CEDS (2024): Probabilidade e Estatística

QUIZ 3

Rennan Dalla Guimarães

2024-10-22

Índice

1 Introdução

O objetivo desse projeto é realizar o quiz 3, que busca colocar em prática aprendizados sobre: Distribuição normal, identificação de distribuições, normalidade e intervalo de confiança.

No repositótio do projeto temos a versão em html e qmd, caso esteja com dificuldade de ler alguma parte do arquivo.

2 Exercício 1: Distribuição Normal

```r
dados <- c(149.3355, 140.3779, 145.7254, 149.8931, 139.6168, 149.1934, 129.6147, 134.7523, 16
           157.5422, 160.2664, 155.4553, 142.5989, 134.9844, 148.5172, 163.1447, 131.0138, 13
           149.4015, 145.6802, 160.3472, 121.1775, 136.7295, 162.2381, 150.7192, 117.8144, 13
           168.0833, 133.9263, 150.9102, 149.4811, 167.4367, 178.0970, 138.4903, 148.6764, 18
           147.0679, 156.1410, 148.8734, 140.9484, 147.6408, 134.5726, 184.6812, 134.6648, 14
```

2.1 a) Realização dos testes de normalidade

i) Teste de Kolmogorov-Smirnov

```r
ks.test(dados, "pnorm", mean=mean(dados), sd=sd(dados))
```

```
    Exact one-sample Kolmogorov-Smirnov test

data:  dados
D = 0.1167, p-value = 0.4688
alternative hypothesis: two-sided
```

Interpretação:

O teste de Kolmogorov-Smirnov compara a distribuição acumulada empírica dos dados com a distribuição normal teórica. O p-valor obtido é usado para testar a hipótese nula de que os dados seguem uma distribuição normal. Se o p-valor for menor que 0,05, rejeitamos a hipótese de normalidade.

ii) Teste de Shapiro-Wilk

```r
shapiro.test(dados)
```

```
    Shapiro-Wilk normality test

data:  dados
W = 0.98185, p-value = 0.6324
```

Interpretação:

O teste de Shapiro-Wilk é utilizado para verificar a normalidade dos dados. Um p-valor maior que 0,05 indica que não podemos rejeitar a hipótese de normalidade.

iii) Teste de Anderson-Darling

```r
library(nortest)
ad.test(dados)
```

```
	Anderson-Darling normality test

data:  dados
A = 0.37902, p-value = 0.3928
```

Interpretação:

O teste de Anderson-Darling é sensível a discrepâncias na cauda da distribuição. Um p-valor maior que 0,05 indica que os dados podem ser considerados normais.

iv) Teste de Lilliefors

```r
lillie.test(dados)
```

```
	Lilliefors (Kolmogorov-Smirnov) normality test

data:  dados
D = 0.1167, p-value = 0.08619
```

Interpretação:

O teste de Lilliefors é uma adaptação do teste de Kolmogorov-Smirnov quando os parâmetros da distribuição normal não são conhecidos e precisam ser estimados. Um p-valor maior que 0,05 indica normalidade.

Conclusão Geral:

Se em todos os testes o p-valor for maior que 0,05, não rejeitamos a hipótese de que os dados seguem uma distribuição normal. Portanto, podemos considerar os dados como normalmente distribuídos.

2.2 b) Probabilidade de que uma chamada demore entre 125 e 150 segundos

Calculando a média e o desvio padrão:

```
media <- mean(dados)
desvio <- sd(dados)
media
```

```
[1] 150.0301
```

```
desvio
```

```
[1] 15.059
```

Calculando a probabilidade:

```
prob_b <- pnorm(150, mean=media, sd=desvio) - pnorm(125, mean=media, sd=desvio)
prob_b
```

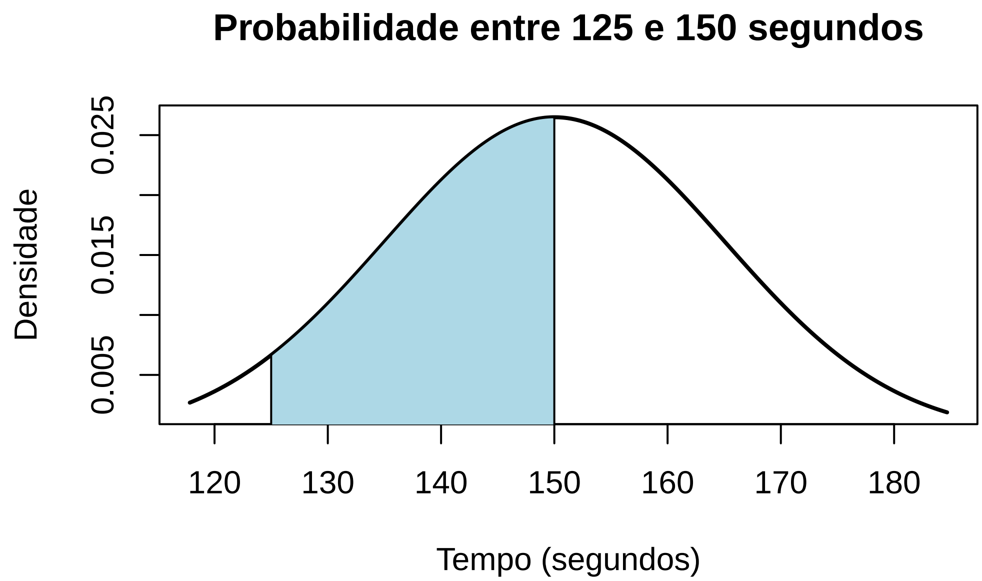
```
[1] 0.4509603
```

Interpretação:

A probabilidade de uma chamada demorar entre 125 e 150 segundos é aproximadamente 45.1%.

Gráfico:

```
x <- seq(min(dados), max(dados), length=1000)
y <- dnorm(x, mean=media, sd=desvio)
plot(x, y, type="l", lwd=2, ylab="Densidade", xlab="Tempo (segundos)", main="Probabilidade en
polygon(c(125, seq(125, 150, length=100), 150), c(0, dnorm(seq(125, 150, length=100), mean=me
```



2.3 c) Probabilidade de que uma chamada demore menos de 125 segundos

```
prob_c <- pnorm(125, mean=media, sd=desvio)
prob_c
```

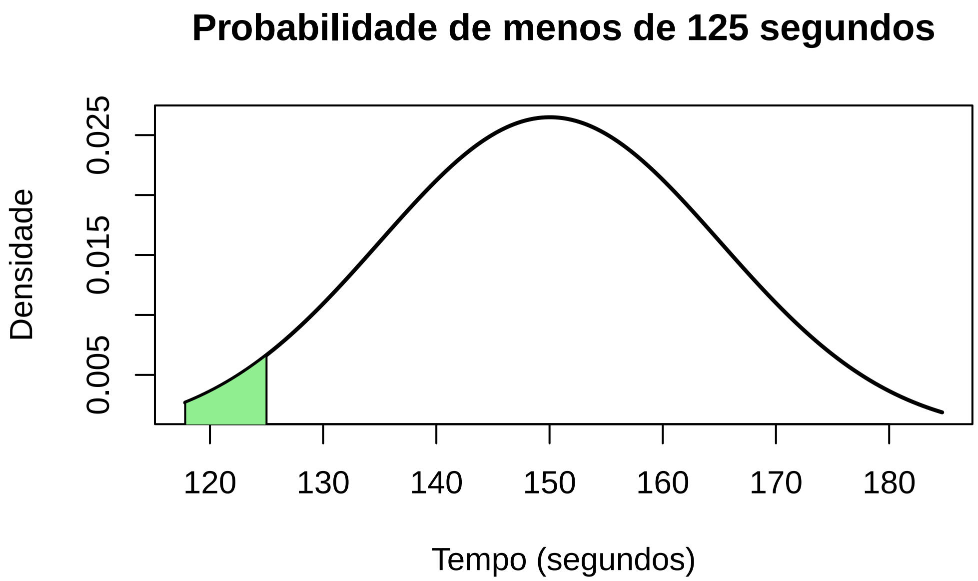
```
[1] 0.04824295
```

Interpretação:

A probabilidade de uma chamada demorar menos de 125 segundos é aproximadamente 4.82%.

Gráfico:

```
plot(x, y, type="l", lwd=2, ylab="Densidade", xlab="Tempo (segundos)", main="Probabilidade de
polygon(c(min(x), seq(min(x), 125, length=100), 125), c(0, dnorm(seq(min(x), 125, length=100)
```



2.4 d) Probabilidade de que uma chamada demore entre 145 e 155 segundos

```
prob_d <- pnorm(155, mean=media, sd=desvio) - pnorm(145, mean=media, sd=desvio)
prob_d
```

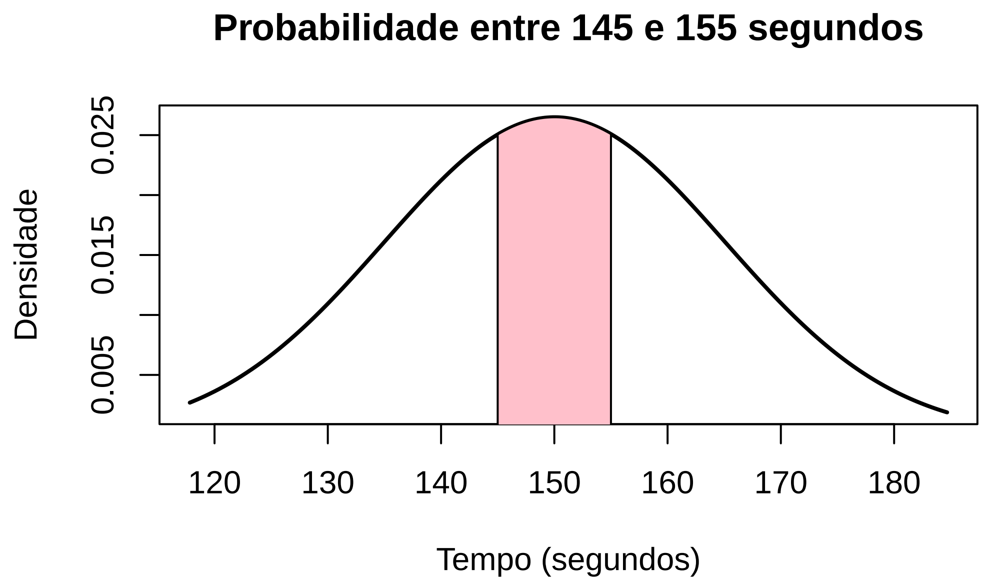
```
[1] 0.2601309
```

Interpretação:

A probabilidade de uma chamada demorar entre 145 e 155 segundos é aproximadamente 26.01%.

Gráfico:

```
plot(x, y, type="l", lwd=2, ylab="Densidade", xlab="Tempo (segundos)", main="Probabilidade en
polygon(c(145, seq(145, 155, length=100), 155), c(0, dnorm(seq(145, 155, length=100), mean=me
```



2.5 e) Probabilidade de que uma chamada demore entre 160 e 165 segundos

```
prob_e <- pnorm(165, mean=media, sd=desvio) - pnorm(160, mean=media, sd=desvio)
prob_e
```

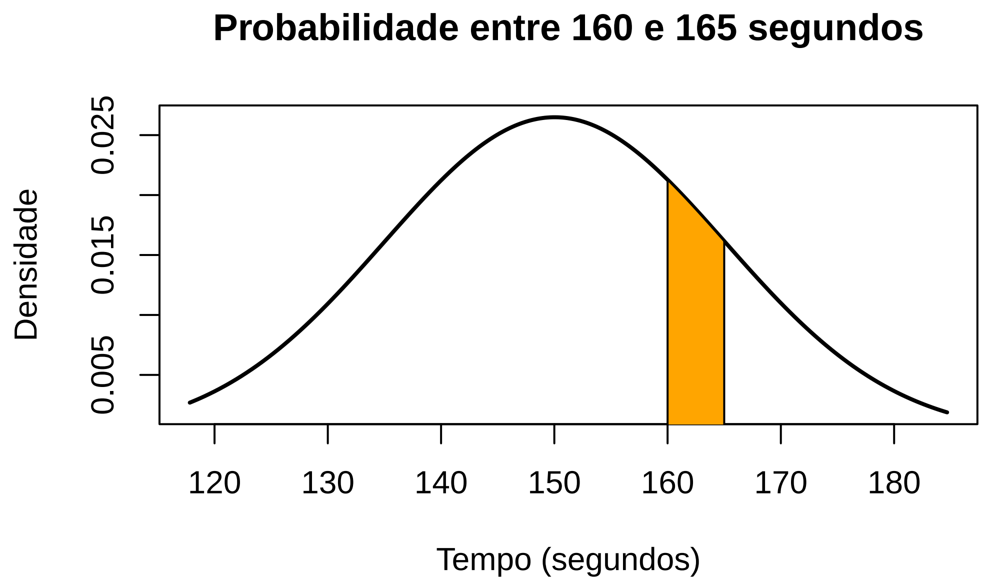
```
[1] 0.09387641
```

Interpretação:

A probabilidade de uma chamada demorar entre 160 e 165 segundos é aproximadamente 9.39%.

Gráfico:

```
plot(x, y, type="l", lwd=2, ylab="Densidade", xlab="Tempo (segundos)", main="Probabilidade er
polygon(c(160, seq(160, 165, length=100), 165), c(0, dnorm(seq(160, 165, length=100), mean=me
```



---

3 Exercício 2: Identificação de Distribuição

Dados da variável aleatória X:

```r
dados <- c(1.9993382, 1.4414849, 2.1477166, 2.1087828, 2.1342892, 2.1844835, 1.5091879, 2.04
           2.1302612, 1.8389897, 1.8924614, 1.9316041, 1.5602204, 1.6991884, 1.7228081, 1.51
           0.6914335, 1.4598759, 2.0017607, 1.5139209, 1.8334780, 1.8847480, 1.9072389, 1.62
           1.7744973, 2.4300455, 1.8958270)
```

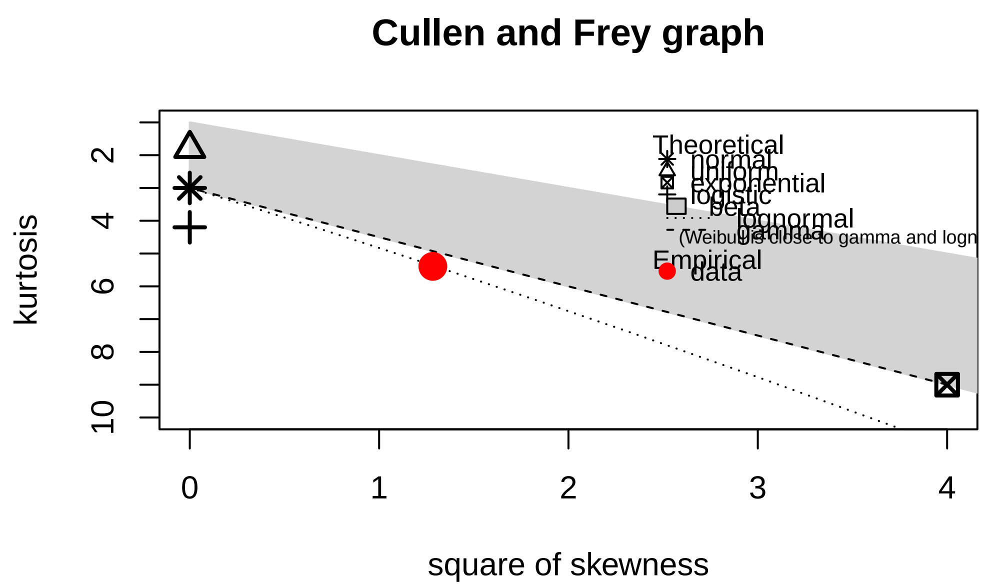
3.1 a) Identificação da distribuição

Utilizando o pacote `fitdistrplus`:

```r
library(fitdistrplus)
```

Análise descritiva:

```r
descdist(dados, discrete = FALSE)
```




```
summary statistics
------
min:  0.6914335   max:  2.430045
median:  1.861869
mean:  1.787556
estimated sd:  0.3498879
estimated skewness:  -1.133072
estimated kurtosis:  5.391445
```

Ajuste das distribuições:

- **Weibull**

```r
ajuste_weibull <- fitdist(dados, "weibull")
```

- **Gamma**

```r
ajuste_gamma <- fitdist(dados, "gamma")
```

- **Lognormal**

```r
ajuste_lognormal <- fitdist(dados, "lnorm")
```

Comparação dos AIC:

```r
aic_values <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal"),
  AIC = c(ajuste_weibull$aic, ajuste_gamma$aic, ajuste_lognormal$aic)
)
aic_values
```

```
  Distribuição      AIC
1      Weibull 21.97718
2        Gamma 31.70738
3    Lognormal 36.10330
```

3.2 b) Comparação dos resultados do teste de Kolmogorov-Smirnov

Weibull

```r
ks.test(dados, "pweibull", shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$estim
```

```

	Exact one-sample Kolmogorov-Smirnov test

data:  dados
D = 0.091733, p-value = 0.9424
alternative hypothesis: two-sided
```

Gamma

```r
ks.test(dados, "pgamma", shape=ajuste_gamma$estimate["shape"], rate=ajuste_gamma$estimate["ra
```

```

	Exact one-sample Kolmogorov-Smirnov test

data:  dados
D = 0.14211, p-value = 0.5332
alternative hypothesis: two-sided
```

Lognormal

```
ks.test(dados, "plnorm", meanlog=ajuste_lognormal$estimate["meanlog"], sdlog=ajuste_lognormal
```

```

    Exact one-sample Kolmogorov-Smirnov test

data:  dados
D = 0.15513, p-value = 0.423
alternative hypothesis: two-sided
```

Justificativa:

Analisando os resultados apresentados, a distribuição Weibull apresenta o melhor ajuste aos dados por duas razões principais:

Menor valor de AIC (Critério de Informação de Akaike):

Weibull: 21.97718 Gamma: 31.70738 Lognormal: 36.10330

O AIC mais baixo da Weibull indica que esta distribuição oferece o melhor compromisso entre a qualidade do ajuste e a complexidade do modelo.

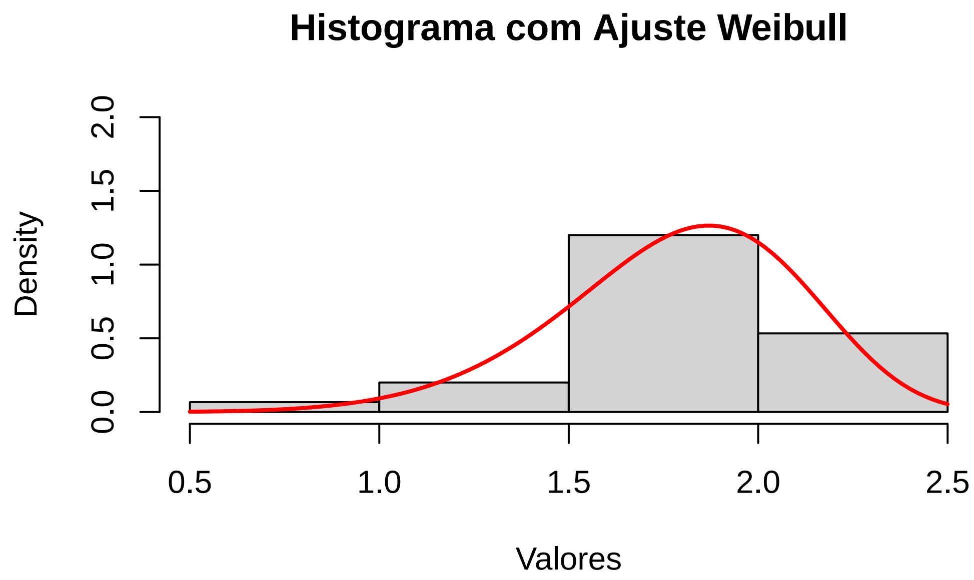
Maior p-valor no teste de Kolmogorov-Smirnov: Weibull: p-valor = 0.9424 Gamma: p-valor = 0.5332 Lognormal: p-valor = 0.423

O p-valor mais alto da Weibull (0.9424) indica que não há evidências para rejeitar a hipótese de que os dados seguem esta distribuição. Quanto maior o p-valor, mais forte é a evidência de que o modelo se ajusta bem aos dados. Em comparação com as outras distribuições, a Weibull apresenta tanto o menor AIC quanto o maior p-valor, o que a torna claramente a melhor escolha para modelar estes dados.

3.3 c) Plotar a função e o histograma para a distribuição escolhida

Histograma com ajuste da Weibull:

```
hist(dados, freq=FALSE, main="Histograma com Ajuste Weibull", xlab="Valores", ylim=c(0, 2))
curve(dweibull(x, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$estimate["sca
      col="red", lwd=2, add=TRUE)
```



3.4 d) Verificar se a área sob a curva estimada é igual a 1

Calculando a integral da função densidade de probabilidade:

```
integrate(function(x) dweibull(x, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibul
          lower=0, upper=Inf)
```

```
1 with absolute error < 3e-08
```

Interpretação:

O resultado deve ser próximo de 1, confirmando que a área sob a curva da distribuição de probabilidade é igual a 1.

3.5 e) Calcular a área no intervalo [1; 1,5] e plotar

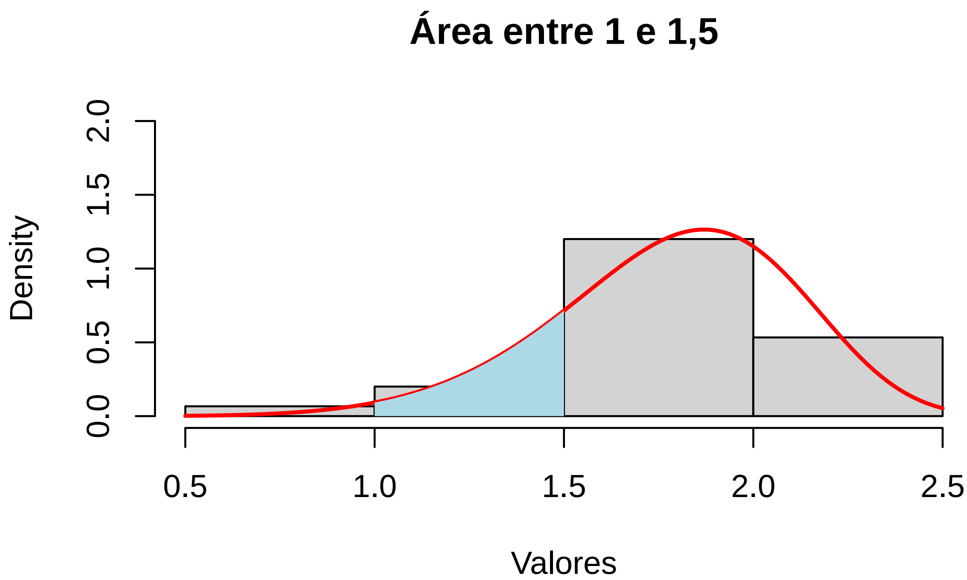
Calculando a probabilidade no intervalo:

```
prob_intervalo <- pweibull(1.5, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$
                  pweibull(1, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$es
prob_intervalo
```

```
[1] 0.1681584
```

Plotando a área:

```r
hist(dados, freq=FALSE, main="Área entre 1 e 1,5", xlab="Valores", ylim=c(0, 2))
curve(dweibull(x, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$estimate["sca
      col="red", lwd=2, add=TRUE)
x_seq <- seq(1, 1.5, length=100)
y_seq <- dweibull(x_seq, shape=ajuste_weibull$estimate["shape"], scale=ajuste_weibull$estima
polygon(c(1, x_seq, 1.5), c(0, y_seq, 0), col="lightblue", border=NA)
```



---

4 Exercício 3: Normalidade e Intervalo de Confiança

Dados de inflação anual (2013 a 2022):

```r
inflacao <- c(5.91, 6.41, 10.67, 6.29, 2.95, 3.75, 4.31, 4.52, 10.06, 5.79)
```

4.1 a) Testes de Shapiro-Wilk e Lilliefors

Shapiro-Wilk

```
shapiro.test(inflacao)
```

```

	Shapiro-Wilk normality test

data:  inflacao
W = 0.88867, p-value = 0.1638
```

Lilliefors

```
lillie.test(inflacao)
```

```

	Lilliefors (Kolmogorov-Smirnov) normality test

data:  inflacao
D = 0.24609, p-value = 0.08736
```

Conclusão:

Se os p-valores forem maiores que 0,05, não rejeitamos a hipótese de normalidade.

4.2 b) Intervalo de confiança de 99% para a média da inflação

Calculando a média e o desvio padrão:

```
media_inf <- mean(inflacao)
desvio_inf <- sd(inflacao)
n <- length(inflacao)
erro_padrao <- desvio_inf / sqrt(n)
```

Calculando o intervalo:

```
t_critico <- qt(0.995, df=n-1)
limite_inferior <- media_inf - t_critico * erro_padrao
limite_superior <- media_inf + t_critico * erro_padrao
c(limite_inferior, limite_superior)
```

```
[1] 3.457911 8.674089
```

4.3 c) Nível de confiança para intervalo com comprimento total igual a 3

Queremos que o comprimento total seja 3, logo o erro máximo é 1,5.

Calculando o t crítico necessário:

```
erro_max <- 1.5
t_necessario <- erro_max / erro_padrao
```

Calculando o nível de confiança correspondente:

```
nivel_conf <- 2*(1 - pt(t_necessario, df=n-1))
nivel_conf
```

```
[1] 0.09443553
```

4.4 d) Intervalo de confiança de 90% para o desvio padrão

Usando a distribuição qui-quadrado:

```
alfa <- 0.10
chi2_inferior <- qchisq(alfa/2, df=n-1)
chi2_superior <- qchisq(1 - alfa/2, df=n-1)
limite_inferior <- sqrt((n-1)*desvio_inf^2 / chi2_superior)
limite_superior <- sqrt((n-1)*desvio_inf^2 / chi2_inferior)
c(limite_inferior, limite_superior)
```

```
[1] 1.850952 4.175218
```

4.5 e) Teste de normalidade para a série histórica de 1999 a 2022

Dados da inflação de 1999 a 2022:

```
inflacao_historica <- c(8.94, 6, 12.53, 7.67, 12.53, 7.6, 7.6, 5.69, 3.14, 4.46, 5.9, 4.31,
                        5.91, 6.41, 10.67, 6.29, 2.95, 3.75, 4.31, 4.52, 10.06, 5.79, 5.79,
```

Testes de normalidade:

Shapiro-Wilk

```r
shapiro.test(inflacao_historica)
```

```

	Shapiro-Wilk normality test

data:  inflacao_historica
W = 0.9045, p-value = 0.02685
```

Lilliefors

```r
lillie.test(inflacao_historica)
```

```

	Lilliefors (Kolmogorov-Smirnov) normality test

data:  inflacao_historica
D = 0.19628, p-value = 0.0175
```

Conclusão:

Baseado nos p-valores, determinar se a série histórica pode ser considerada normalmente distribuída.

---

5 Exercício 4: Identificação de Distribuição

Neste exercício, iremos identificar a distribuição adequada para cada conjunto de dados fornecido, utilizando técnicas estatísticas e ferramentas do R.

5.1 Preparação do Ambiente

5.2 a) Conjunto de dados (a)

5.2.1 Dados

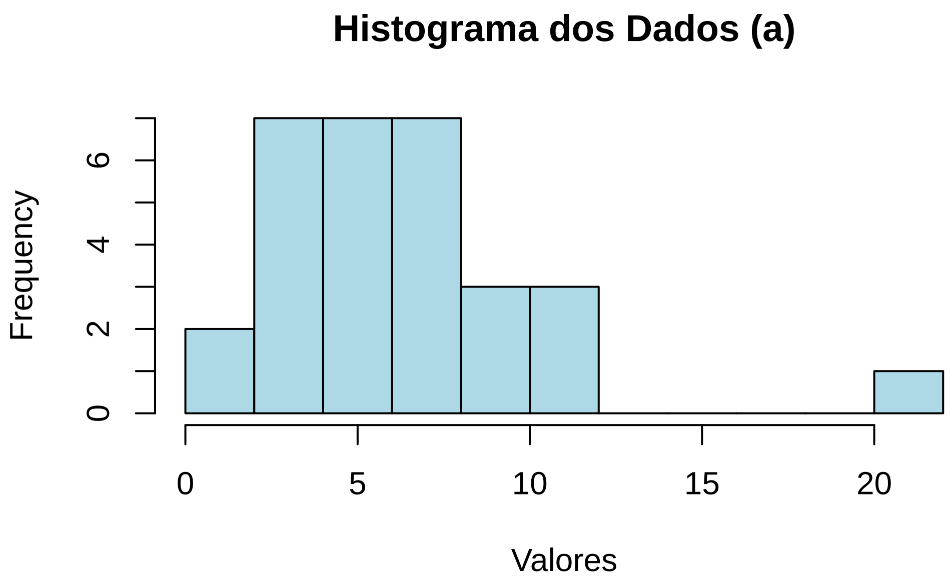
```r
dados_a <- c(20.8625807,  7.2445709,  4.4659396,  3.2712081,  4.9300651,  5.7444213,  6.6700
             11.1750446,  2.3753017,  3.5425386,  0.5978486,  6.8869953,  6.1102197,  8.2716
             9.7465462,  3.3991988,  1.8557047, 11.3983705,  3.6847590,  2.3327479,  6.13643
             4.4686122,  7.8007834,  4.7649257,  3.8829371,  5.9986131,  5.5163819,  9.69517
             10.1645820,  6.1304865)
```

5.2.2 Análise Descritiva

```r
summary(dados_a)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 0.5978  3.7343  5.8715  6.3042  7.6617 20.8626 
```

```r
hist(dados_a, breaks=10, col="lightblue", main="Histograma dos Dados (a)", xlab="Valores")
```



5.2.3 Ajuste das Distribuições

Realizamos o ajuste para as distribuições Weibull, Gamma, Lognormal e Normal.

```r
ajuste_weibull_a <- fitdist(dados_a, "weibull")
ajuste_gamma_a <- fitdist(dados_a, "gamma")
ajuste_lognormal_a <- fitdist(dados_a, "lnorm")
ajuste_normal_a <- fitdist(dados_a, "norm")
```

5.2.4 Comparação dos Critérios de Informação (AIC)

```r
aic_values_a <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal", "Normal"),
  AIC = c(ajuste_weibull_a$aic, ajuste_gamma_a$aic, ajuste_lognormal_a$aic, ajuste_normal_a$a
)
aic_values_a
```

```
  Distribuição      AIC
1      Weibull 161.6723
2        Gamma 160.3733
3    Lognormal 163.2588
4       Normal 169.7731
```

Interpretação:

A distribuição com o menor AIC é a que melhor se ajusta aos dados. Observamos que a distribuição Gamma apresenta o menor AIC.

5.2.5 Teste de Kolmogorov-Smirnov

Testamos a aderência dos dados à distribuição Gamma.

```r
ks.test(dados_a, "pgamma", shape=ajuste_gamma_a$estimate["shape"], rate=ajuste_gamma_a$estima
```

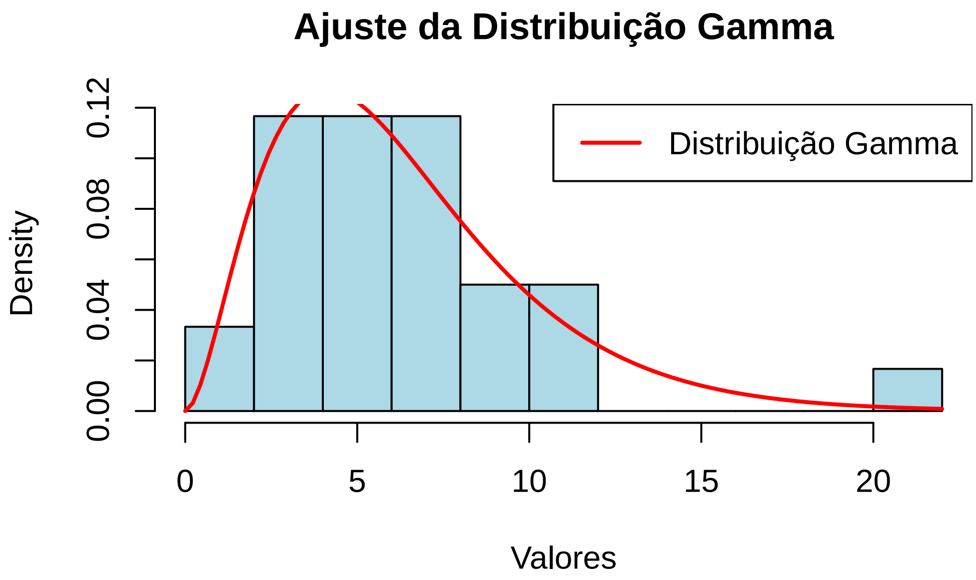
```

    Exact one-sample Kolmogorov-Smirnov test

data:  dados_a
D = 0.079856, p-value = 0.9828
alternative hypothesis: two-sided
```

Interpretação:

O p-valor do teste de Kolmogorov-Smirnov é maior que 0,05, não rejeitando a hipótese nula de que os dados seguem uma distribuição Gamma.

5.2.6 Plotagem do Histograma com a Curva Ajustada

```
hist(dados_a, breaks=10, freq=FALSE, col="lightblue", main="Ajuste da Distribuição Gamma", xl
curve(dgamma(x, shape=ajuste_gamma_a$estimate["shape"], rate=ajuste_gamma_a$estimate["rate"])
      col="red", lwd=2, add=TRUE)
legend("topright", legend="Distribuição Gamma", col="red", lwd=2)
```



5.3 b) Conjunto de dados (b)

5.3.1 Dados

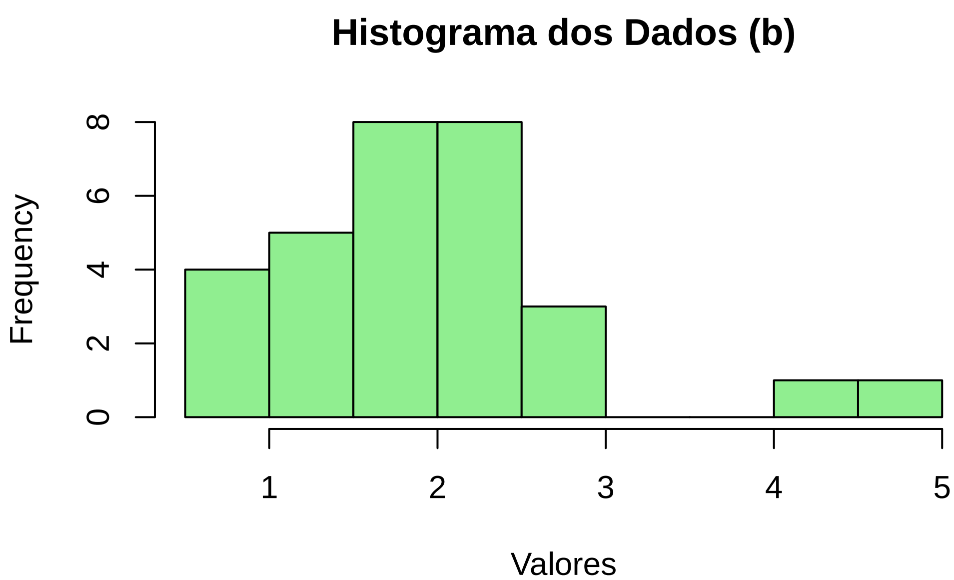
```
dados_b <- c(1.4940354, 2.0164275, 1.9513521, 1.5298282, 0.6815670, 2.4267801, 0.6762800, 1.7
             4.1632638, 2.5472784, 2.2174151, 0.6058986, 1.7432601, 1.1199216, 1.7135932, 2.8
             0.8537880, 1.5511504, 2.3262178, 2.3267933, 1.3916375, 4.7439947, 2.1864812, 2.0
             1.7489244, 1.8191036, 2.0845146, 1.2229195, 1.0115042, 2.7931222)
```

5.3.2 Análise Descritiva

```r
summary(dados_b)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 0.6059  1.4172  1.7840  1.9184  2.2990  4.7440 
```

```r
hist(dados_b, breaks=10, col="lightgreen", main="Histograma dos Dados (b)", xlab="Valores")
```



5.3.3 Ajuste das Distribuições

```r
ajuste_weibull_b <- fitdist(dados_b, "weibull")
ajuste_gamma_b <- fitdist(dados_b, "gamma")
ajuste_lognormal_b <- fitdist(dados_b, "lnorm")
ajuste_normal_b <- fitdist(dados_b, "norm")
```

5.3.4 Comparação dos Critérios de Informação (AIC)

```
aic_values_b <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal", "Normal"),
  AIC = c(ajuste_weibull_b$aic, ajuste_gamma_b$aic, ajuste_lognormal_b$aic, ajuste_normal_b$a
)
aic_values_b
```

```
  Distribuição      AIC
1      Weibull 79.57745
2        Gamma 77.32663
3    Lognormal 77.84694
4       Normal 83.27721
```

Interpretação:

A distribuição Gamma apresenta o menor AIC (77.32663), indicando o melhor ajuste aos dados.

5.3.5 Teste de Kolmogorov-Smirnov

```
ks.test(dados_b, "pgamma", shape=ajuste_gamma_b$estimate["shape"], rate=ajuste_gamma_b$estima
```

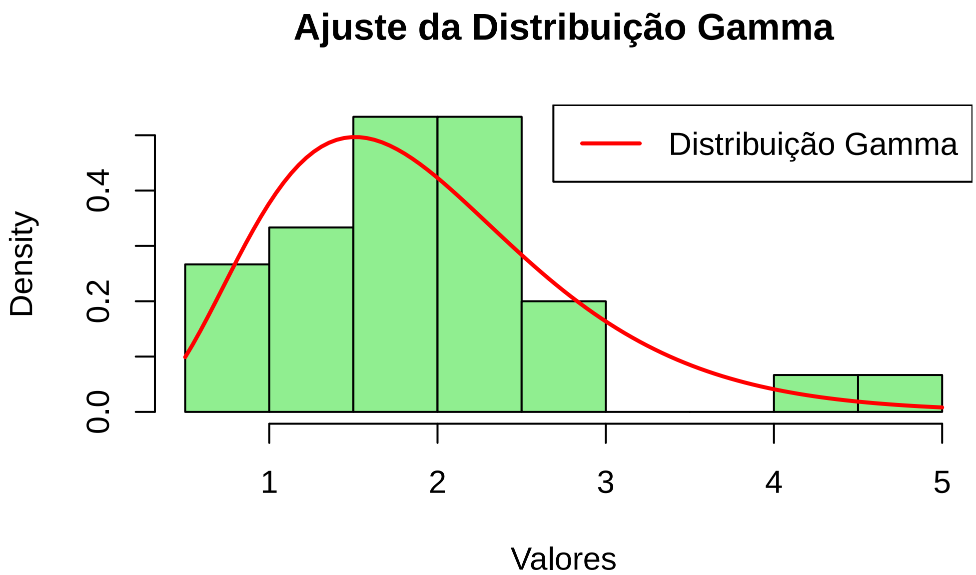
```

	Exact one-sample Kolmogorov-Smirnov test

data:  dados_b
D = 0.094142, p-value = 0.9304
alternative hypothesis: two-sided
```

Interpretação:

- O p-valor é maior que 0,05, não rejeitando a hipótese de que os dados seguem uma distribuição Gamma.
- Portanto, para o conjunto de dados (b), a distribuição Gamma é a que melhor se ajusta.

5.3.6 Plotagem do Histograma com a Curva Ajustada

```
hist(dados_b, breaks=10, freq=FALSE, col="lightgreen", main="Ajuste da Distribuição Gamma", x
curve(dgamma(x, shape=ajuste_gamma_b$estimate["shape"], rate=ajuste_gamma_b$estimate["rate"])
      col="red", lwd=2, add=TRUE)
legend("topright", legend="Distribuição Gamma", col="red", lwd=2)
```



5.4 c) Conjunto de dados (c)

5.4.1 Dados

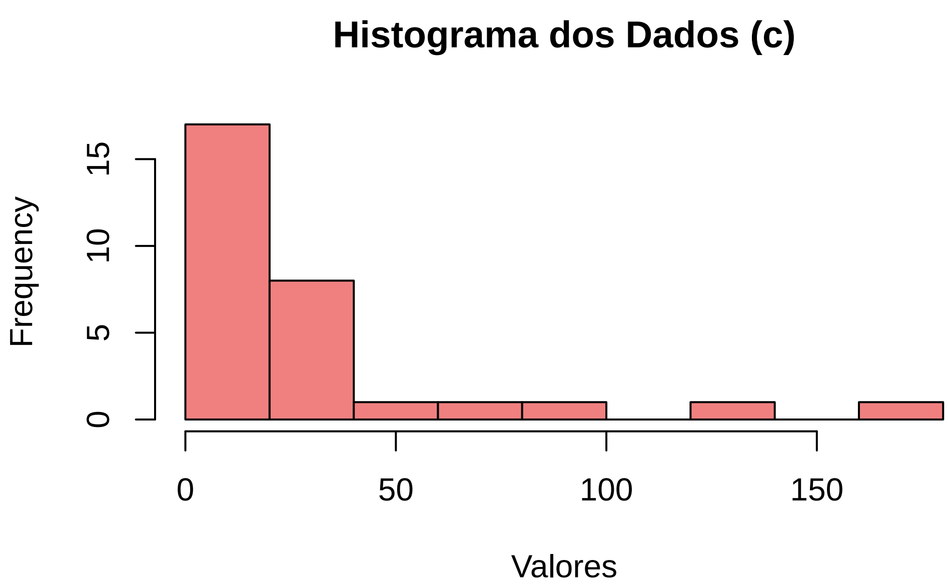
```
dados_c <- c(9.534149, 12.878719, 35.635908, 39.158389, 10.091099, 133.714299, 15.684000, 3.1
             16.073085, 57.767201, 29.543033, 24.672685, 11.955565, 2.132028, 17.455254, 20.5
             6.293823, 22.717485, 83.353863, 18.544482, 66.437399, 4.616951, 18.931367, 1.464
             21.180916, 179.315876, 24.941790, 14.105447, 7.680880, 17.688369)
```

5.4.2 Análise Descritiva

```
summary(dados_c)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  1.464  10.557  18.116  30.911  28.393 179.316 
```

```
hist(dados_c, breaks=10, col="lightcoral", main="Histograma dos Dados (c)", xlab="Valores")
```



5.4.3 Ajuste das Distribuições

```
ajuste_weibull_c <- fitdist(dados_c, "weibull")
ajuste_gamma_c <- fitdist(dados_c, "gamma")
ajuste_lognormal_c <- fitdist(dados_c, "lnorm")
ajuste_normal_c <- fitdist(dados_c, "norm")
```

5.4.4 Comparação dos Critérios de Informação (AIC)

```
aic_values_c <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal", "Normal"),
  AIC = c(ajuste_weibull_c$aic, ajuste_gamma_c$aic, ajuste_lognormal_c$aic, ajuste_normal_c$a
)
aic_values_c
```

```
  Distribuição      AIC
1      Weibull 269.7264
2        Gamma 269.8562
3    Lognormal 265.9106
4       Normal 308.5494
```

Interpretação:

A distribuição Lognormal apresenta o menor AIC, indicando o melhor ajuste.

5.4.5 Teste de Kolmogorov-Smirnov

```
ks.test(dados_c, "plnorm", meanlog=ajuste_lognormal_c$estimate["meanlog"], sdlog=ajuste_logno
```

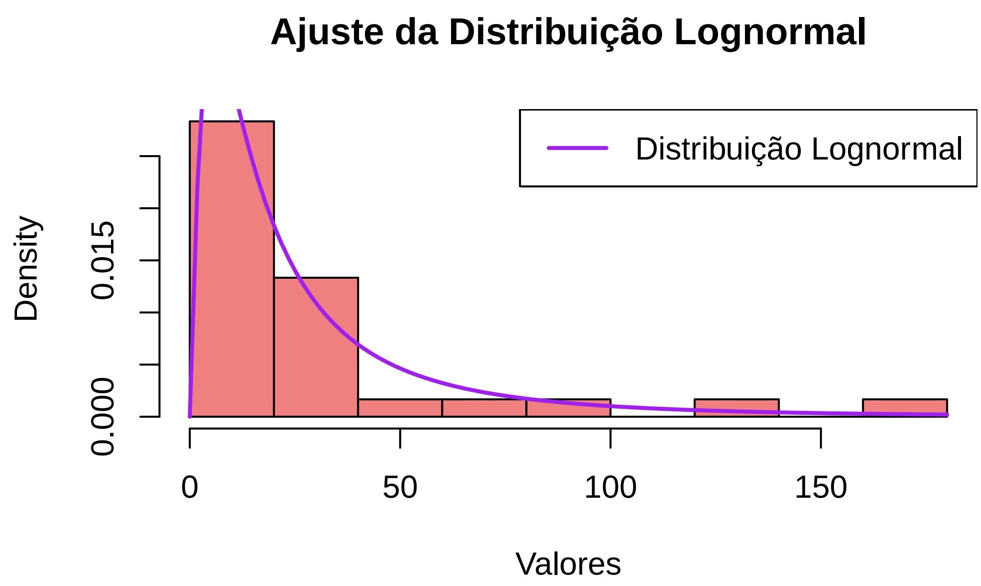
```
    Exact one-sample Kolmogorov-Smirnov test

data:  dados_c
D = 0.10729, p-value = 0.8441
alternative hypothesis: two-sided
```

Interpretação:

O p-valor é maior que 0,05, não rejeitando a hipótese de aderência à distribuição Lognormal.

5.4.6 Plotagem do Histograma com a Curva Ajustada

```
hist(dados_c, breaks=10, freq=FALSE, col="lightcoral", main="Ajuste da Distribuição Lognormal
curve(dlnorm(x, meanlog=ajuste_lognormal_c$estimate["meanlog"], sdlog=ajuste_lognormal_c$esti
      col="purple", lwd=2, add=TRUE)
legend("topright", legend="Distribuição Lognormal", col="purple", lwd=2)
```



5.5 d) Conjunto de dados (d)

5.5.1 Dados

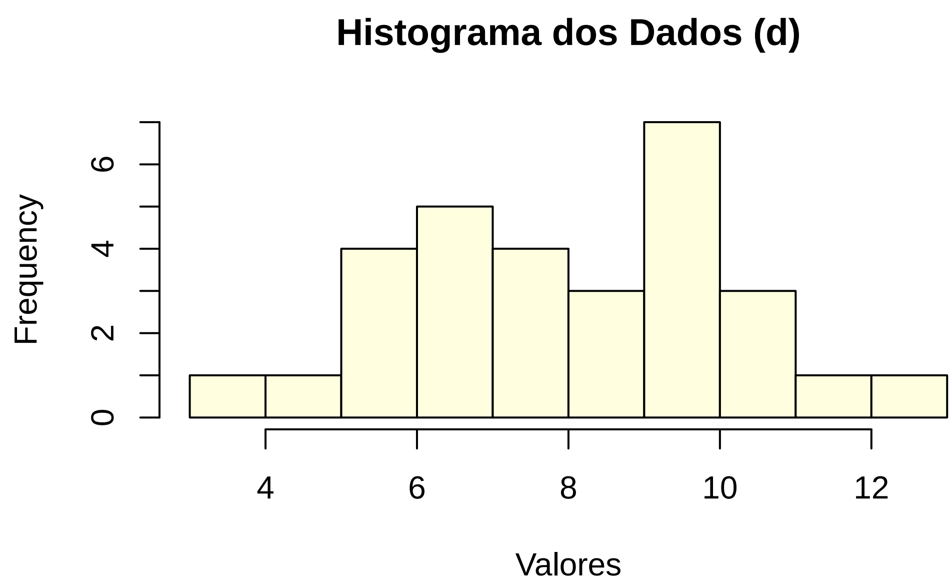
```
dados_d <- c(4.391658,  5.364267, 10.707930,  5.431008,  6.904122,  6.960462, 12.741468,  8.0
             8.434530,  9.747057,  6.440681,  7.623020,  9.276933,  8.711818,  5.250229,  6.4
             9.717008,  9.317296,  9.011653, 11.758927, 10.844472,  9.644711,  7.541715,  7.5
             9.654606,  6.222452,  5.207637)
```

5.5.2 Análise Descritiva

```r
summary(dados_d)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  3.478   6.451   7.859   7.994   9.652  12.741 
```

```r
hist(dados_d, breaks=10, col="lightyellow", main="Histograma dos Dados (d)", xlab="Valores")
```



5.5.3 Ajuste das Distribuições

```r
ajuste_weibull_d <- fitdist(dados_d, "weibull")
ajuste_gamma_d <- fitdist(dados_d, "gamma")
ajuste_lognormal_d <- fitdist(dados_d, "lnorm")
ajuste_normal_d <- fitdist(dados_d, "norm")
```

5.5.4 Comparação dos Critérios de Informação (AIC)

```
aic_values_d <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal", "Normal"),
  AIC = c(ajuste_weibull_d$aic, ajuste_gamma_d$aic, ajuste_lognormal_d$aic, ajuste_normal_d$a
)
aic_values_d
```

```
  Distribuição      AIC
1      Weibull 136.3711
2        Gamma 137.5245
3    Lognormal 139.0259
4       Normal 136.6819
```

Interpretação:

A distribuição Weibull apresenta o menor AIC (136.37110), indicando o melhor ajuste aos dados.

5.5.5 Teste de Kolmogorov-Smirnov para a Distribuição Weibull

```
ks.test(dados_d, "pweibull", shape=ajuste_weibull_d$estimate["shape"], scale=ajuste_weibull_d
```

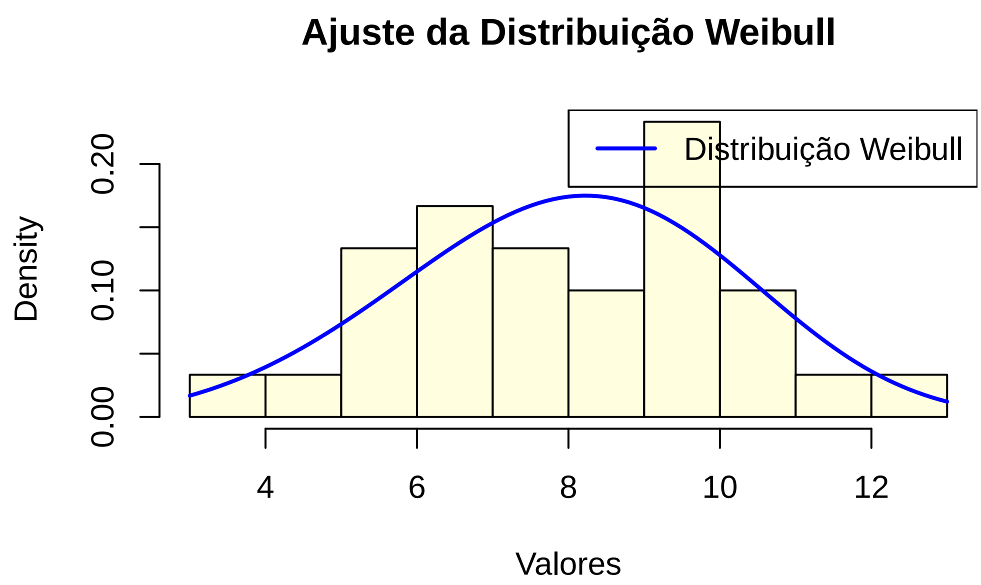
```
	Exact one-sample Kolmogorov-Smirnov test

data:  dados_d
D = 0.074926, p-value = 0.9913
alternative hypothesis: two-sided
```

Interpretação:

- O p-valor é maior que 0,05, não rejeitando a hipótese de que os dados seguem uma distribuição Weibull.
- Portanto, para o conjunto de dados (d), a distribuição Weibull é a que melhor se ajusta.

5.5.6 Plotagem do Histograma com a Curva Ajustada

```r
hist(dados_d, breaks=10, freq=FALSE, col="lightyellow", main="Ajuste da Distribuição Weibull"
curve(dweibull(x, shape=ajuste_weibull_d$estimate["shape"], scale=ajuste_weibull_d$estimate["
      col="blue", lwd=2, add=TRUE)
legend("topright", legend="Distribuição Weibull", col="blue", lwd=2)
```



5.6 e) Conjunto de dados (e)

5.6.1 Dados

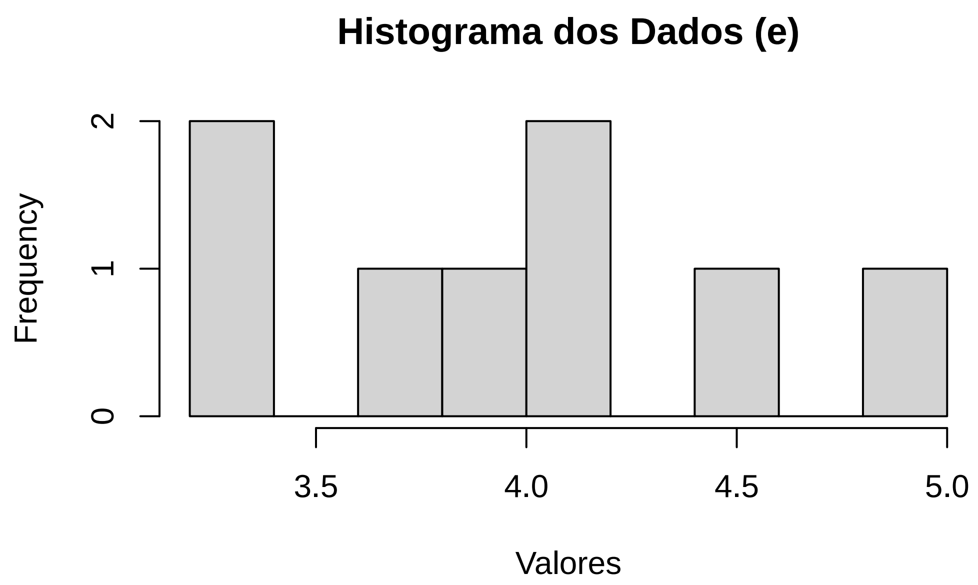
```r
dados_e <- c(3.816942, 4.123619, 4.575150, 3.214129, 4.854917, 3.647232, 4.003734, 3.261923)
```

5.6.2 Análise Descritiva

```r
summary(dados_e)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  3.214   3.551   3.910   3.937   4.237   4.855 
```

```
hist(dados_e, breaks=10, col="lightgray", main="Histograma dos Dados (e)", xlab="Valores")
```



5.6.3 Ajuste das Distribuições

```
ajuste_weibull_e <- fitdist(dados_e, "weibull")
ajuste_gamma_e <- fitdist(dados_e, "gamma")
ajuste_lognormal_e <- fitdist(dados_e, "lnorm")
ajuste_normal_e <- fitdist(dados_e, "norm")
```

5.6.4 Comparação dos Critérios de Informação (AIC)

```
aic_values_e <- data.frame(
  Distribuição = c("Weibull", "Gamma", "Lognormal", "Normal"),
  AIC = c(ajuste_weibull_e$aic, ajuste_gamma_e$aic, ajuste_lognormal_e$aic, ajuste_normal_e$a
)
aic_values_e
```

```
  Distribuição      AIC
1      Weibull 17.48481
2        Gamma 16.78578
3    Lognormal 16.74810
4       Normal 16.95620
```

Interpretação:

A distribuição Lognormal apresenta o menor AIC (16.74810), indicando o melhor ajuste aos dados.

5.6.5 Teste de Kolmogorov-Smirnov para a Distribuição Lognormal

```
ks.test(dados_e, "plnorm", meanlog=ajuste_lognormal_e$estimate["meanlog"], sdlog=ajuste_logno
```

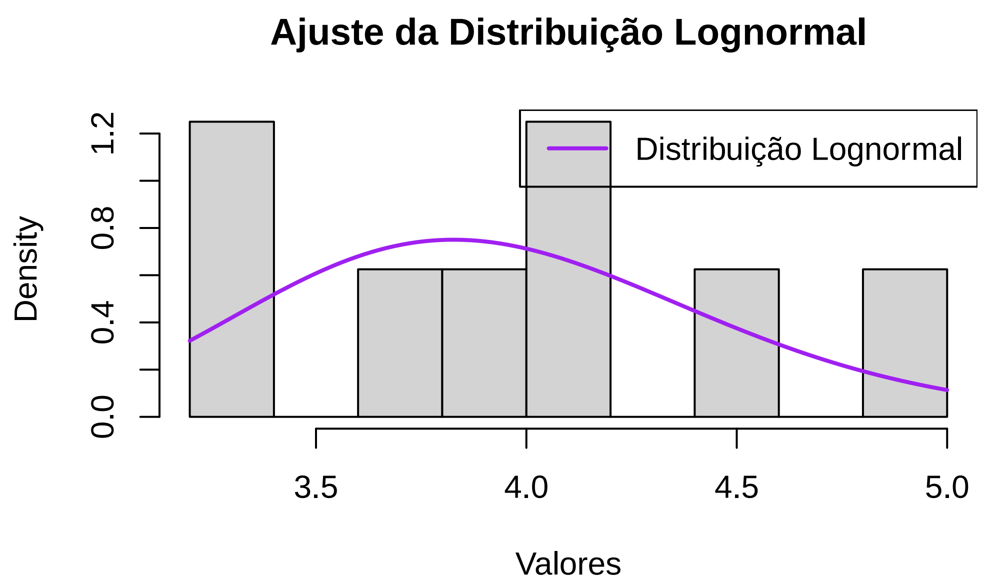
```
    Exact one-sample Kolmogorov-Smirnov test

data:  dados_e
D = 0.15288, p-value = 0.9774
alternative hypothesis: two-sided
```

Interpretação:

- O p-valor é maior que 0,05, não rejeitando a hipótese de que os dados seguem uma distribuição Lognormal.
- Portanto, para o conjunto de dados (e), a distribuição Lognormal é a que melhor se ajusta.

5.6.6 Plotagem do Histograma com a Curva Ajustada

```
hist(dados_e, breaks=10, freq=FALSE, col="lightgray", main="Ajuste da Distribuição Lognormal"
curve(dlnorm(x, meanlog=ajuste_lognormal_e$estimate["meanlog"], sdlog=ajuste_lognormal_e$esti
      col="purple", lwd=2, add=TRUE)
legend("topright", legend="Distribuição Lognormal", col="purple", lwd=2)
```


Ajuste da Distribuição Lognormal
Distribuição Lognormal
Density
0.0
0.4
0.8
1.2
3.5
4.0
4.5
5.0
Valores