testes, como o de Kolmogorov Smirnorv, mas optou-se pelo primeiro Shapiro-Wilk por conta do tamanho da amostra.

```
shapiro.test(amostra1)

        Shapiro-Wilk normality test

data:  amostra1
W = 0.87348, p-value = 7.137e-05
```

Ao analisarmos o resultado notamos que $p - valor = 7,137e - 05 < 0.05$, ou seja, ao nível de significância de 5%, rejeitamos a hipótese de normalidade nos nossos dados amostrais. A seguir, temos o histograma da nossa amostra:

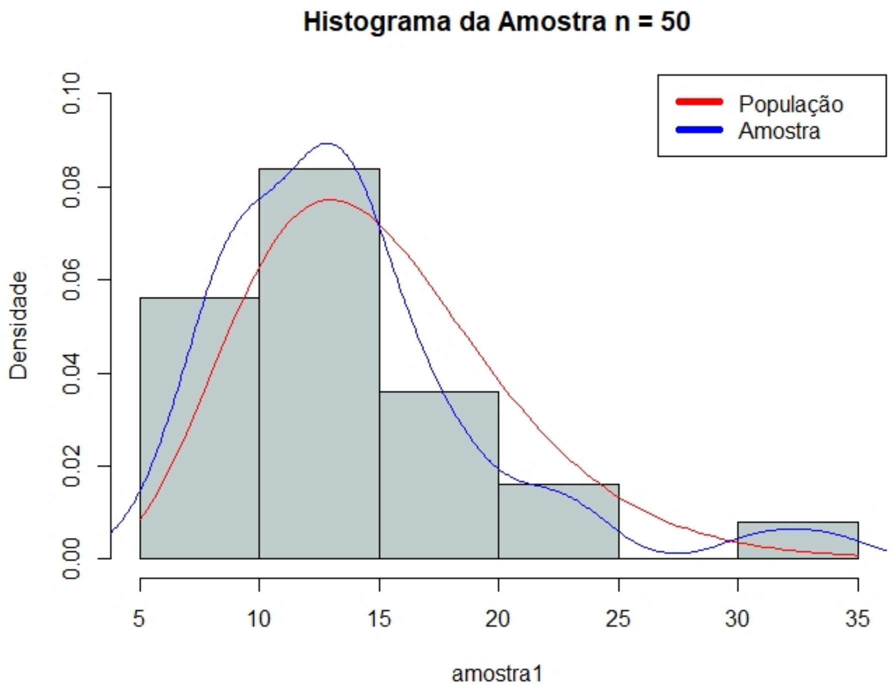

Figura 4.5: Histograma com a curva populacional e a curva amostral

A figura acima representa o histograma de uma amostra de tamanho n = 50 de uma variável $X \sim \chi(15)$. Nele foi plotado a curva populacional da distribuição na cor vermelha e a curva correspondente a amostra na cor azul.

Questão 2

Seja $X \sim t(21)$. Responda ao que se pede:
a. Calcule p = $P(X \leq 0,859)$.

```
pa = pt(0.859, 21);pa;round(pa,3)
[1] 0.79998
[1] 0.8
```

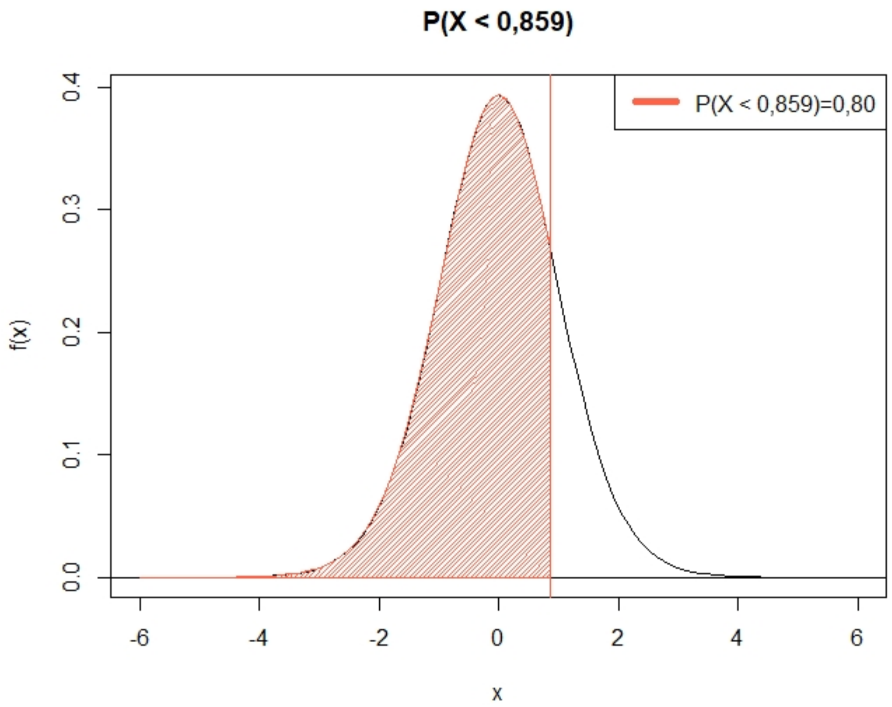

Figura 4.6: p = $P(X \leq 0,859)$

No gráfico acima podemos visualizar a área hachurada correspondente à probabilidade $P(X \leq 0,859)$. A área equivale a 0.80.

b. Qual o sexto decil de X?

```
d6 = qt(0.60, 21);d6;round(d6, 3) # Decil 6 de X
[1] 0.2565799
[1] 0.257
```

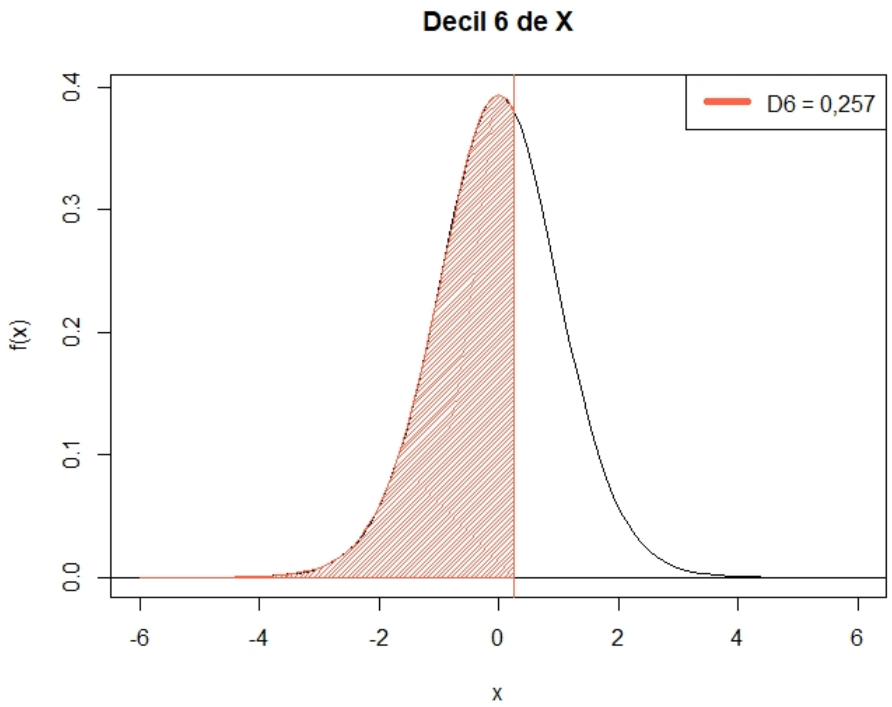

Figura 4.7: Decil 6 de X

No gráfico acima, a linha vermelha na vertical corresponde a 0,257, valor referente ao sexto decil de $X \sim t(21)$. Ou seja, até o valor 0,257 estão acumulados 60% dos dados populacionais.

4.2 Análise de Medidas Populacionais e de Medidas Amostrais de $X \sim t(21)$

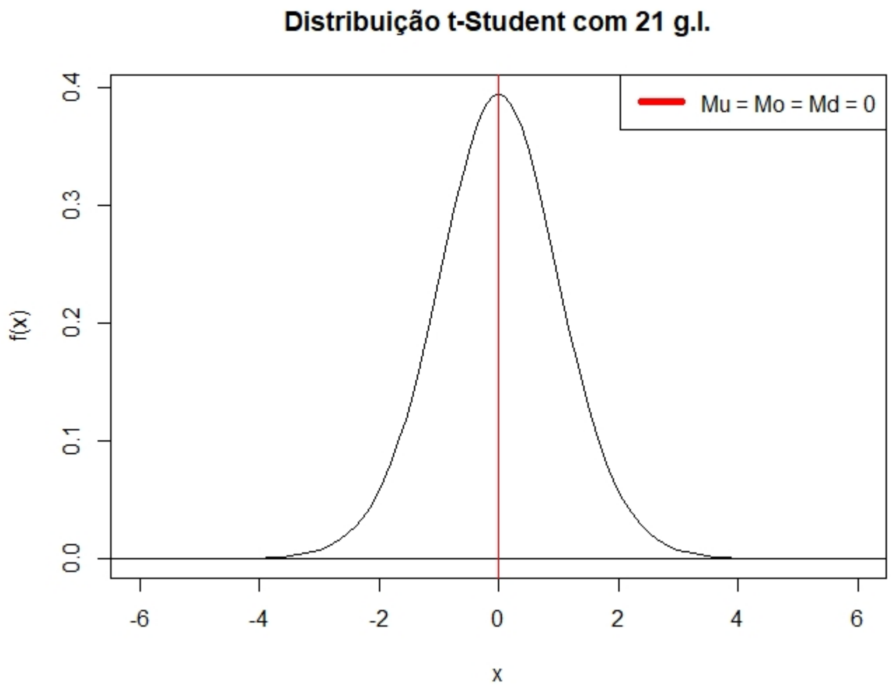

Figura 4.8: Medidas populacionais de $X \sim t(21)$

No gráfico acima está destacado algumas medidas de tendência central (Média, Moda e Mediana) de uma variável t-Student com k = 21 graus de liberdade. Trata-se uma distribuição simétrica em torno do zero, sua média é dada por $\mu = 0, k > 0$ e sua Moda e Mediana também são zero. $\mu = Mo = Md = 0$. Além disso, a variância $\sigma^2 = \frac{k}{k-2}, k > 2$, ou seja, $\sigma^2 = 1,1$.

Após ser gerada uma amostra de tamanho 50 usando o RStudio, implementando uma semente com o número de matrícula 521353, foram calculadas algumas medidas descritivas da amostra de análise. A seguir, temos as estatísticas pontuais de uma amostra de tamanho n = 50 de uma variável com distribuição T de Student com 21 graus de liberdade, onde DP significa Desvio Padrão da amostra.

Estatísticas Pontuais da Amostra						
Mínimo	**1º Quartil**	**Mediana**	**Média**	**3º Quartil**	**Máximo**	**DP**
-2,54072	-0,65627	-0,14323	-0,09844	0,58420	2,15228	0,945

Tabela 4.2: Estatísticas Pontuais de uma Amostra n = 50 da Variável $X \sim t(21)$

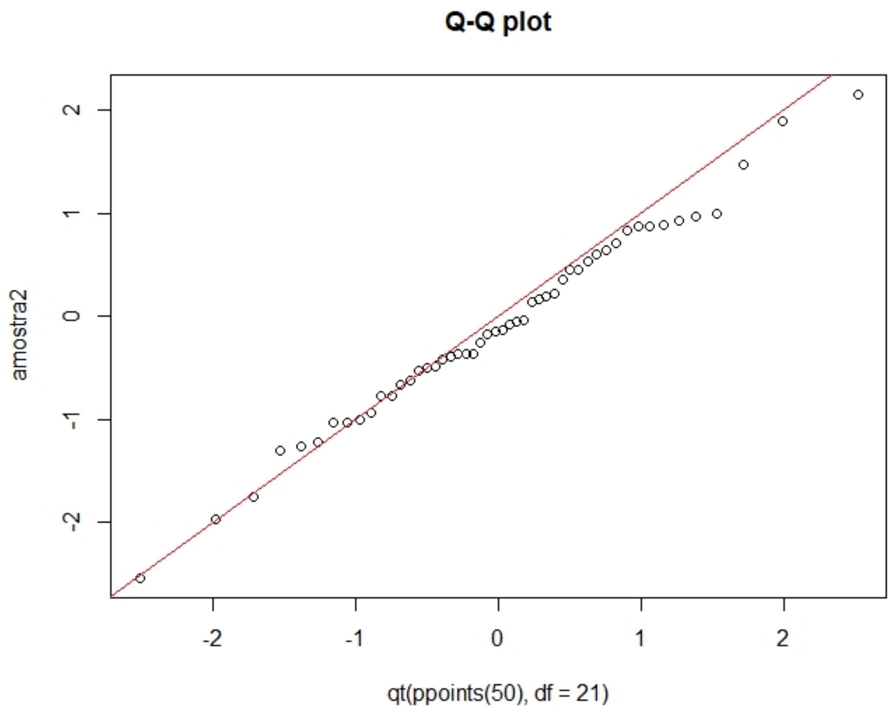

Figura 4.9: QQ Plot de uma Amostra n = 50 da Variável $X \sim t(21)$

Na figura 4.11 temos a representação de um qqplot de uma Amostra n = 50 da Variável $X \sim t(21)$. Em seguida, foi realizado um teste Shapiro-Wilk para a confirmação ou não da hipótese de normalidade dos dados amostrais.

```
shapiro.test(amostra2)

        Shapiro-Wilk normality test

data:  amostra2
W = 0.99108, p-value = 0.9681
```

Ao analisarmos o resultado notamos que $p - valor = 0.9681 > 0.05$, ou seja, ao nível de significância de 5%, podemos suspeitar normalidade nos nossos dados amostrais. A seguir, temos o histograma da nossa amostra:

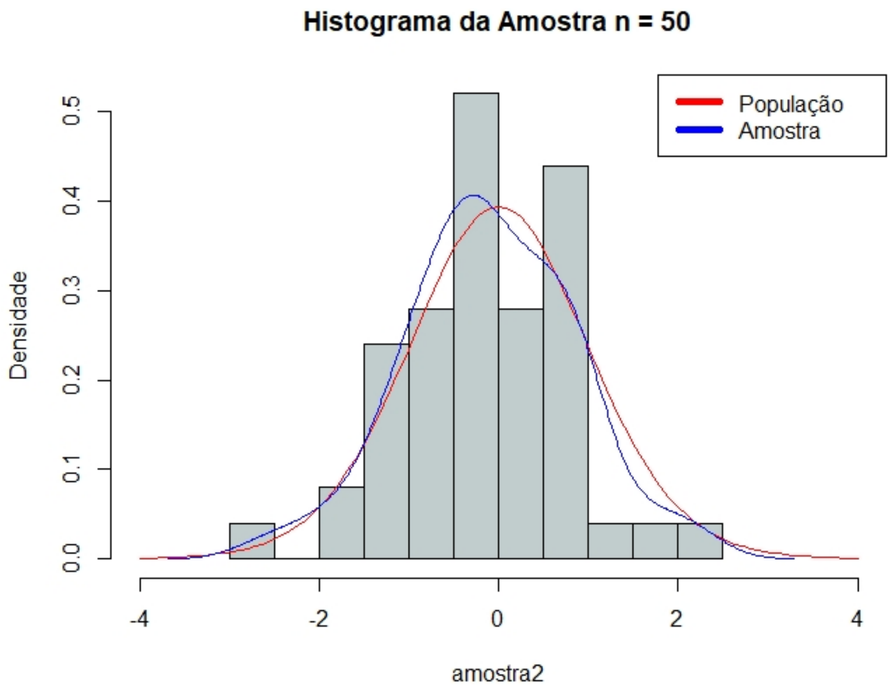

Figura 4.10: Histograma com a curva populacional e a curva amostral

A figura acima representa o histograma de uma amostra de tamanho n = 50 de uma variável $X \sim t(21)$. Nele foi plotado a curva populacional da distribuição na cor vermelha e a curva correspondente a amostra na cor azul.

Questão 3

Seja $X \sim F(15, 15)$. Responda ao que se pede:
a. Calcule pa $= P(X \leq 2,40)$.

```
pa = pf(2.40,15,15);pa;round(pa,2)
[1] 0.9497299
[1] 0.95
```

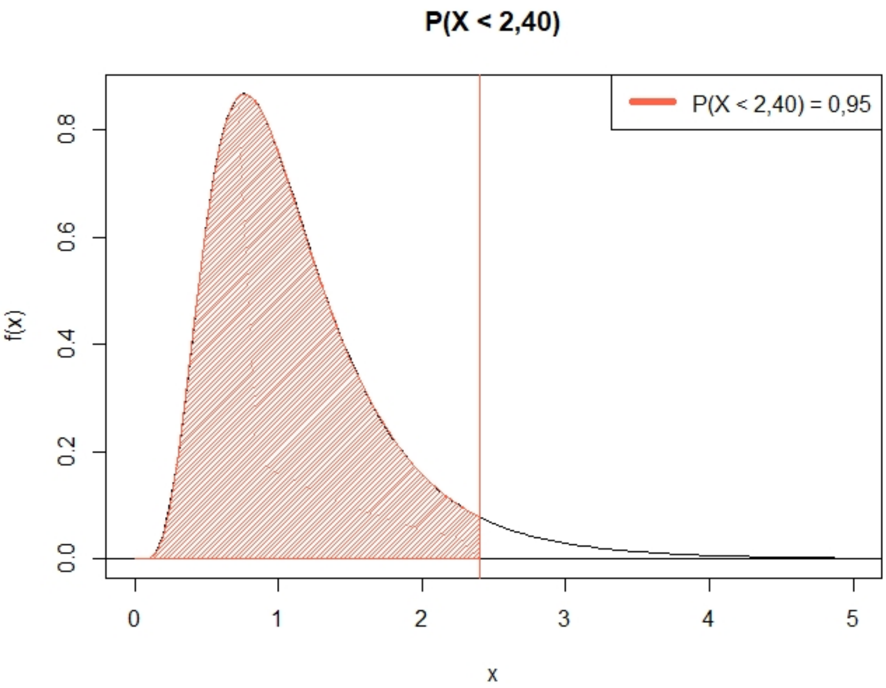

Figura 4.11: $P(X \leq 2,40)$

No gráfico acima podemos visualizar a área hachurada correspondente à probabilidade $P(X \leq 2,40)$. A área equivale a 0.95.

b. Qual o quinto percentil de X?

```
p5 = qf(0.05, 15, 15, lower.tail = TRUE);p5;round(p5,2)
[1] 0.4160691
[1] 0.42
```

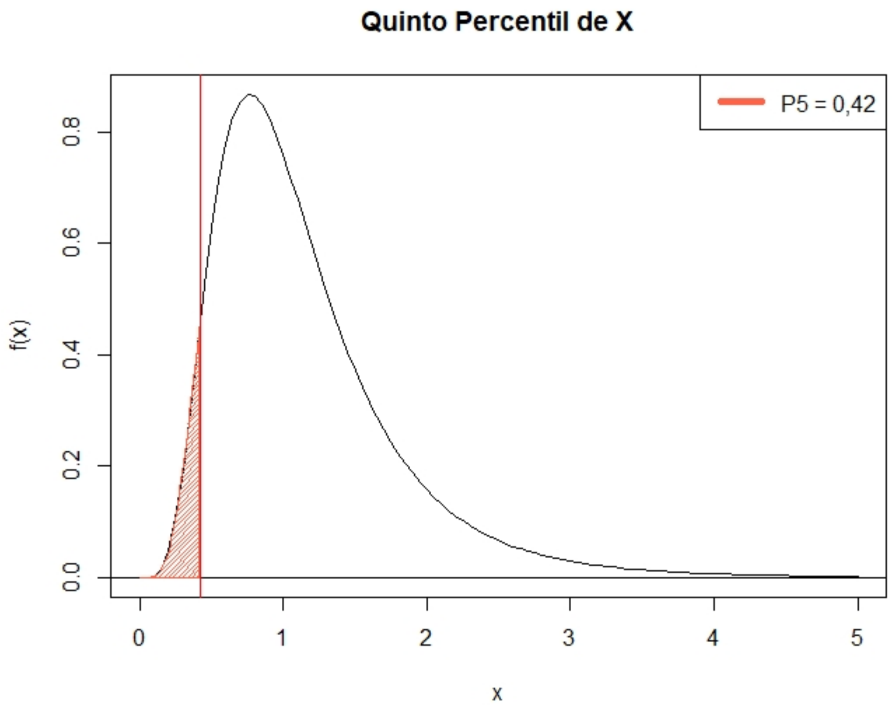

Figura 4.12: Quinto percentil de X

No gráfico acima, a linha vermelha corresponde ao valor 0,42, referente ao quinto percentil de $X \sim F(15, 15)$. Ou seja, até 0,42 estão acumulados 5% dos dados populacionais.

4.3 Análise de Medidas Populacionais e de Medidas Amostrais de $X \sim F(15, 15)$

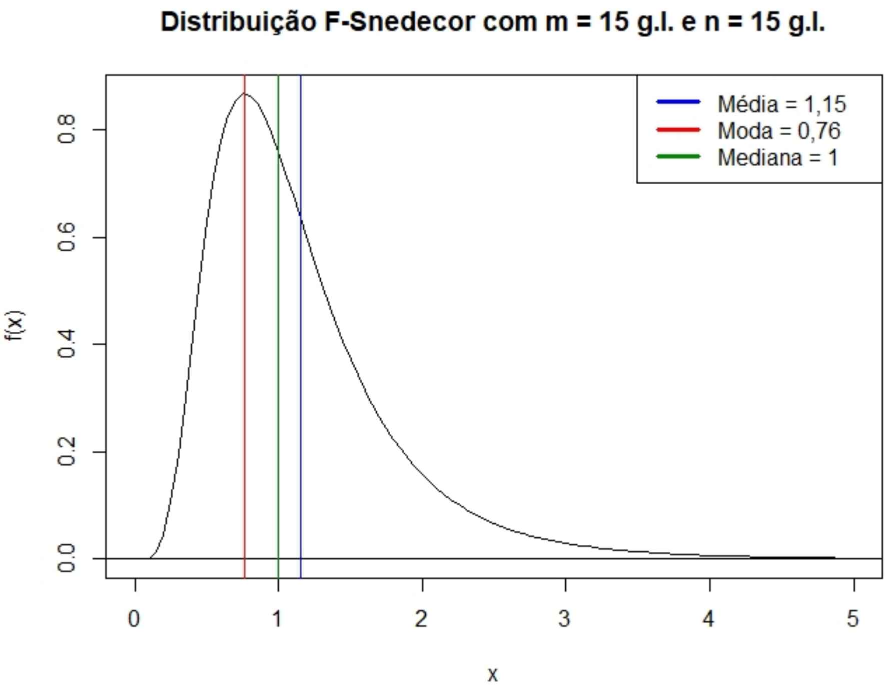

Figura 4.13: Medidas populacionais de $X \sim F(15, 15)$

No gráfico acima está destacado algumas medidas de tendência central (Média, Moda e Mediana) de uma variável F-Snedecor com m = 15 graus de liberdade e n = 15 graus de liberadade. A média é dada por $\mu = \frac{n}{n-2}, n > 2$. A moda é dada por $Mo = \frac{n(m-2)}{m(n-2)}$. Logo, $\mu = 1,15$, $Mo = 0,76$, já a mediana, a medida que divide o conjunto de dados meio a meio está entre a média e a moda. Ela foi obtida pelo seguinte comando do RStudio:

```
#Mediana
Md = qf(0.5, 15, 15, lower.tail = TRUE)
[1] 1
```

Após ser gerada uma amostra de tamanho 50 usando o RStudio, implementando uma semente com o número de matrícula 521353, foram calculadas algumas medidas descritivas da amostra de análise. A seguir, temos as estatísticas pontuais de uma amostra de tamanho n = 50 de uma variável com distribuição F-Snedecor com m = 15 graus de liberdade e n = 15 graus de liberdade, onde DP significa Desvio Padrão da amostra.

Estatísticas Pontuais da Amostra						
Mínimo	**1º Quartil**	**Mediana**	**Média**	**3º Quartil**	**Máximo**	**DP**
0,2559	0,7890	1,1835	1,2938	1.5465	2,15228	0,844

Tabela 4.3: Estatísticas Pontuais de uma Amostra n = 50 da Variável $X \sim F(15, 15)$

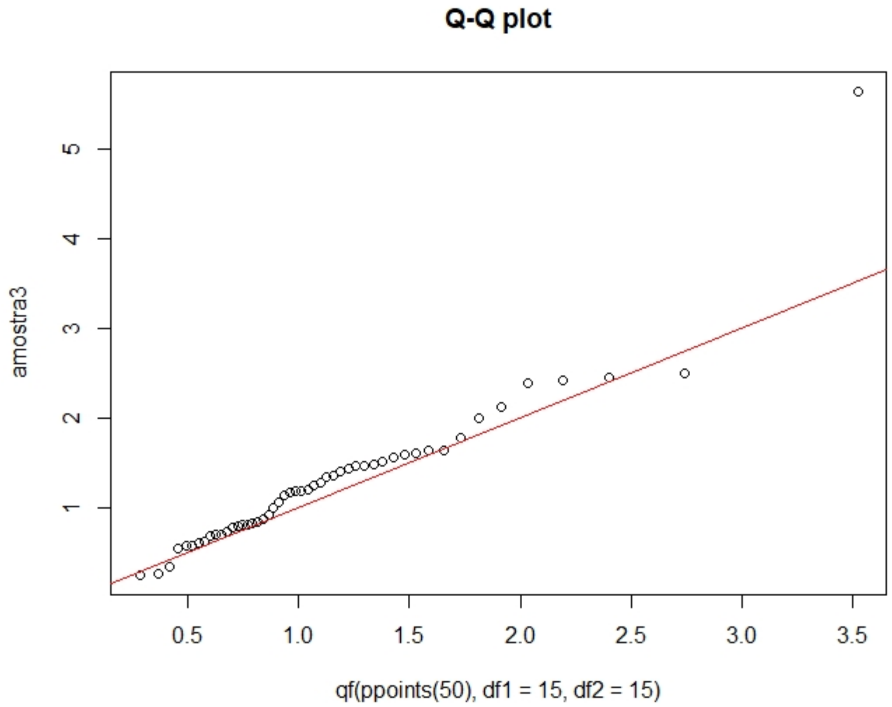

Figura 4.14: QQ Plot de uma Amostra n = 50 da Variável $X \sim F(15, 15)$

Na figura 4.14 temos a representação de um qqplot de uma Amostra n = 50 da Variável $X \sim F(15, 15)$. Em seguida, foi realizado um teste Shapiro-Wilk para a confirmação ou não da hipótese de normalidade dos dados amostrais.

```
shapiro.test(amostra3)

        Shapiro-Wilk normality test

data:  amostra3
W = 0.76031, p-value = 1.195e-07
```

Ao analisarmos o resultado notamos que $p - valor = 1,195e - 07 < 0.05$, ou seja, ao nível de significância de 5%, rejeitamos a hipótese de normalidade nos nossos dados amostrais. A seguir, temos o histograma da nossa amostra:

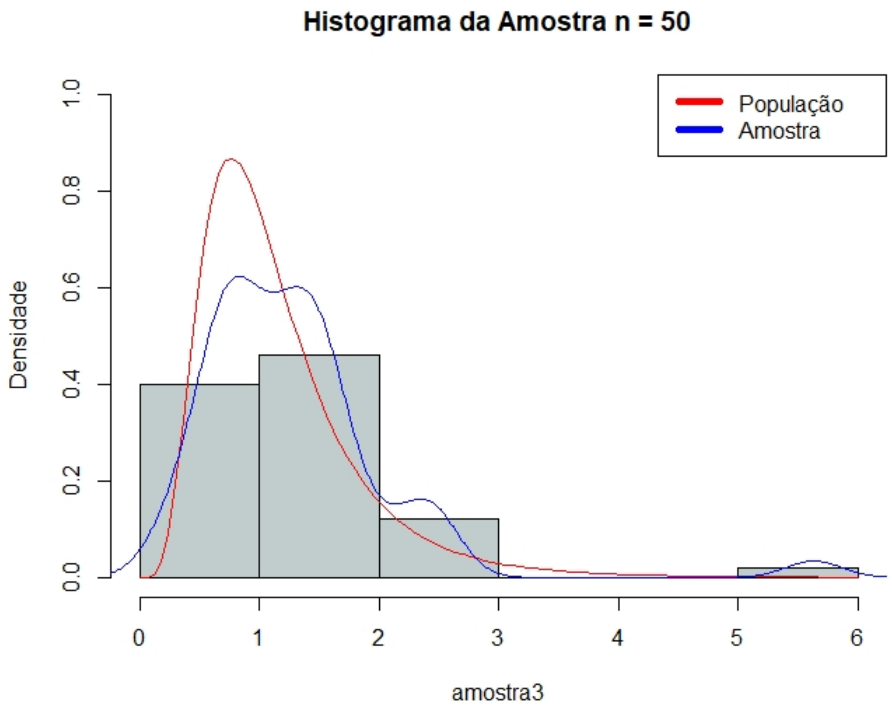

Figura 4.15: Histograma com a curva populacional e a curva amostral

A figura acima representa o histograma de uma amostra de tamanho n = 50 de uma variável $X \sim F(15, 15)$. Nele foi plotado a curva populacional da distribuição na cor vermelha e a curva correspondente a amostra na cor azul.

Questão 4

Seja $X \sim N(40, 25)$. Responda ao que se pede:
a. Calcule p = $P(|X - 40| < 12)$.

```
mu=40;sigma=5
z=12/sigma;z
[1] 2.4
p=2*0.4918;p
[1] 0.9836
```

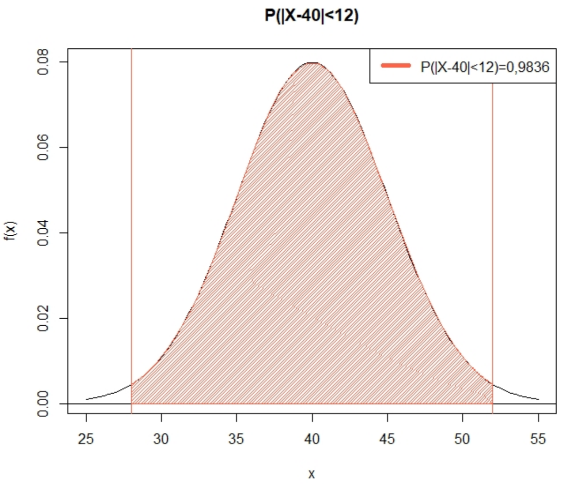

Figura 4.16: $P(|X-40|<12)$

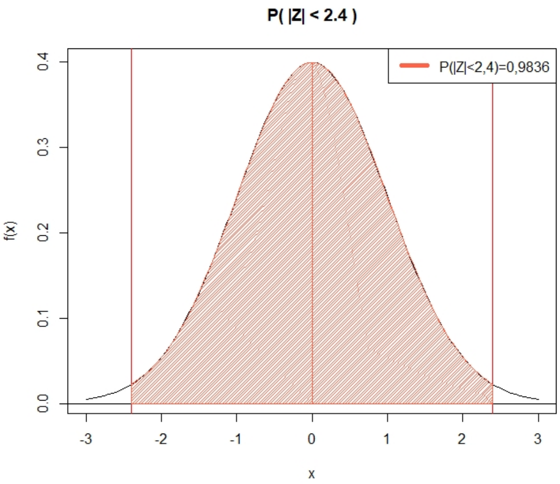

Figura 4.17: $P(|Z|<2,4)$

Nos gráficos acima temos a área hachurada da probabilidade $P(|X-40|<12)$ e de sua padronização, $P(|Z|<2,4)$, respectivamente. O valor da área hachurada é de 0,9836.

b. Qual o percentil de ordem 5 de X?

```
# Percentil de ordem 5
round(qnorm(0.05,40,5),3)
[1] 31.776
```

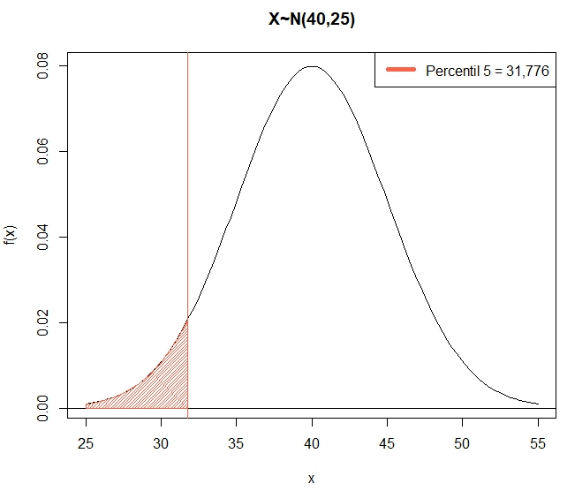

Figura 4.18: Percentil 5 de uma Normal $X \sim N(40,25)$

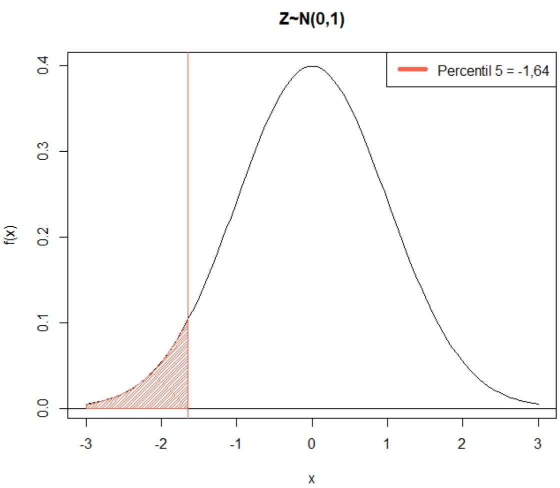

Figura 4.19: Percentil 5 de uma Normal Padrão $Z \sim N(0,1)$

Nos gráficos acima temos o quinto percentil de uma $X \sim N(40,25)$ e o quinto percentil de uma $Z \sim N(0,1)$, respectivamente. O valor do quinto percentil de uma $X \sim N(40,25)$ é 31,776 e o valor do quinto percentil de uma $Z \sim N(0,1)$ é -1,64.

4.4 Estatísticas Pontuais de uma Amostra n = 50 da Variável $X \sim N(40, 25)$

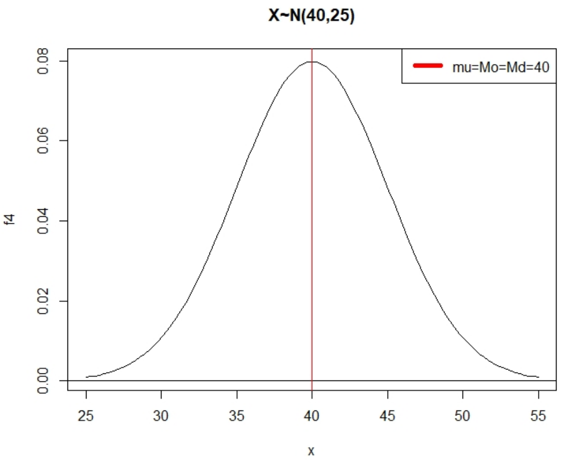

Figura 4.20: Medidas populacionais de $X \sim N(40, 25)$

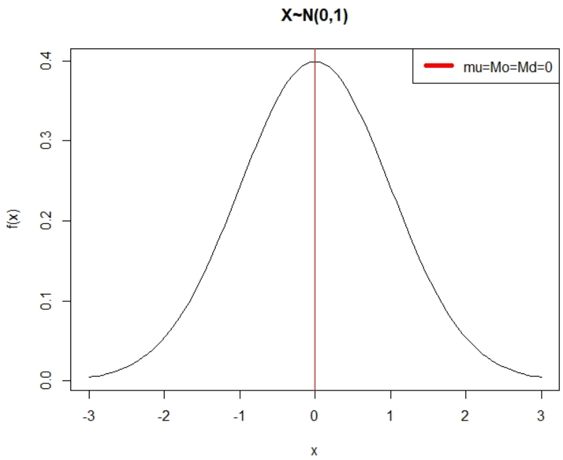

Figura 4.21: Medidas de uma Normal Padrão $Z \sim N(0, 1)$

No gráfico acima está destacado com a linha vermelha na vertical algumas medidas de tendência central (Média, Moda e Mediana) de uma variável Normal com média $\mu = 40$ e variância $\sigma^2 = 25$. Como a distribuição normal é simétrica, $\mu = Mo = Md$, no caso da normal padrão, $\mu = Mo = Md = 0$.

Após ser gerada uma amostra de tamanho 50 usando o RStudio, implementando uma semente com o número de matrícula 521353, foram calculadas algumas medidas descritivas da amostra de análise. A seguir, temos as estatísticas pontuais de uma amostra de tamanho n = 50 de uma variável com distribuição normal com média $\mu = 40$ e variância $\sigma^2 = 25$, onde DP significa Desvio Padrão da amostra.

Estatísticas Pontuais da Amostra						
Mínimo	**1º Quartil**	**Mediana**	**Média**	**3º Quartil**	**Máximo**	**DP**
28,82	36,58	40,21	39,89	42,79	53,25	4,789

Tabela 4.4: Estatísticas Pontuais de uma Amostra n = 50 da Variável $X \sim N(40, 25)$

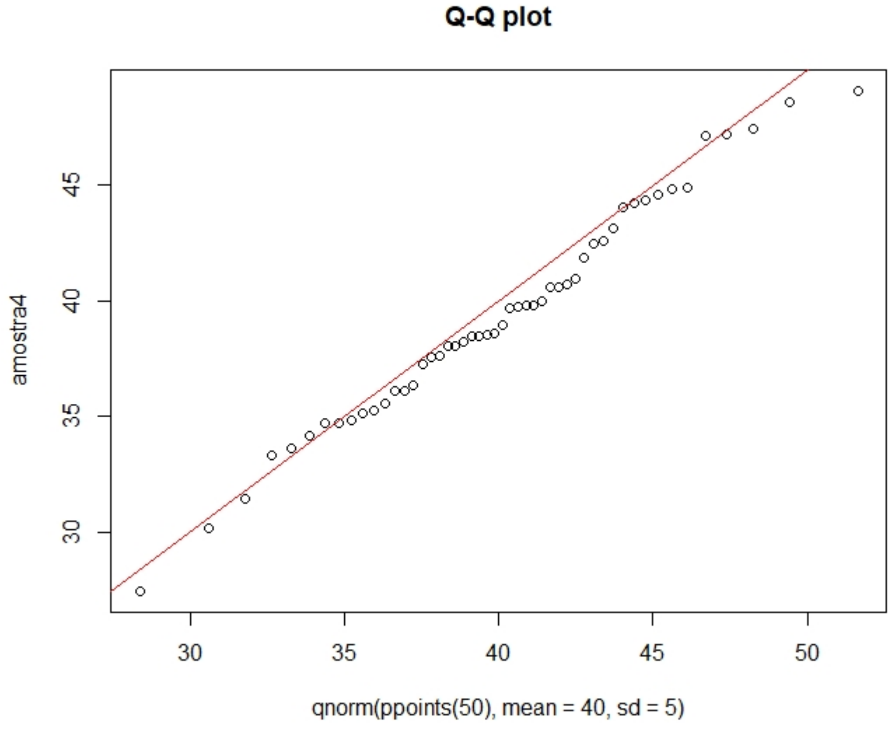

Figura 4.22: QQ Plot de uma Amostra n = 50 da Variável $X \sim N(40, 25)$

Na figura 4.18 temos a representação de um qqplot de uma Amostra n = 50 da Variável $X \sim N(40, 25)$. Em seguida, foi realizado um teste denominado Shapiro-Wilk para a confirmação ou não da hipótese de normalidade dos dados amostrais.

```
Shapiro-Wilk normality test

data:  amostra4
W = 0.98623, p-value = 0.8231
```

Ao analisarmos o resultado notamos que $p - valor = 0,8231 > 0.05$, ou seja, ao nível de significância de 5%, podemos suspeitar a normalidade dos nossos dados amostrais. A seguir, temos o histograma da nossa amostra:

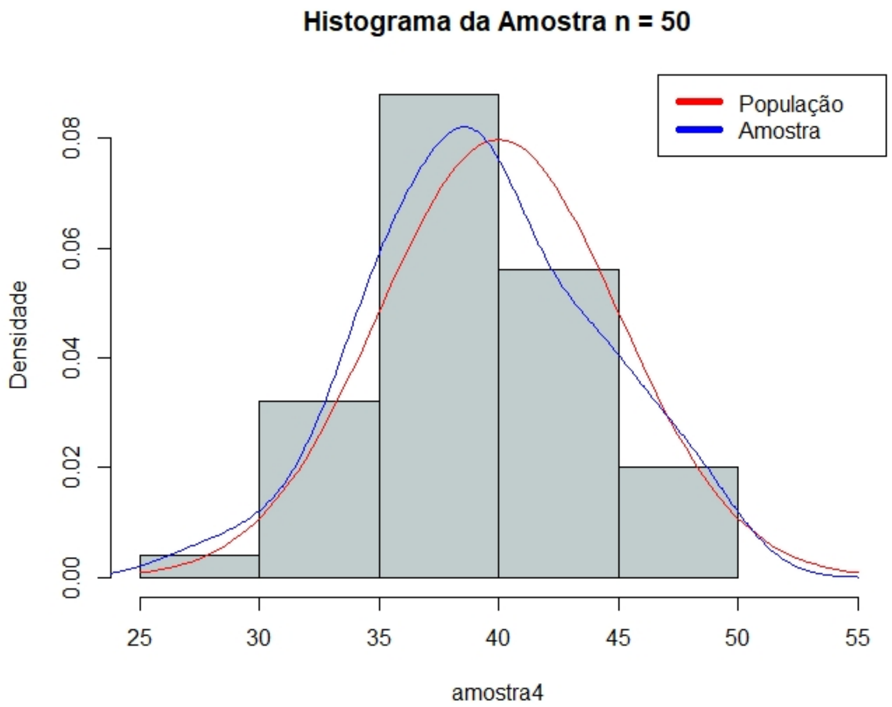

Figura 4.23: Histograma com a curva populacional e a curva amostral

A figura acima representa o histograma de uma amostra de tamanho n = 50 de uma variável $X \sim N(40, 25)$. Nele foi plotado a curva populacional da distribuição na cor vermelha e a curva correspondente a amostra na cor azul.

5 Anexo

Neste anexo estão presentes os códigos em R que foram utilizados para a realização do relatório.

5.1 Códigos da Questão 1

```
# Dados populacionais:
r = 15 # g.l.
mu = r # Media
Mdn = p5 # Mediana
Mo = max(r-2, 0) # Moda

# Definindo a função e plotando os gráficos:
f1 = function(x) dchisq(x, 15)
plot(f1, 0, 36, main = 'Distribuição Qui-Quadrado com 15 g.l.', ylab =
↪ 'f(x)')
abline(h=0, col='black')
abline(v=mu, col = 'blue') # Media
abline(v=Mo, col = 'red') # Moda
abline(v=Mdn, col = 'green4') # Mediana
#Legenda
legend("topright", legend = c('Média = 15', 'Moda = 13', 'Mediana =
↪ 14,339'),
       col = c('blue', 'red', 'green4'), lwd = 5)

# GRÁFICO HACHURADO A
plot(x, f1,type="l", main="P(X < 11, 721)", ylab="f(x)")
abline(h=0, col='black')
abline(v=11.721, col = 'tomato')
ax=c(0,0,x[x<11.721],11.721,11.721)
ay=c(0,dchisq(c(0,x[x<11.721],11.721),15),0)
polygon(ax,ay, dens = 50, col = 'tomato')
#Legenda
legend("topright", legend = c('P(X < 11, 721) = 0,30'),
       col = c('tomato'), lwd = 5)

# GRAFICO HACHURADO B
plot(f2, -6, 6, main = 'Decil 6 de X', ylab = 'f(x)')
```

```
abline(h=0, col='black')
abline(v=0.257, col = 'tomato')
polygon(x = c(-6, seq(-6, 0.257, l=50),0.257),
        y = c(0, f2(seq(-6, 0.257,l=50)), 0),
        col = 'tomato', density = 40)
#Legenda
legend("topright", legend = c('D6 = 0,257'),
       col = c('tomato'), lwd = 5)
```

```
 # Dados amostrais:
> set.seed(521353)
> amostra1 = rchisq(50, 15);sort(amostra1)
 [1]  5.654735  6.386517  7.546076  8.049069  8.082717  8.386685
 ↪  8.498309  8.651397  8.796119  8.939903
[11]  9.074225  9.337691  9.475687  9.787520 10.162185 10.463472
↪  11.119402 11.272567 11.934300 12.019123
[21] 12.021504 12.376646 12.423886 12.434691 12.522610 12.874508
↪  13.050868 13.148501 13.414650 13.745850
[31] 13.769147 13.990007 14.472856 14.610422 14.741284 15.046443
↪  15.154507 15.211572 15.606661 16.411732
[41] 17.660721 17.676548 17.687292 18.337885 20.057428 21.562039
↪  22.562398 23.703908 30.889265 33.810590
> mean(amostra1);median(amostra1);var(amostra1);sd(amostra1)
[1] 13.69228
[1] 12.69856
[1] 31.47466
[1] 5.610228
> summary(amostra1)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.655   9.554  12.699  13.692  15.197  33.811

# Histograma
hist(amostra1, prob = T, col = 'azure3',
     main = 'Histograma da Amostra n = 50', ylab = 'Densidade', ylim =
      ↪  c(0, 0.1))
curve(f1, add = T, col = 'red')
lines(density(amostra1), col = 'blue')
#Legenda
```

```
legend("topright", legend = c('População', 'Amostra'),
       col = c('red', 'blue'), lwd = 5)
```

5.2 Códigos da Questão 2

```
# Dados populacionais:
n = 21 # g.l.
mu = 0 # Media
Mo = 0 # Moda
Mdn = qt(0.50, 21);Mdn;round(Mdn, 3) # Mediana
[1] 0
[1] 0

# Gerando a função e plotando o gráfico:
f2 = function(x) dt(x, 21)
plot(f2, -6, 6, main = 'Distribui??o t-Student com 21 g.l.', ylab =
↪  'f(x)')
abline(h=0, col='black')
abline(v=0, col = 'red') # Media, Moda e Mediana
#Legenda
legend("topright", legend = c('Mu = Mo = Md = 0'),
       col = c('red'), lwd = 5)

# GRAFICO HACHURADO A
plot(f2, -6, 6, main = 'P(X < 0,859)', ylab = 'f(x)')
abline(h=0, col='black')
abline(v=0.859, col = 'tomato')
polygon(x = c(-6, seq(-6, 0.859, l=50),0.859),
        y = c(0, f2(seq(-6, 0.859,l=50)), 0),
        col = 'tomato', density = 40, text())
#Legenda
legend("topright", legend = c('P(X < 0,859)=0,80'),
       col = c('tomato'), lwd = 5)

# GRAFICO HACHURADO B
plot(f2, -6, 6, main = 'Decil 6 de X', ylab = 'f(x)')
abline(h=0, col='black')
```

```
abline(v=0.257, col = 'tomato')
polygon(x = c(-6, seq(-6, 0.257, l=50),0.257),
        y = c(0, f2(seq(-6, 0.257,l=50)), 0),
        col = 'tomato', density = 40)
#Legenda
legend("topright", legend = c('D6 = 0,257'),
       col = c('tomato'), lwd = 5)
```

```
> # Dados amostrais:
> set.seed(521353)
> amostra2 = rt(50, 21);sort(amostra2)
 [1] -2.54071701 -1.97339468 -1.75524201 -1.30753159 -1.27019853
 ↪  -1.21827561 -1.03695757 -1.02602062
 [9] -1.00320321 -0.94128934 -0.77171990 -0.76771431 -0.66593646
 ↪  -0.62725381 -0.52338401 -0.50604485
[17] -0.48780968 -0.42678960 -0.39581382 -0.36722976 -0.36553545
↪  -0.36297382 -0.25598797 -0.17997905
[25] -0.14562282 -0.14082963 -0.08120123 -0.05115865 -0.03317002
↪  0.14417172  0.16970091  0.19215454
[33]  0.21950061  0.36192037  0.44599508  0.45083054  0.53336395
↪  0.60114779  0.63483314  0.70431365
[41]  0.82939981  0.86686862  0.86870980  0.88614871  0.92285985
↪  0.96773866  0.99175241  1.47140372
[49]  1.89188206  2.15228091
> mean(amostra2);median(amostra2);var(amostra2);sd(amostra2)
[1] -0.09844016
[1] -0.1432262
[1] 0.8937156
[1] 0.9453653
> summary(amostra2)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-2.54072 -0.65627 -0.14323 -0.09844  0.58420  2.15228

# Histograma
hist(amostra2, prob = T, col = 'azure3',
     main = 'Histograma da Amostra n = 50', ylab = 'Densidade', xlim =
      ↪  c(-4, 4))
curve(f2, add = T, col = 'red')
```

```
lines(density(amostra2), col = 'blue')
#Legenda
legend("topright", legend = c('População', 'Amostra'),
       col = c('red', 'blue'), lwd = 5)
```

5.3 Códigos da Questão 3

```
> # Dados populacionais:
> m = 15
> n = 15
> Md = n/(n-2);Md;round(Md, 2) # Media
[1] 1.153846
[1] 1.15
> Mo = (n*(m-2))/(m*(n+2));Mo;round(Mo,2) # Moda
[1] 0.7647059
[1] 0.76
> Mdn = qf(0.5, 15, 15, lower.tail = TRUE); Mdn;round(Mdn, 2) # Mediana
[1] 1
[1] 1

# Definindo a função e plotando o gráfico
f3 = function(x) df(x, 15, 15)
plot(f3, 0, 5, main = 'Distribui??o F-Snedecor com m = 15 g.l. e n = 15
↪  g.l.', ylab = 'f(x)')
abline(h=0, col='black')
abline(v=Md, col = 'blue') # Media
abline(v=Mo, col = 'red') # Moda
abline(v=Mdn, col = 'green4') # Mediana
#Legenda
legend("topright", legend = c('MÃ©dia = 1,15 ', 'Moda = 0,76', 'Mediana
↪  = 1'),
       col = c('blue', 'red', 'green4'), lwd = 3)

# GRAFICO HACHURADO A
plot(f3, 0, 5, main = 'P(X < 2,40)', ylab = 'f(x)')
abline(h=0, col='black')
polygon(x = c(0, seq(0, 2.40, l=50),2.40),
```

```
        y = c(0, f3(seq(0, 2.40,l=50)), 0),
        col = 'tomato', density = 40)
abline(v=2.40, col = 'tomato')
#Legenda
legend("topright", legend = c('P(X < 2,40) = 0,95'),
       col = c('tomato'), lwd = 5)

# GRAFICO HACHURADO B
plot(f3, 0, 5, main = 'Quinto Percentil de X', ylab = 'f(x)')
abline(h=0, col='black')
abline(v=0.42, col = 'red')
polygon(x = c(0, seq(0, p5, l=50),p5),
        y = c(0, f3(seq(0, p5,l=50)), 0),
        col = 'tomato', density = 40)
#Legenda
legend("topright", legend = c('P5 = 0,42'),
       col = c('tomato'), lwd = 5)
```

```
> # Dados amostrais:
> set.seed(521353)
> amostra3 = rf(50, 15, 15);sort(amostra3)
 [1] 0.2558783 0.2605782 0.3449324 0.5564487 0.5740207 0.5877041
 ↪  0.6100806 0.6256613 0.6868603 0.7020321
[11] 0.7120911 0.7303108 0.7853953 0.7998617 0.8075718 0.8213799
↪  0.8350141 0.8397655 0.8836933 0.9268834
[21] 1.0074810 1.0676796 1.1346915 1.1786101 1.1815635 1.1854966
↪  1.2096038 1.2516273 1.2757353 1.3370329
[31] 1.3639342 1.4124286 1.4411331 1.4752905 1.4760184 1.4870153
↪  1.5089071 1.5590838 1.5993043 1.6055381
[41] 1.6386199 1.6417508 1.7879840 2.0038421 2.1301879 2.3834726
↪  2.4218501 2.4513129 2.4971395 5.6313080
> mean(amostra3);median(amostra3);var(amostra3);sd(amostra3)
[1] 1.293836
[1] 1.18353
[1] 0.7139469
[1] 0.8449538
> summary(amostra3)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
 0.2559  0.7890  1.1835  1.2938  1.5465  5.6313

# Histograma
hist(amostra3, prob = T, col = 'azure3',
     main = 'Histograma da Amostra n = 50', ylab = 'Densidade', ylim =
      ↪  c(0, 1))
curve(f3, add = T, col = 'red')
lines(density(amostra3), col = 'blue')
#Legenda
legend("topright", legend = c('População', 'Amostra'),
       col = c('red', 'blue'), lwd = 5)
```

5.4 Códigos da Questão 4

```
#GRAFICO HACHURADO A
f = function(x) dnorm(x,40,5)
par(bg = 'whitesmoke') #Cor de fundo
plot(f, 25,55, ylab = 'f(x)') # O xaxt="n" remove os valores do eixo X
abline(h=0, col = 'black')
abline(v=28, col = 'tomato')
abline(v=52, col = 'tomato')
title( main = "Densidade de X~N(40,25)",  cex.main = 1.2 )

ax = c(28,seq(28, 52, length.out = 30), 52 )
ay = c(0,lapply(seq(28, 52, length.out = 30),f),0)
polygon(ax,ay, dens=40,col='tomato')
#Legenda
legend("topright", legend = c("P(|X-40|<12)=0,9836"),
       col = c('tomato'), lwd = 5, cex = 0.8)

# GRAFICO HACHURADO AZ
plot(f4_1, -3,3, main = "P( |Z| < 2.4 )", ylab = 'f(x)')
abline(h=0, col = "black")
abline(v=-2.4, col = "red")
abline(v=2.4, col = "red")

polygon(x = c(-2.4, seq(-2.4, 0, l=50), 0),
```

```
         y = c(0, f4_1(seq(-2.4, 0,l=50)), 0),
         col = "tomato", density = 40)

polygon(x = c(0, seq(0, 2.4, l=50), 2.4),
         y = c(0, f4_1(seq(0, 2.4,l=50)), 0),
         col = "tomato", density = 40)
#Legenda
legend("topright", legend = c("P(|Z|<2,4)=0,9836"),
       col = c('tomato'), lwd = 5, cex = 0.8)

# GRAFICO HACURADO BZ
p5 = qnorm(0.05);p5;round(p5,2)
plot(f4_1, -3,3, main = 'Z~N(0,1)', ylab = 'f(x)')
abline(h=0, col = "black")
abline(v=p5, col = "tomato")
polygon(x = c(-3, seq(-3, p5, l=50), p5),
        y = c(0, f4_1(seq(-3, p5,l=50)), 0),
        col = "tomato", density = 40)
#Legenda
legend("topright", legend = c('P(Z < 0,05) = -1,64'),
       col = c('tomato'), lwd = 5, cex = 0.8)

# GRAFICO HACURADO B
p_5 = round(qnorm(0.05,40,5),3); p_5
plot(f4, 25,55, main = 'X~N(40,25)', ylab = 'f(x)')
abline(h=0, col = "black")
abline(v=p_5, col = "tomato")
polygon(x = c(25, seq(25, p_5, l=50), p_5),
        y = c(0, f4(seq(25, p_5,l=50)), 0),
        col = "tomato", density = 40)
#Legenda
legend("topright", legend = c('Percentil 5 = 31,776'),
       col = c('tomato'), lwd = 5, cex = 0.8)
```

```
> # Dados amostrais:
> set.seed(521353)
> amostra4 = rnorm(50, 40, 5);sort(amostra4)
 [1] 27.42092 30.19400 31.41218 33.31501 33.60373 34.17730 34.70986
 ↪   34.71543 34.84299 35.15000 35.25296
```

```
[12] 35.55742 36.08377 36.11533 36.32256 37.28335 37.56769 37.60322
 ↪  38.05689 38.05932 38.21501 38.47709
[23] 38.47769 38.54449 38.58357 38.93847 39.68057 39.75875 39.78096
 ↪  39.80131 39.99056 40.57064 40.60735
[34] 40.69141 40.97098 41.87143 42.50155 42.59595 43.16848 44.06603
 ↪  44.23649 44.33774 44.58312 44.84209
[45] 44.92658 47.12998 47.19723 47.47262 48.58052 49.11730
> mean(amostra4);median(amostra4);var(amostra4);sd(amostra4)
[1] 39.3432
[1] 38.76102
[1] 22.9415
[1] 4.789728
> summary(amostra4)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  27.42   36.09   38.76   39.34   42.57   49.12

# Criando o QQ plot
qqplot(qnorm(ppoints(50), mean = 40, sd = 5), amostra4, main = "Q-Q
 ↪  plot")
# Adicionando uma linha de referencia
abline(0, 1, col = "red")

# Histograma
hist(amostra4, prob = T, col = 'azure3',
     main = 'Histograma da Amostra n = 50', ylab = 'Densidade', xlim =
      ↪  c(25, 55))
curve(f4, add = T, col = 'red')
lines(density(amostra4), col = 'blue')
#Legenda
legend("topright", legend = c('Popula??o', 'Amostra'),
       col = c('red', 'blue'), lwd = 5)
```