Probabilidade e Estatística - QUIZ 5

Instituto Tecnológico de Aeronáutica

Alexandre A. A. M. de Abreu

1. Distribuição Normal

A amostra de dados para realização dos testes de normalidade é apresentada a seguir:

```r
dados <- c(149.3355, 140.3779, 145.7254, 149.8931, 139.6168, 149.1934,
    129.6147, 134.7523, 167.8030, 171.7407, 157.5422, 160.2664,
    155.4553, 142.5989, 134.9844, 148.5172, 163.1447, 131.0138,
    130.2423, 167.2239, 149.4015, 145.6802, 160.3472, 121.1775,
    136.7295, 162.2381, 150.7192, 117.8144, 137.3630, 158.6373,
    168.0833, 133.9263, 150.9102, 149.4811, 167.4367, 178.0970,
    138.4903, 148.6764, 181.0990, 167.3345, 147.0679, 156.1410,
    148.8734, 140.9484, 147.6408, 134.5726, 184.6812, 134.6648,
    146.8130, 167.4161)
z.dados <- scale(dados)
```

a) Testes de Normalidade

i) Kolmogorov-Smirnov

```r
ks.test(z.dados, "pnorm", 0, 1)
```

```
	One-sample Kolmogorov-Smirnov test

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

ii) Shapiro-Wilk

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

```
	Shapiro-Wilk normality test

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

iii) Anderson-Darlin

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

```
	Anderson-Darling normality test

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

iv) Lilliefors

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

```
	Lilliefors (Kolmogorov-Smirnov) normality test

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

Interpretação dos resultados:

Considerando um nível de significância de 5%, não se rejeita a hipótese de que os dados amostrais seguem a distrição normal, pois os p-valores são maiores do que 0,05 para todos os testes de normalidade realizados anteriormente.

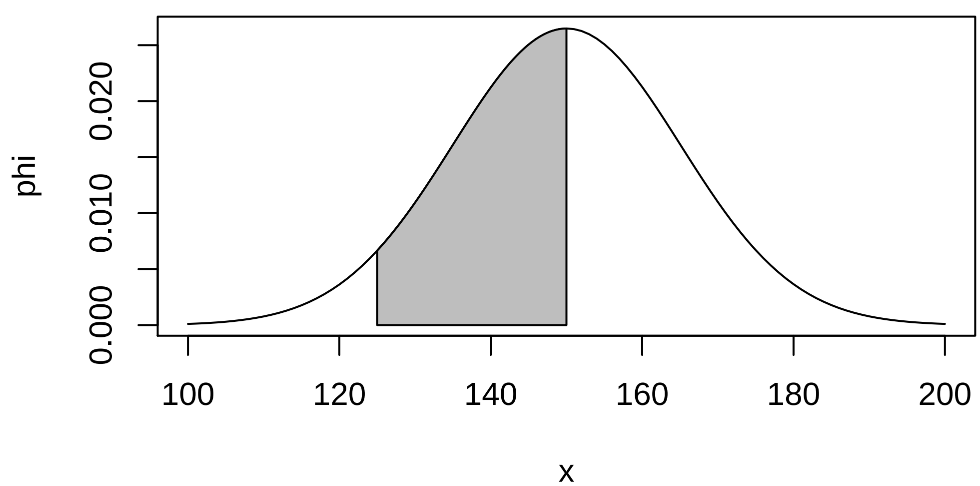
---

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

```r
media <- mean(dados)
sigma <- sd(dados)
z1 <- (125  - media) / sigma
z2 <- (150  - media) / sigma
probabilidade <- pnorm(z2) - pnorm(z1)
probabilidade
```

```
[1] 0.4509603
```

```r
phi <- function(x) {
  (1/(sigma*sqrt(2*pi))) *exp((-1/2)*((x-media)/sigma)^2)
}
x = c(125, seq(125, 150, l=100), 150)
y = c(0, phi(seq(125, 150, l=100)), 0)
plot(phi, 100, 200)
polygon(x = x, y = y, col="gray")
```



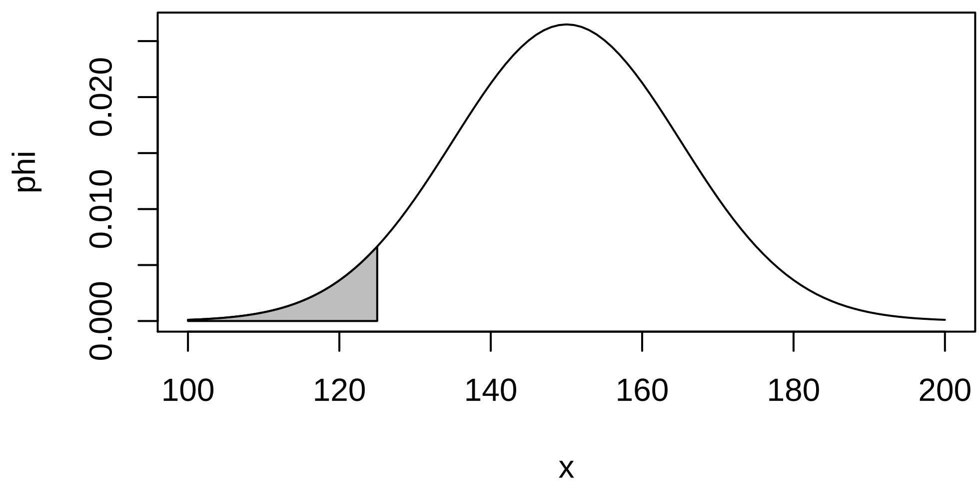
A probabilidade de que uma chamada demore entre 125 e 150 segundos é de 45,09603%.

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

```
probabilidade <- integrate(phi, -Inf, 125)
probabilidade
```

```
0.04824296 with absolute error < 9e-05
```

```
x = c(100, seq(100, 125, l=100), 125)
y = c(0, phi(seq(100, 125, l=100)), 0)
plot(phi, 100, 200)
polygon(x = x, y = y, col="gray")
```



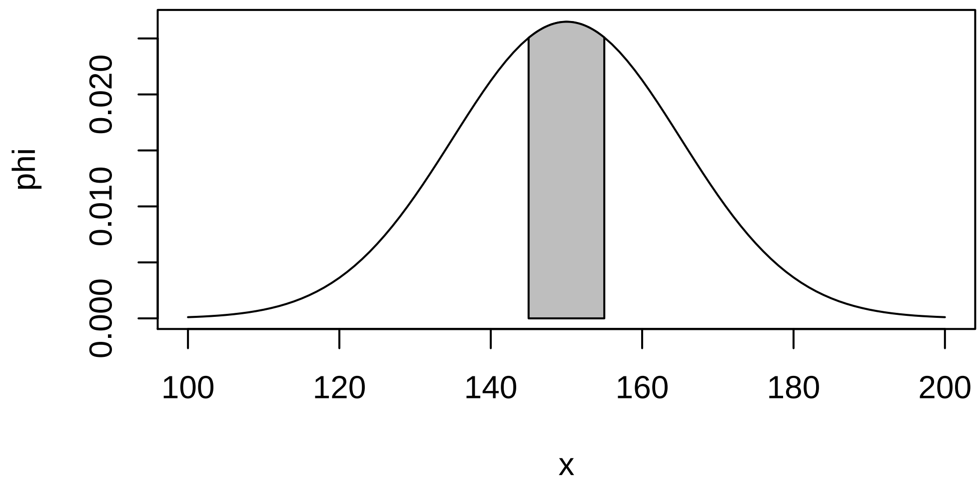
A probabilidade de que uma chamada demore menos de 125 segundos é de 4,824296%.

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

```
z1 <- (145  - media) / sigma
z2 <- (155  - media) / sigma
probabilidade <- pnorm(z2) - pnorm(z1)
probabilidade
```

```
[1] 0.2601309
```

```
x = c(145, seq(145, 155, l=100), 155)
y = c(0, phi(seq(145, 155, l=100)), 0)
plot(phi, 100, 200)
polygon(x = x, y = y, col="gray")
```



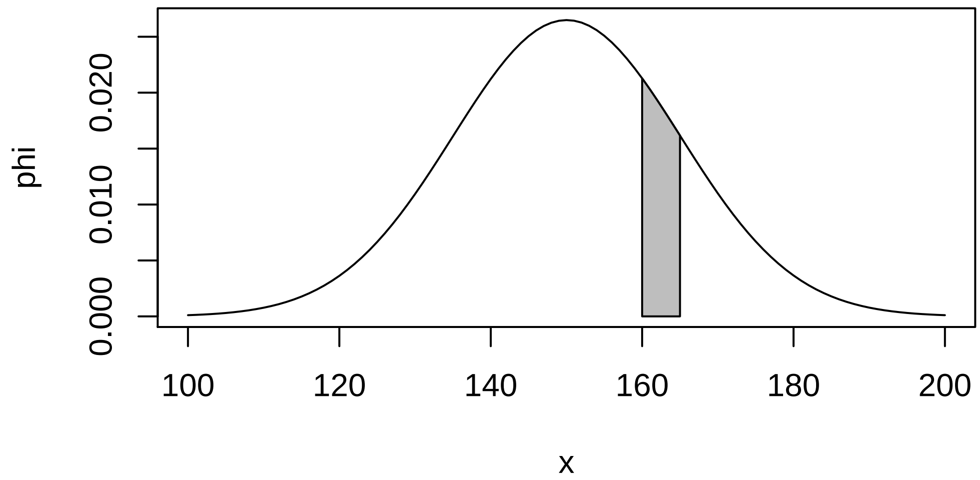
A probabilidade de que uma chamada demore entre 145 e 155 segundos é de 26,01309%.

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

```
z1 <- (160  - media) / sigma
z2 <- (165  - media) / sigma
probabilidade <- pnorm(z2) - pnorm(z1)
probabilidade
```

```
[1] 0.09387641
```

```
x = c(160, seq(160, 165, l=100), 165)
y = c(0, phi(seq(160, 165, l=100)), 0)
plot(phi, 100, 200)
polygon(x = x, y = y, col="gray")
```



A probabilidade de que uma chamada demore entre 160 e 165 segundos é de 9,387641%.

---

2. Identificação de distribuição

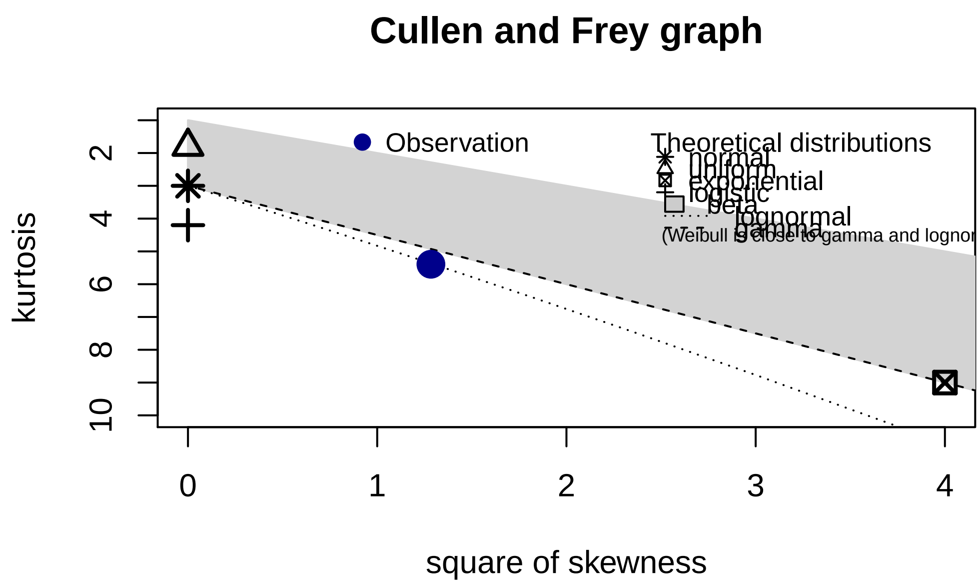
Dados de uma variável aleatória X:

```
dados <- c(1.9993382, 1.4414849, 2.1477166, 2.1087828, 2.1342892,
 ↪  2.1844835, 1.5091879, 2.0467623, 1.0642741, 2.1302612, 1.8389897,
 ↪  1.8924614, 1.9316041, 1.5602204, 1.6991884, 1.7228081, 1.5197833,
 ↪  1.7659242, 0.6914335, 1.4598759, 2.0017607, 1.5139209, 1.8334780,
 ↪  1.8847480, 1.9072389, 1.6294414, 1.9068617, 1.7744973, 2.4300455,
 ↪  1.8958270)
```

a) Faça a identificação da distribuição.

Pelo diagrama de Cullen e Frei, os dados parecem seguir uma distribuição Lognormal ou Weibull.

```
library(fitdistrplus)
library(logspline)
descdist(dados, discrete = FALSE)
```

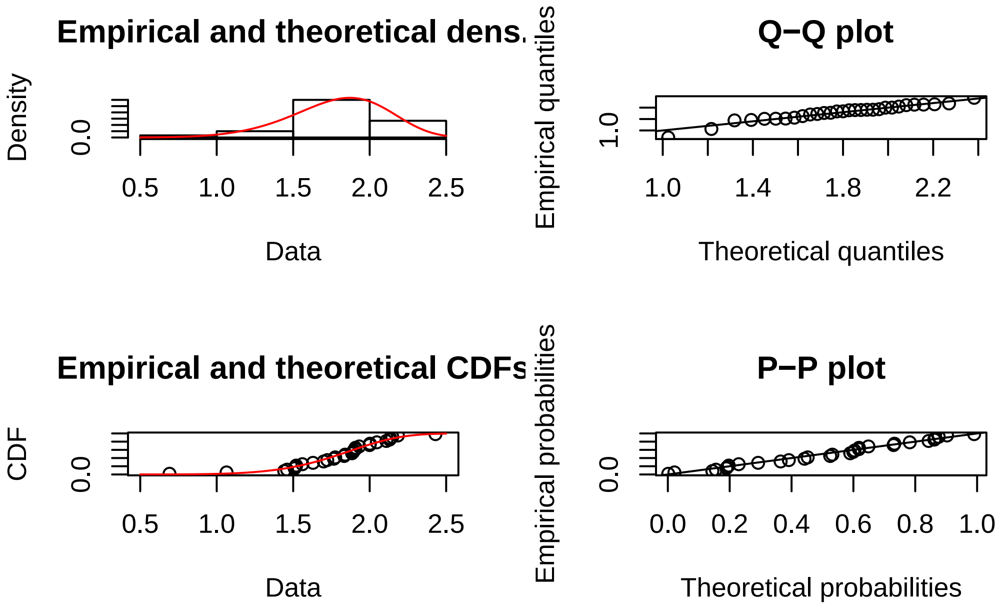



```
summary statistics
------
min:  0.6914335   max:  2.430045
median:  1.861869
mean:  1.787556
estimated sd:  0.3498879
estimated skewness:  -1.133072
estimated kurtosis:  5.391445
```

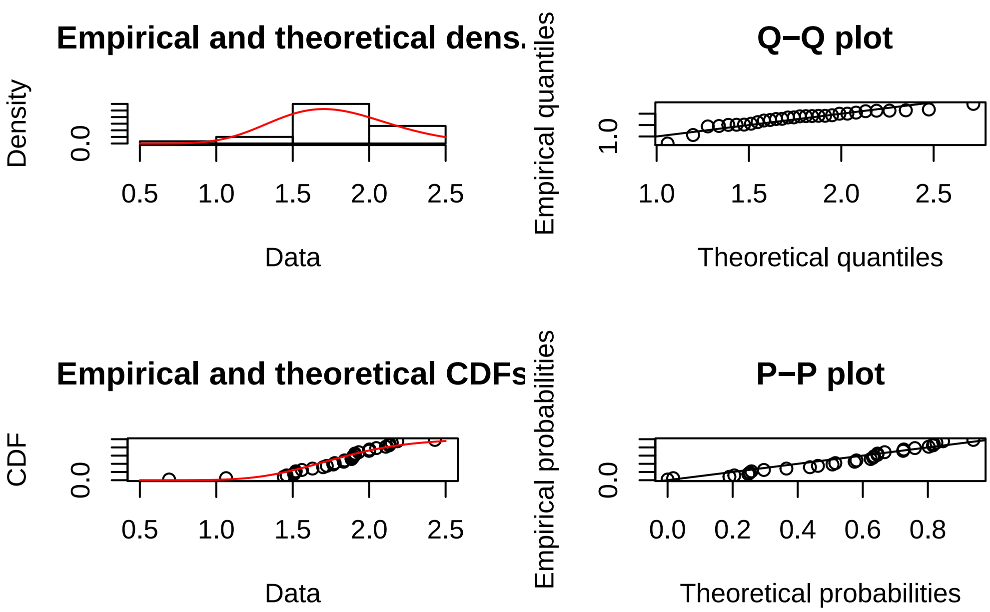
Os gráficos de ajuste de distribuição, dão indícios de que a distribuição Weibull é a que melhor se ajusta aos dados.

```
dados_norm <- (dados - min(dados))/(max(dados) - min(dados))
weibull <- fitdist(dados, "weibull")
gamma <- fitdist(dados, "gamma")
lognormal <- fitdist(dados, "lnorm")
normal <- fitdist(dados, "norm")
beta <- fitdist(dados_norm, "beta", method = "mme")
```

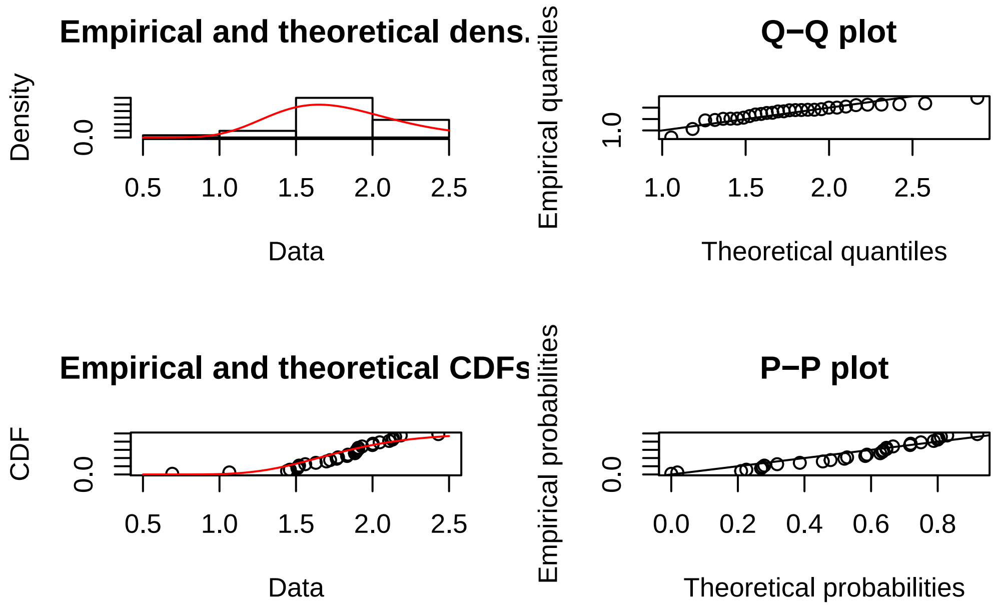
```
plot(weibull)
```



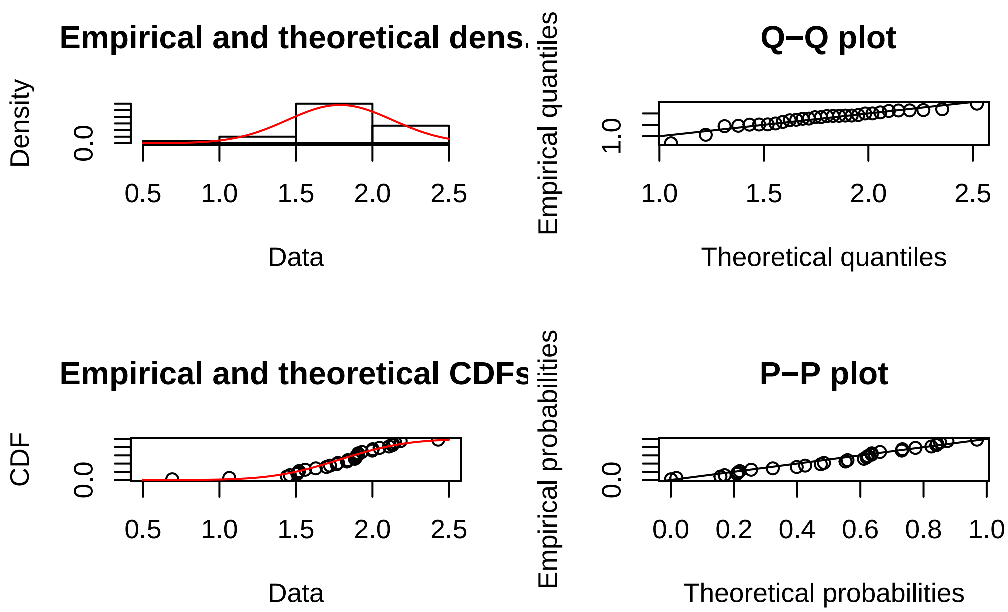

```
plot(gamma)
```




```
plot(lognormal)
```




```
plot(normal)
```



O Critério de Informação de Akaike (AIC) confirma que **a distribuição Weibull** é a que melhor se ajusta aos dados. Seguem os valores:

```r
cat(paste("Weibull: ", weibull$aic, "\nGamma: ", gamma$aic,
 ↪  "\nLognormal: ", lognormal$aic, "\nNormal: ", normal$aic))
```

```
Weibull:  21.9771768698319
Gamma:  31.7072545492645
Lognormal:  36.1032950711012
Normal:  25.110725948175
```

Ajuste de distribuição e teste de Kolmogorov-Smirnov:

```r
mle <- fitdist(dados, "weibull", method="mle")
mle$estimate
```

```
   shape    scale
6.513198 1.918411
```

```r
ks.test(dados, "pweibull", shape = 6.513198, scale = 1.918411,
 ↪  exact=FALSE)
```

```

	One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.091733, p-value = 0.9624
alternative hypothesis: two-sided
```

Portanto, com um nível de significância de 5%, não se rejeita a hipótese de que os dados se ajustem a uma distribuição Weibull, pois o p-valor é 0,9624.

b) Compare os resultados gerados pelo teste de Kolmogorov-Smirnov considerando as distribuições Gama, Lognormal e Weibull.

```r
mle <- fitdist(dados, "gamma", method="mle")
mle
```

```
Fitting of the distribution ' gamma ' by maximum likelihood
Parameters:
      estimate Std. Error
shape 20.98456   5.375681
rate  11.73920   3.043447
```

```r
ks.test(dados, "pgamma", 20.98456, 11.73920, exact = FALSE)
```

```
    One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.14176, p-value = 0.5829
alternative hypothesis: two-sided
```

```r
mle <- fitdist(dados, "lnorm", method="mle")
mle
```

```
Fitting of the distribution ' lnorm ' by maximum likelihood
Parameters:
         estimate Std. Error
meanlog 0.5568348 0.04322583
sdlog   0.2367576 0.03056282
```

```r
ks.test(dados, "plnorm", meanlog=0.5568348, sdlog=0.2367576,
↪   exact=FALSE)
```

```
    One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.15513, p-value = 0.4658
alternative hypothesis: two-sided
```

```r
mle <- fitdist(dados, "weibull", method="mle")
mle
```

```
Fitting of the distribution ' weibull ' by maximum likelihood
Parameters:
      estimate Std. Error
shape 6.513198 0.94665004
scale 1.918411 0.05622354
```

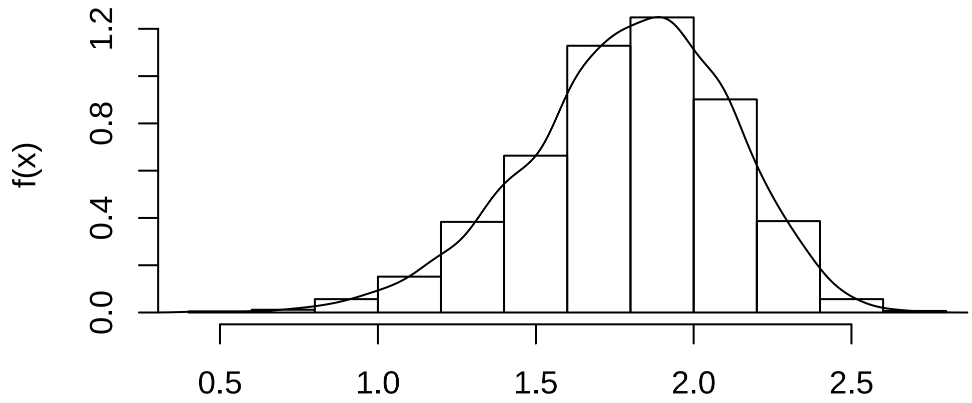
```r
ks.test(dados, "pweibull", shape = 6.513198, scale = 1.918411,
 ↪  exact=FALSE)
```

```

    One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.091733, p-value = 0.9624
alternative hypothesis: two-sided
```

Todos os testes apresentaram p-valores que não justificam rejeitar as respectivas hipóteses nulas (que os dados seguem as distribuições Gamma, Normal e Weibull, respectivamente), com um nível de significância de 5%. Entretanto, a escolha da função de distribuição que melhor ajusta os dados é a Weibull, pois além de apresentar p-valor $= 0.9624$, é a que obteve melhor (menor) Critério de Informação de Akaike.

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

```r
w_2_1 <- rweibull(3000, shape = 6.513198, scale = 1.918411)
hist(w_2_1, lwd = 1, ylab = "f(x)", xlab = "", freq = F, main = "")
lines(x = density(w_2_1))
```



d) Verifique se a área sob a curva estimada é igual a 1.

```
w <- function(x, alpha = 6.513198, beta = 1.918411)
 ↪  {(alpha/(beta^alpha))*(x^(alpha-1))*(exp(-(x/beta)^alpha))}
x <- seq(0, 5, by = .01)
area <- integrate(w, 0, Inf)
area
```

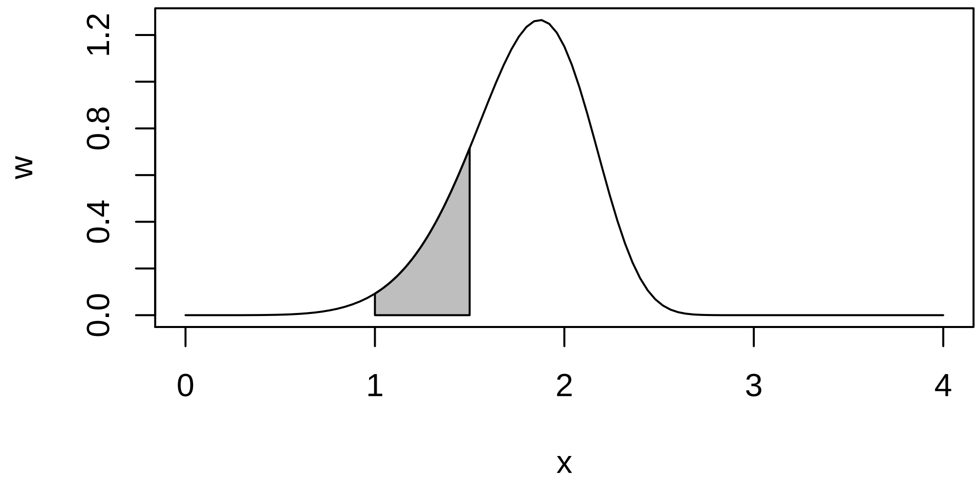
```
1 with absolute error < 3e-08
```

e) Para a distribuição escolhida, calcule a área gerada no intervalo [1; 1,5]. Plotar e destacar essa área.

```
a1 <- pweibull(1, shape = 6.513198, scale = 1.918411)
a2 <- pweibull(1.5, shape = 6.513198, scale = 1.918411)
probabilidade <- a2 - a1
probabilidade
```

```
[1] 0.1681586
```

```
q1 <- qweibull(a1, shape = 6.513198, scale = 1.918411)
q2 <- qweibull(a2, shape = 6.513198, scale = 1.918411)
x = c(q1, seq(q1, q2, l=100), q2)
y = c(0, w(seq(q1, q2, l=100)), 0)
plot(w, 0, 4)
polygon(x = x, y = y, col="gray")
```



---

3. Normalidade e intervalo de confiança

Amostra aleatória simples com os valores de inflação para os anos de 2013 a 2022 coletados do site do Banco Central do Brasil:

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

a) Faça os testes de Shapiro-Wilk e de Lilliefors para verificar a normalidade. Qual é sua conclusão?

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

```

	Shapiro-Wilk normality test

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

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

```
    Lilliefors (Kolmogorov-Smirnov) normality test

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

Considerando um nível de significância de 5%, não se rejeita a hipótese de que os dados amostrais de inflação seguem a distrição normal, considerando que ambos p-valores são maiores do que 0,05.

b) Usando esses dados, construa um intervalo de confiança de 99% para a média da inflação.

```r
t.test(inflacao, conf.level = .99)
```

```
    One Sample t-test

data:  inflacao
t = 7.5586, df = 9, p-value = 3.473e-05
alternative hypothesis: true mean is not equal to 0
99 percent confidence interval:
 3.457911 8.674089
sample estimates:
mean of x
    6.066
```

```r
x_bar <- mean(inflacao)
s <- sd(inflacao)
alpha <-  0.01
n <- length(inflacao)
gl <- n - 1
t <- qt(alpha/2, gl, lower.tail = FALSE)
LI <- x_bar - t*(s/(n^0.5))
LS <- x_bar + t*(s/(n^0.5))
cat(paste0("[", LI, ", ", LS, "]"))
```

```
[3.45791074446933, 8.67408925553067]
```

Portanto, pode-se afirmar com nível de confiaça de 99% que a média da inflação estará no intervalo [3.457911, 8.674089].

c) Os especialistas têm a opinião de que o intervalo calculado para a média é muito amplo e querem um intervalo de comprimento total igual a 3. Encontre o nível de confiança para esse novo intervalo.

```
# Calcule o valor de t
t <- ((3/2) * sqrt(n)) / s
LI <- x_bar - t*(s/(n^0.5))
LS <- x_bar + t*(s/(n^0.5))
cat(paste0("Intervalo: [", LI, ", " , LS, "]"))
```

```
Intervalo: [4.566, 7.566]
```

```
alpha <- 2 * (pt(t, df = gl, lower.tail=FALSE))
alpha
```

```
[1] 0.09443553
```

```
c <- 1 - alpha
c
```

```
[1] 0.9055645
```

O intervalo [4.566,7.566] de comprimento total igual a 3 é obtido para um alpha = 0.09443553 e, portanto, o nível de confiança necessário é de 90,55645%.

d) Construa um intervalo de confiança de 90% para o desvio padrão.

```
c <- 0.9
q2_1 <- qchisq((1 - c) / 2, df = gl)
q2_2 <- qchisq(1 - (1 - c) / 2, df = gl)

# Intervalo de confiança para o desvio padrão
```

```r
intervalo <- (n - 1) * s^2 / c(q2_2, q2_1)
intervalo
```

```
[1]  3.426025 17.432443
```

Portanto, o intervalo de confiaça de 90% para média da amostra da inflação é [3.426025 17.432443].

e) Agora, teste a normalidade para toda a série histórica desde o início do regime de metas. Ou seja, utilize os dados de inflação efetiva de 1999 até 2022. Qual a conclusão?

```r
inflacao <- c(8.94, 5.97, 7.67, 12.53, 9.3, 7.6, 5.69, 3.14, 4.46, 5.9,
 ↪  4.31, 5.91, 6.5, 5.84, 5.91, 6.41, 10.67, 6.29, 2.95, 3.75, 4.31,
 ↪  4.52, 10.06, 5.79)
length(inflacao)
```

```
[1] 24
```

Como a amostra tem menos de 30 observações, será utilizado o teste de Shapiro-Wilk.

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

```

	Shapiro-Wilk normality test

data:  inflacao
W = 0.92757, p-value = 0.08602
```

Considerando um nível de significância de 5%, não rejeita-se a hipótese de que os dados da inflação para os anos de 1999 até 2022 seguem a distrição normal. É possível chegar na mesma conclusão aplicando-se os testes de Anderson-Darling e Kolmogorov-Smirnov.

```r
ad.test(inflacao)
```

```
	Anderson-Darling normality test

data:  inflacao
A = 0.70952, p-value = 0.05561
```

```r
ks.test(scale(inflacao), "pnorm", 0, 1)
```

```
	One-sample Kolmogorov-Smirnov test

data:  scale(inflacao)
D = 0.19749, p-value = 0.3065
alternative hypothesis: two-sided
```

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

```
	Lilliefors (Kolmogorov-Smirnov) normality test

data:  inflacao
D = 0.19749, p-value = 0.01631
```

Entretanto, pelo teste de Lilliefors (p-value = 0.01631), rejeita-se a hipótese de que os dados da inflação para os anos de 1999 até 2022 seguem a distrição normal, com nível de 5% significância.

---

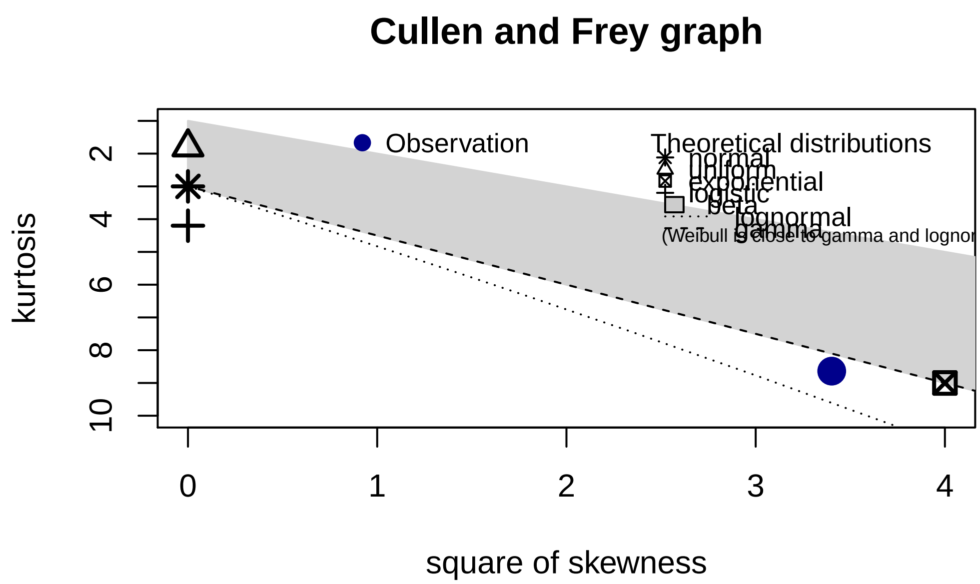
4. Identificação de distribuição

Identifique a distribuição de cada um dos conjuntos de dados mostrado a seguir.

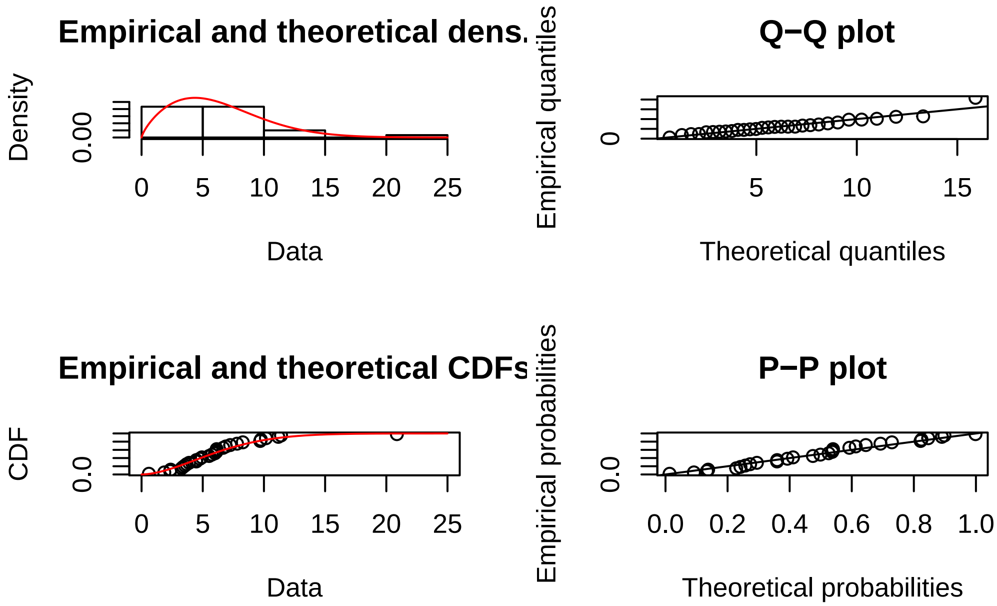
a)

```r
dados <- c(20.8625807, 7.2445709, 4.4659396, 3.2712081, 4.9300651,
↪  5.7444213, 6.6700987, 11.1750446, 2.3753017, 3.5425386, 0.5978486,
↪  6.8869953, 6.1102197, 8.2716973, 9.7465462, 3.3991988, 1.8557047,
↪  11.3983705, 3.6847590, 2.3327479, 6.1364329, 4.4686122, 7.8007834,
↪  4.7649257, 3.8829371, 5.9986131, 5.5163819, 9.6951710, 10.1645820,
↪  6.1304865)

id_dist <- function (dados) {
  descdist(dados, discrete = FALSE)

  dados_norm <- (dados - min(dados)) / (max(dados) - min(dados))
  weibull <- fitdist(dados, "weibull", method="mle")
  gamma <- fitdist(dados, "gamma", method="mle")
  lognormal <- fitdist(dados, "lnorm", method="mle")
  normal <- fitdist(dados, "norm", method="mle")
  beta <- fitdist(dados_norm, "beta", method = "mme")
  uniforme <- fitdist(dados, "unif", method = "mle")
  exponencial <- fitdist(dados, "exp", method = "mle")
  logistica <- fitdist(dados, "logis", method = "mle")
  cat("Weibull: ")
  plot(weibull)

  cat("Gamma: ")
  plot(gamma)

  cat("Lognormal: ")
  plot(lognormal)

  cat("Normal: ")
  plot(normal)

  cat("Beta: ")
  plot(beta)

  cat("Uniforme: ")
  plot(uniforme)

  cat("Exponencial: ")
  plot(exponencial)

  cat("Logística: ")
```

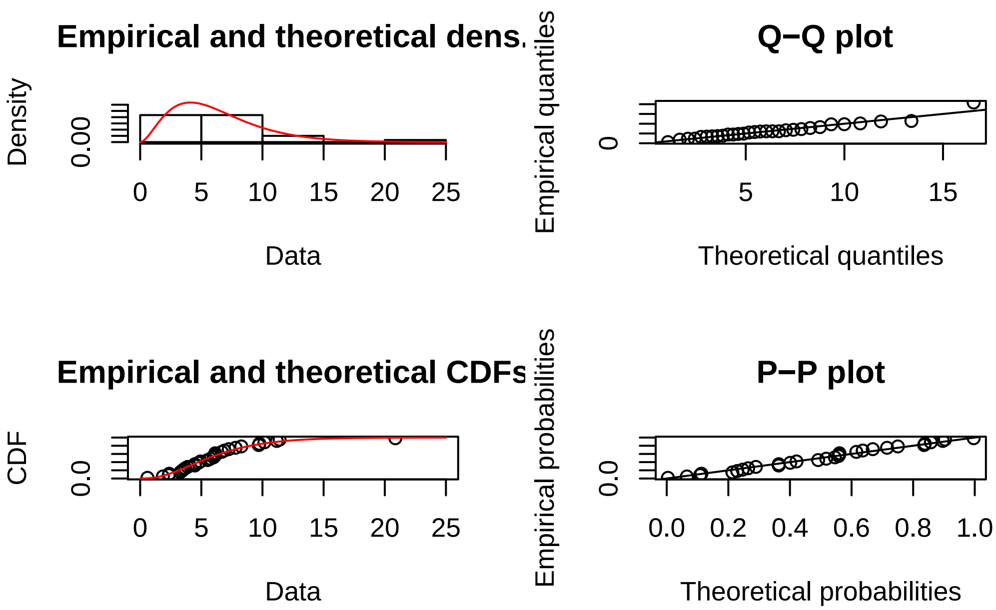
```
  plot(logistica)

  cat(paste("Critério de Informação de Akaike\n"))
  cat(paste("Weibull: ", weibull$aic, "\n"))
  cat(paste("Gamma: ", gamma$aic, "\n"))
  cat(paste("Lognormal: ", lognormal$aic, "\n"))
  cat(paste("Normal: ", normal$aic, "\n"))
  cat(paste("Beta: ", beta$aic, "\n"))
  cat(paste("Uniforme: ", uniforme$aic, "\n"))
  cat(paste("Exponencial: ", exponencial$aic, "\n"))
  cat(paste("Logística: ", logistica$aic, "\n"))
}

id_dist(dados);
```



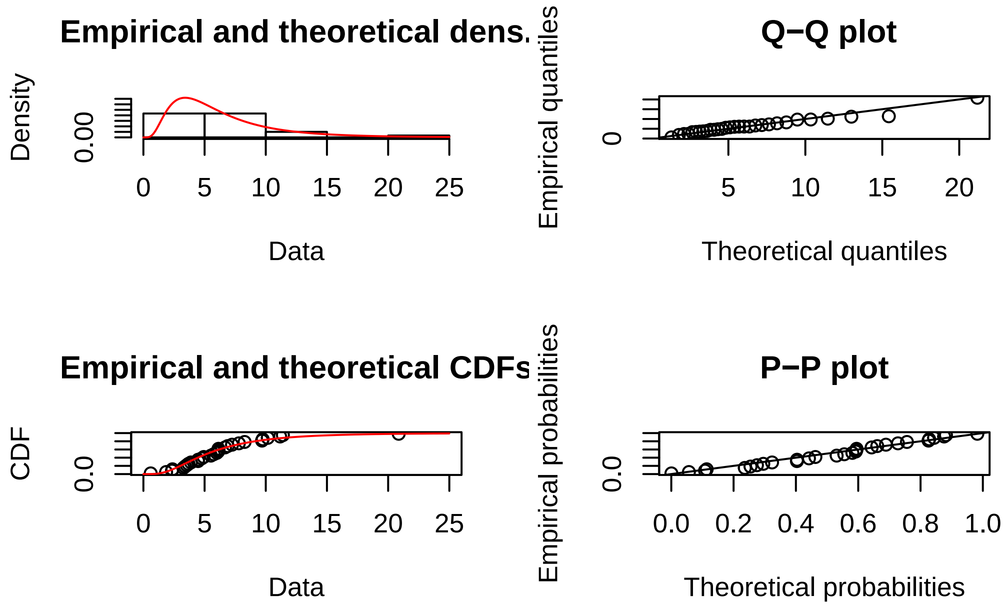

```
Weibull:
```



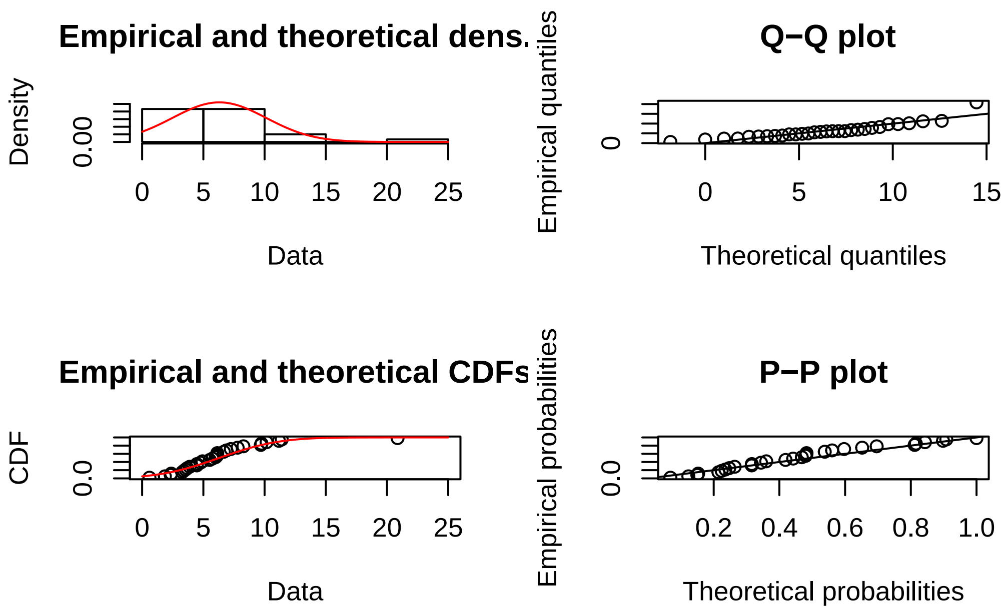
Gamma:



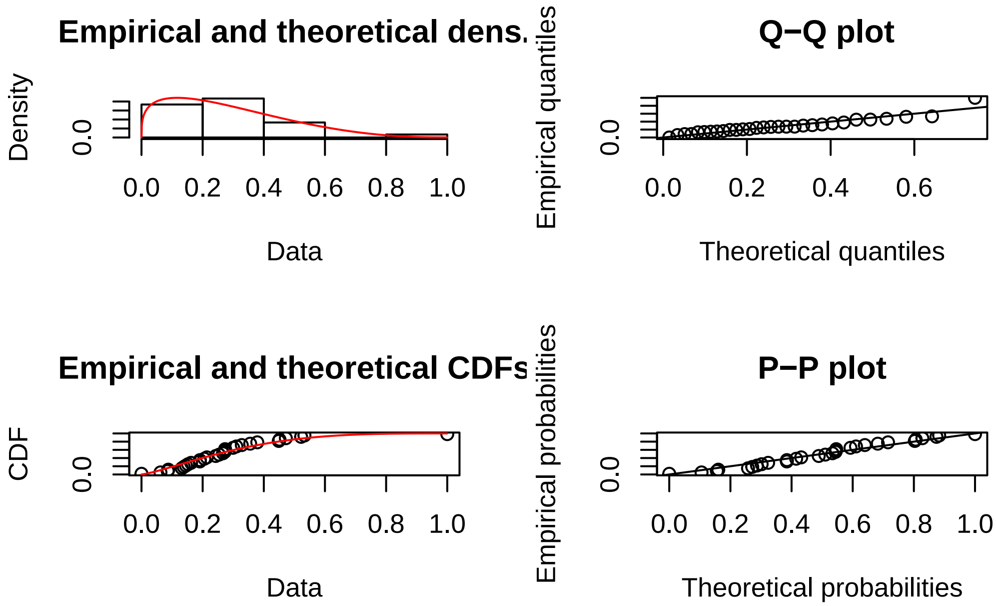
Lognormal:


Empirical and theoretical dens.
Density
Data
Q−Q plot
Empirical quantiles
Theoretical quantiles
Empirical and theoretical CDFs
CDF
Data
P−P plot
Empirical probabilities
Theoretical probabilities

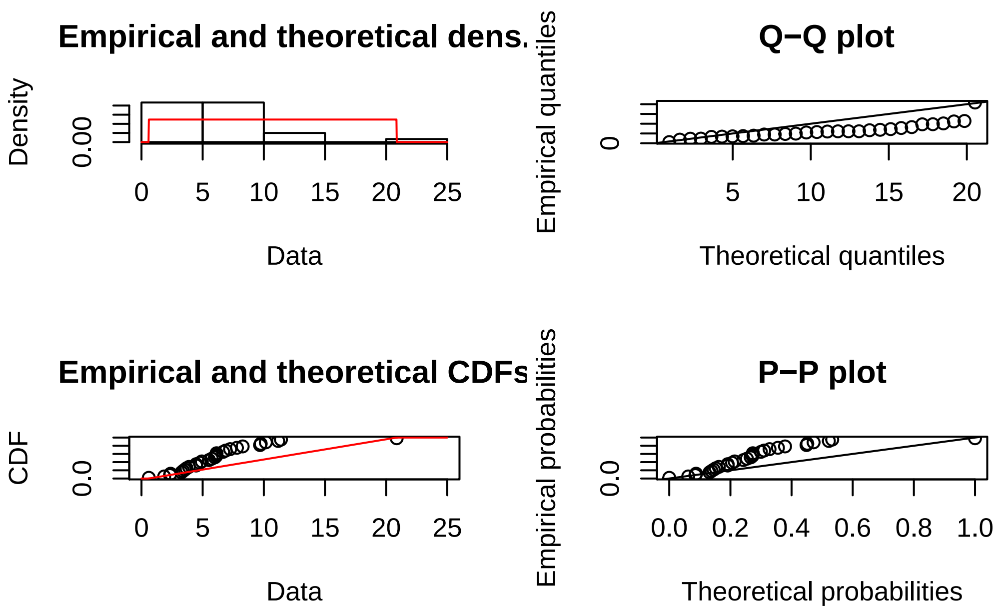
Normal:


Empirical and theoretical dens.
Density
Data
Q−Q plot
Empirical quantiles
Theoretical quantiles
Empirical and theoretical CDFs
CDF
Data
P−P plot
Empirical probabilities
Theoretical probabilities

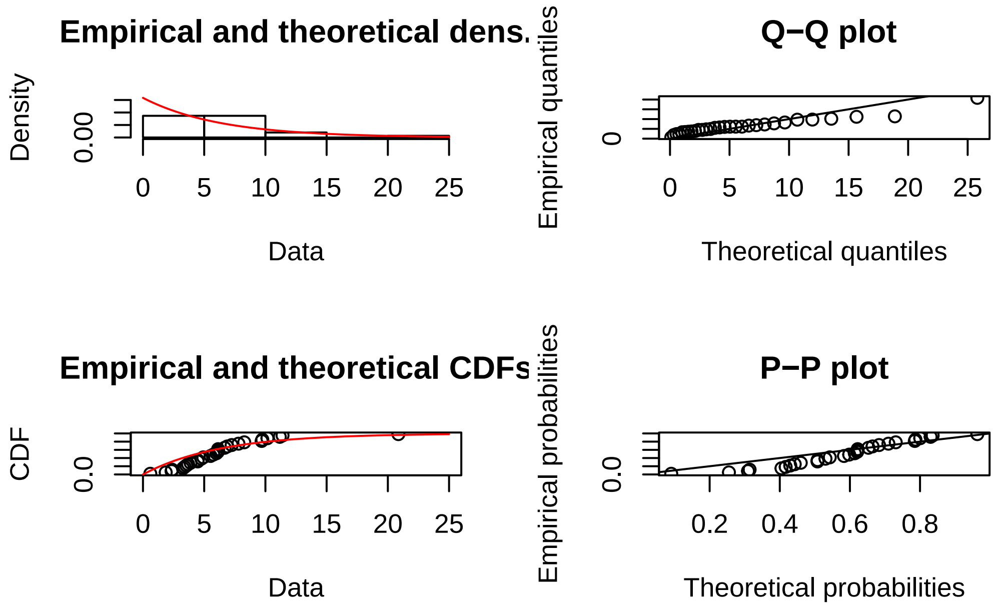
Beta:



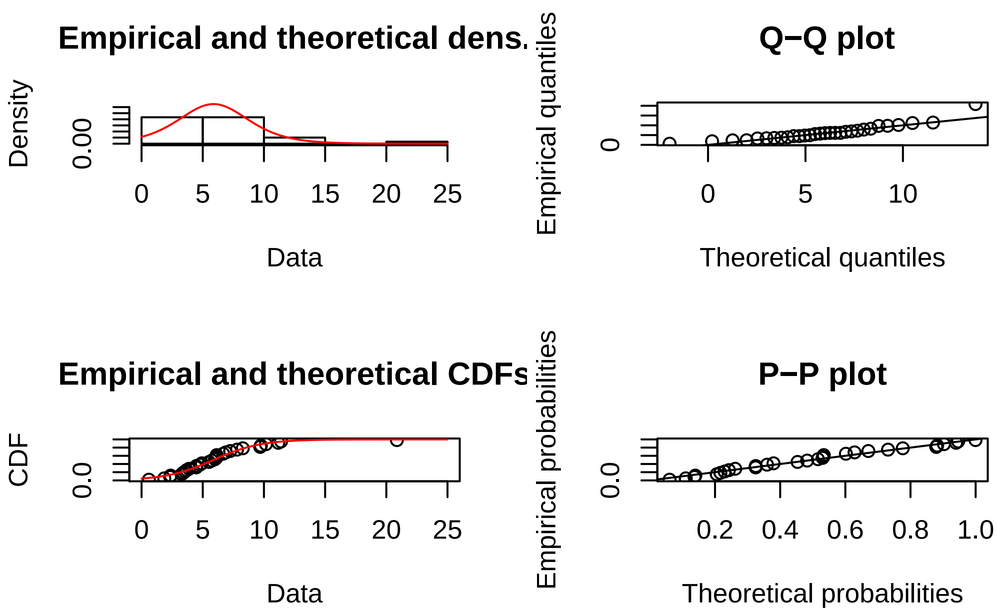
Uniforme:



Exponencial:




```
Logística:
```




```
Critério de Informação de Akaike
Weibull:  161.672292309238
Gamma:  160.373248793533
Lognormal:  163.258756195955
```

```
Normal:  169.773122301626
Beta:  Inf
Uniforme:  184.532922409344
Exponencial:  172.47257861511
Logística:  164.7748413195
```

A distribuição que apresenta menor Critério de Informação de Akaike é a **Gamma**. Portanto, realiza-se o teste de Kolmogorov-Smirnov e não se rejeita a hipótese de que os dados seguem a distribuição **Gamma**, com um nível de significância de 5%.

```r
fitdist(dados, "gamma", method="mle")
```

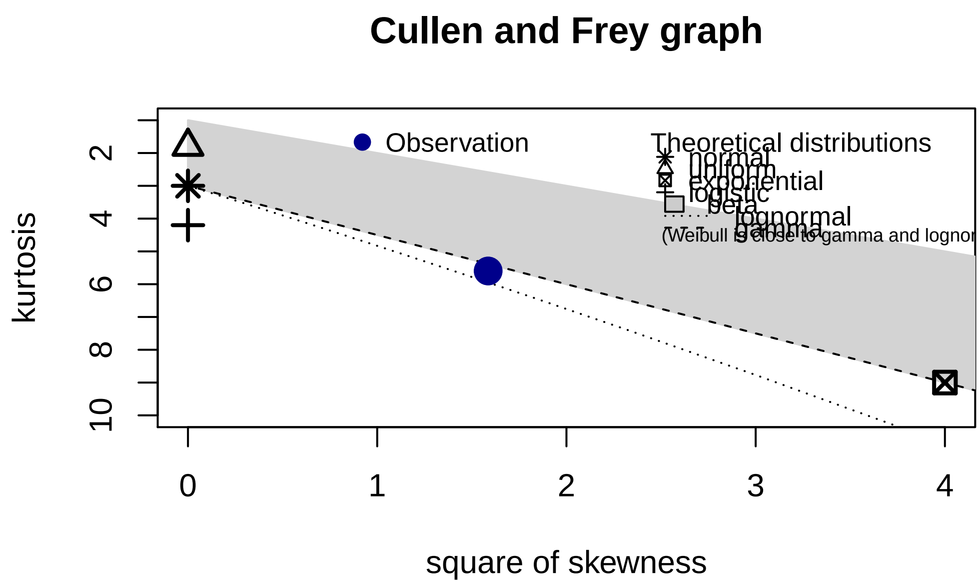
```
Fitting of the distribution ' gamma ' by maximum likelihood
Parameters:
       estimate Std. Error
shape 2.8718513  0.7027011
rate  0.4555631  0.1217956
```

```r
ks.test(dados, "pgamma", 2.8718513, 0.4555631, exact=FALSE)
```

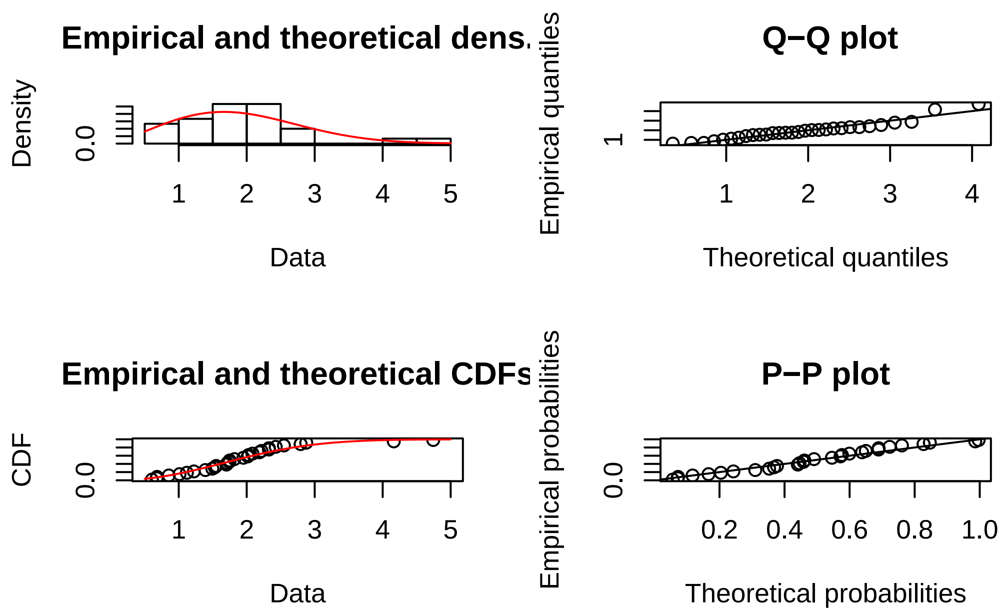
```

	One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.079846, p-value = 0.9909
alternative hypothesis: two-sided
```

b)

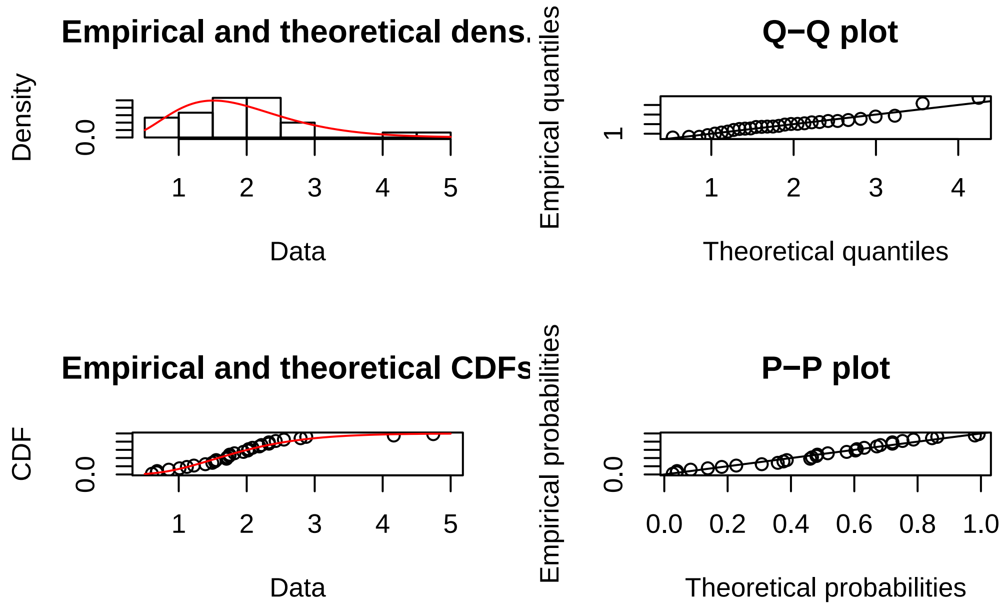
```r
dados <- c(1.4940354, 2.0164275, 1.9513521, 1.5298282, 0.6815670,
 ↪  2.4267801, 0.6762800, 1.7018986, 4.1632638, 2.5472784, 2.2174151,
 ↪  0.6058986, 1.7432601, 1.1199216, 1.7135932, 2.8758657, 0.8537880,
 ↪  1.5511504, 2.3262178, 2.3267933, 1.3916375, 4.7439947, 2.1864812,
 ↪  2.0269031,1.7489244, 1.8191036, 2.0845146, 1.2229195, 1.0115042,
 ↪  2.7931222)
id_dist(dados);
```



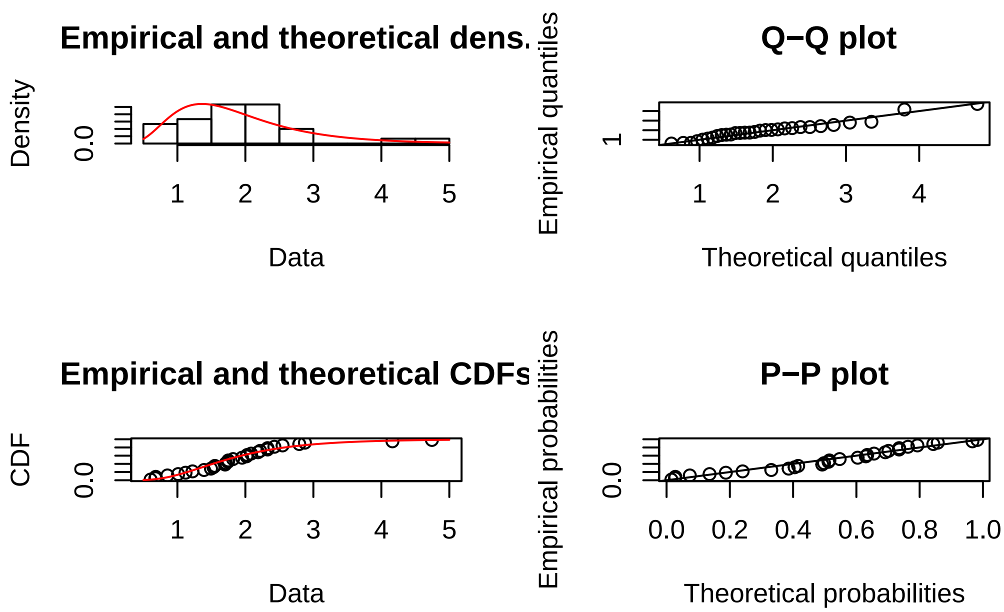
Weibull:



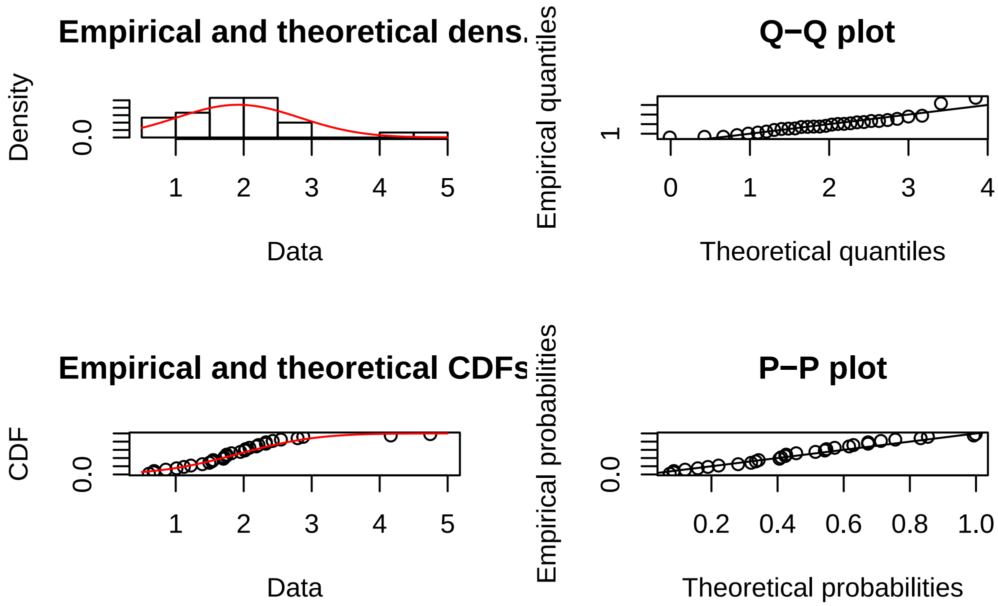
Gamma:



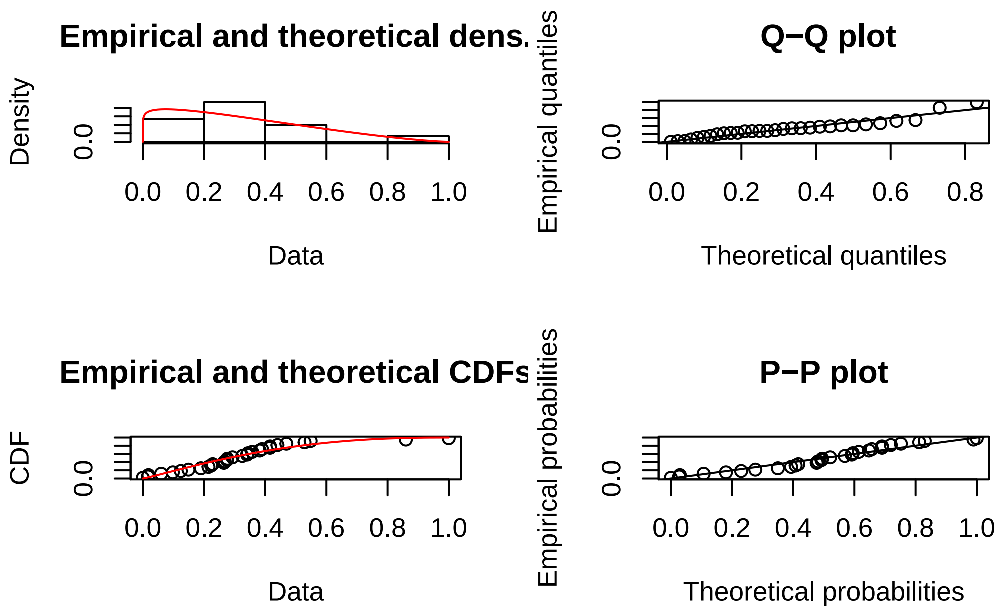
Lognormal:



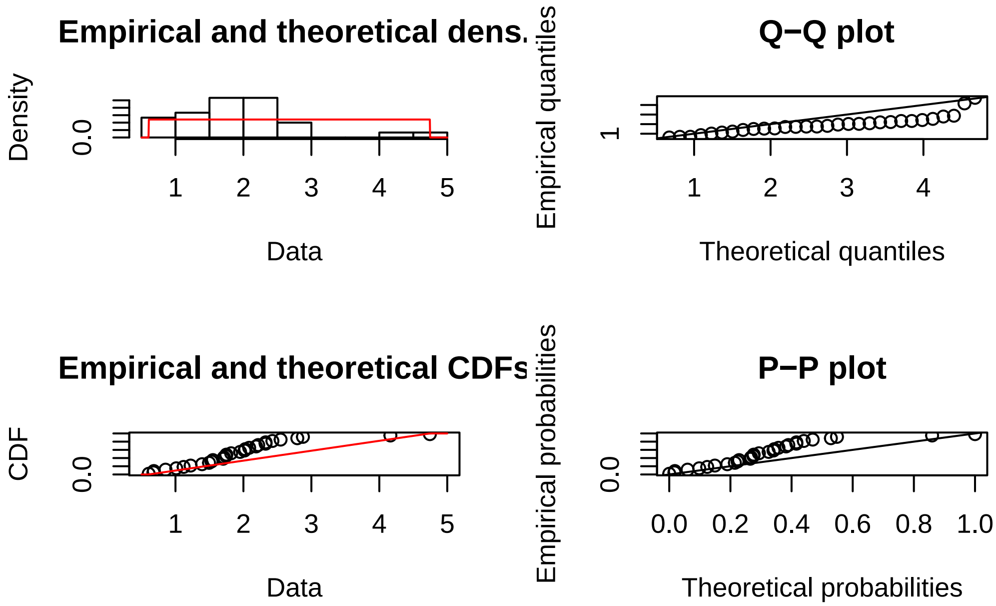
Normal:



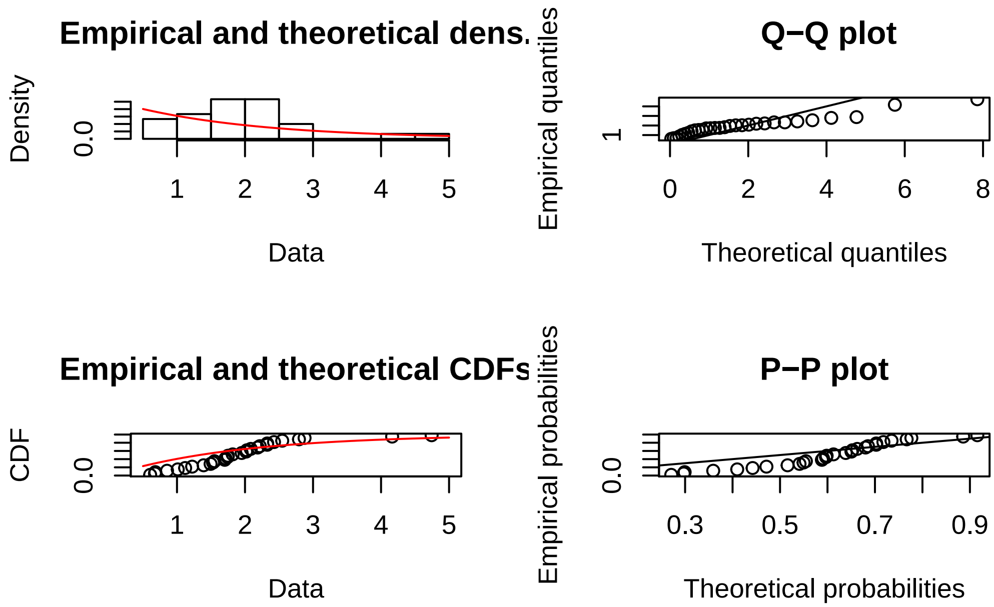
Beta:



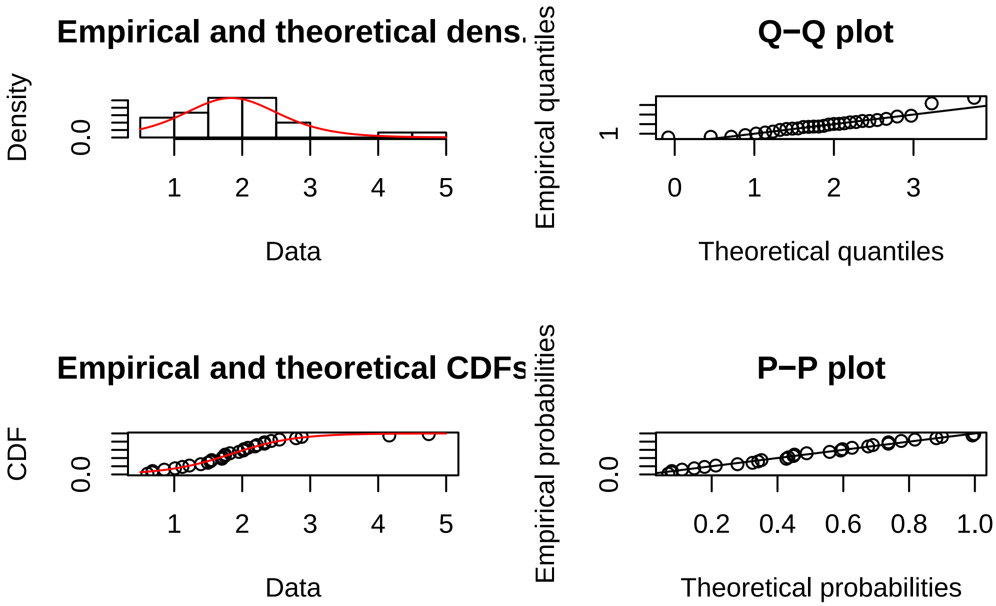
Uniforme:



Exponencial:



Logística:




```
Critério de Informação de Akaike
Weibull:  79.5774526646433
Gamma:  77.3266281106984
Lognormal:  77.8469359969155
Normal:  83.2772121593977
Beta:  Inf
Uniforme:  89.2141481699972
Exponencial:  101.089198302578
Logística:  80.2541161161851
```

A distribuição que apresenta menor Critério de Informação de Akaike é a **Gamma**. Portanto, realiza-se o teste de Kolmogorov-Smirnov e não se rejeita a hipótese de que estes dados também seguem a distribuição **Gamma**, com um nível de significância de 5%.

```r
fitdist(dados, "gamma", method="mle")
```

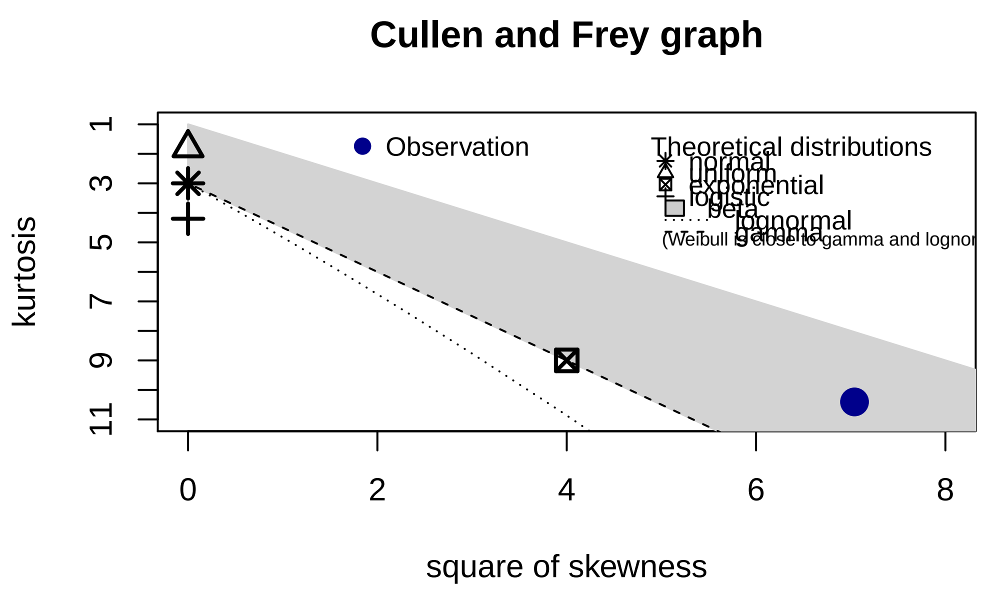
```
Fitting of the distribution ' gamma ' by maximum likelihood
Parameters:
      estimate Std. Error
shape 4.697015  1.1722338
rate  2.448444  0.6449308
```

```r
ks.test(dados, "pgamma", 4.697015, 2.448444, exact=FALSE)
```

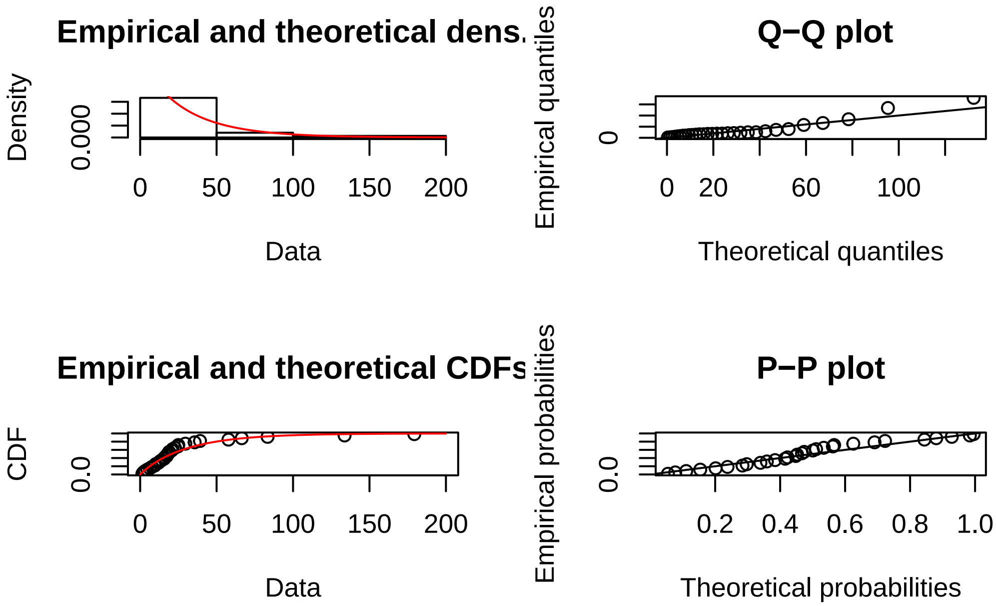
```
	One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.094125, p-value = 0.9531
alternative hypothesis: two-sided
```

c)

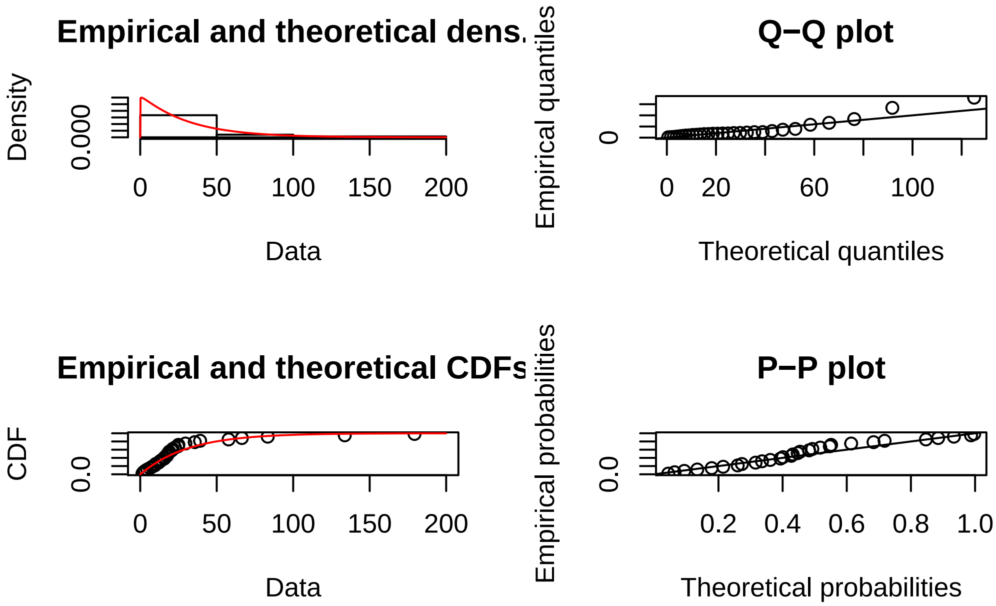
```r
dados <- c(9.534149, 12.878719, 35.635908, 39.158389, 10.091099,
 ↪  133.714299, 15.684000, 3.179206, 16.073085, 57.767201, 29.543033,
 ↪  24.672685, 11.955565, 2.132028, 17.455254, 20.569096, 6.293823,
 ↪  22.717485, 83.353863, 18.544482, 66.437399, 4.616951, 18.931367,
 ↪  1.464430, 21.180916, 179.315876, 24.941790, 14.105447,
 ↪  7.680880,17.688369)
id_dist(dados);
```



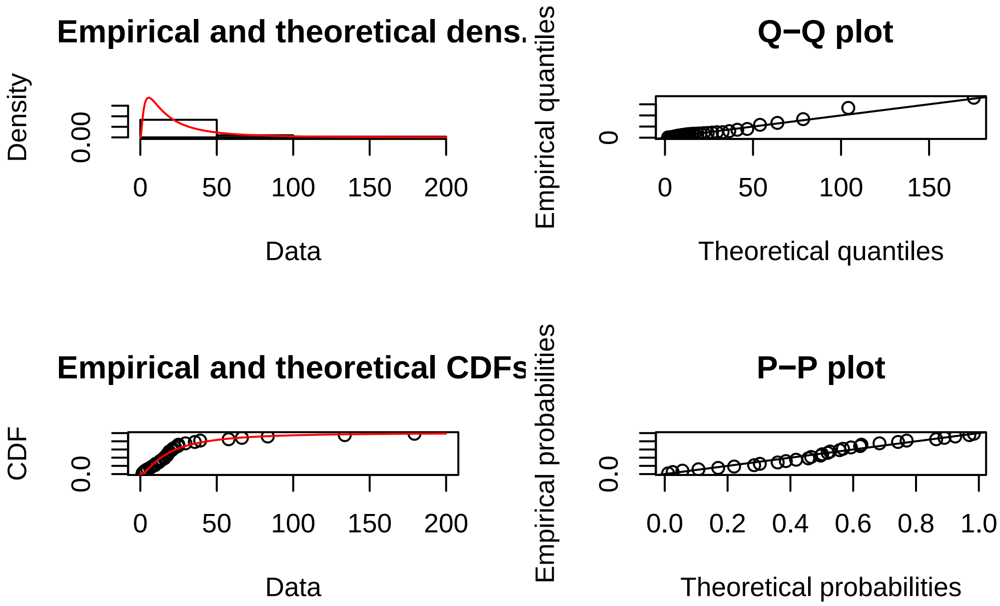

```
Weibull:
```



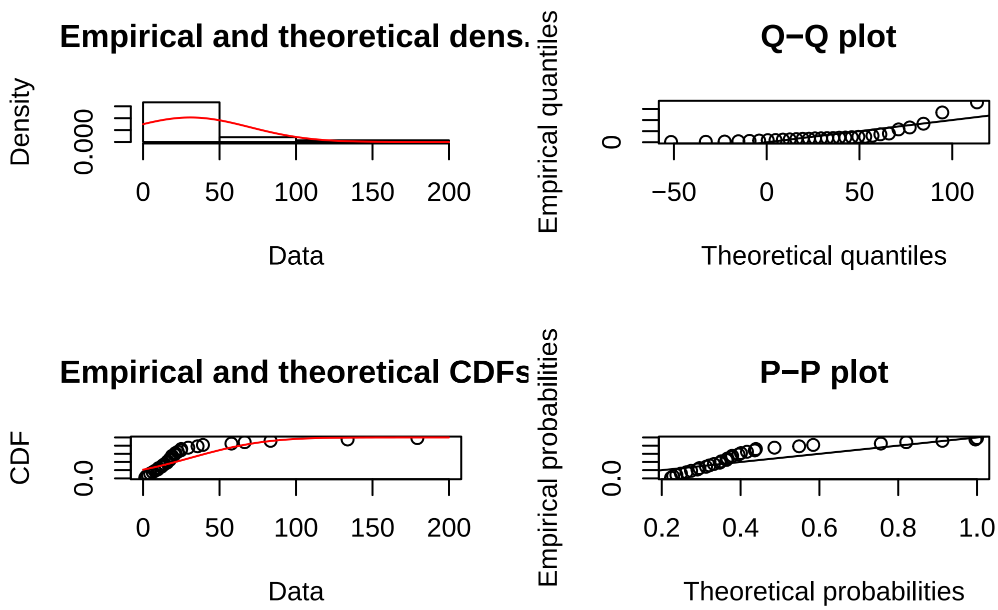
Gamma:



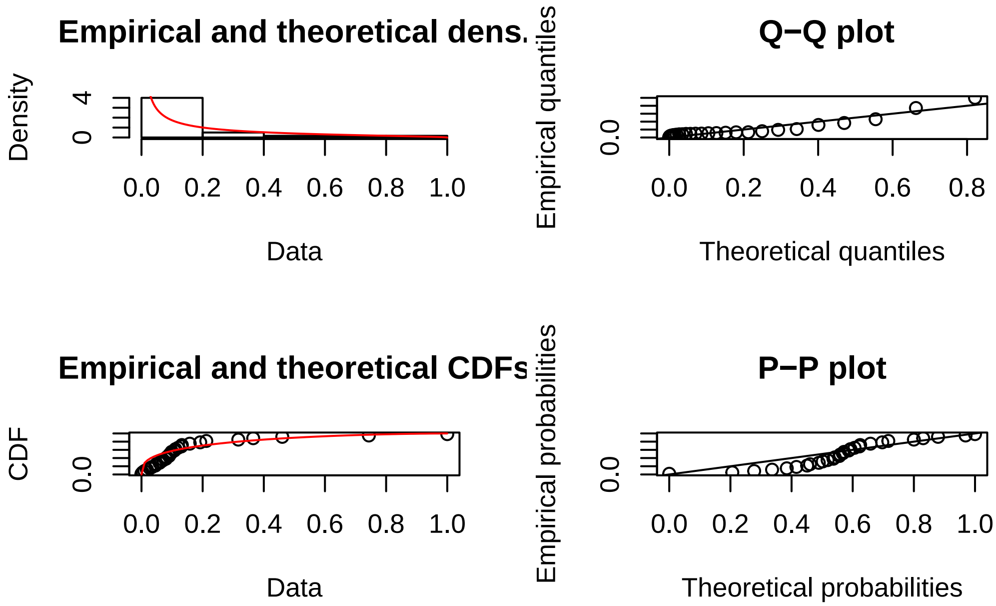
Lognormal:


Empirical and theoretical dens.
Density
0.00
0
50
100
150
200
Data
Q−Q plot
Empirical quantiles
0
0
50
100
150
Theoretical quantiles
Empirical and theoretical CDFs
CDF
0.0
0
50
100
150
200
Data
P−P plot
Empirical probabilities
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Theoretical probabilities

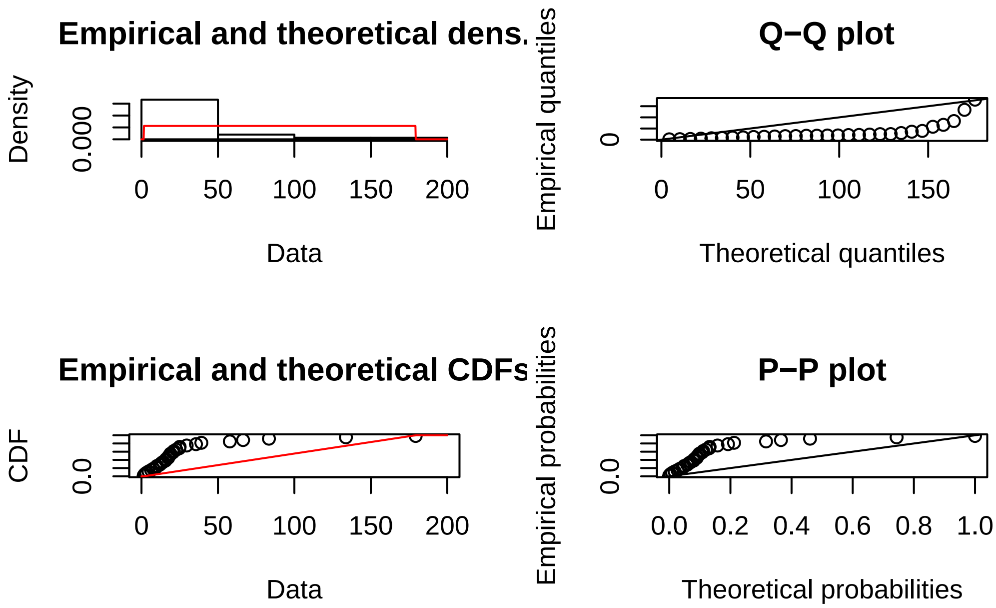
Normal:


Empirical and theoretical dens.
Density
0.000
0
50
100
150
200
Data
Q−Q plot
Empirical quantiles
0
−50
0
50
100
Theoretical quantiles
Empirical and theoretical CDFs
CDF
0.0
0
50
100
150
200
Data
P−P plot
Empirical probabilities
0.0
0.2
0.4
0.6
0.8
1.0
Theoretical probabilities

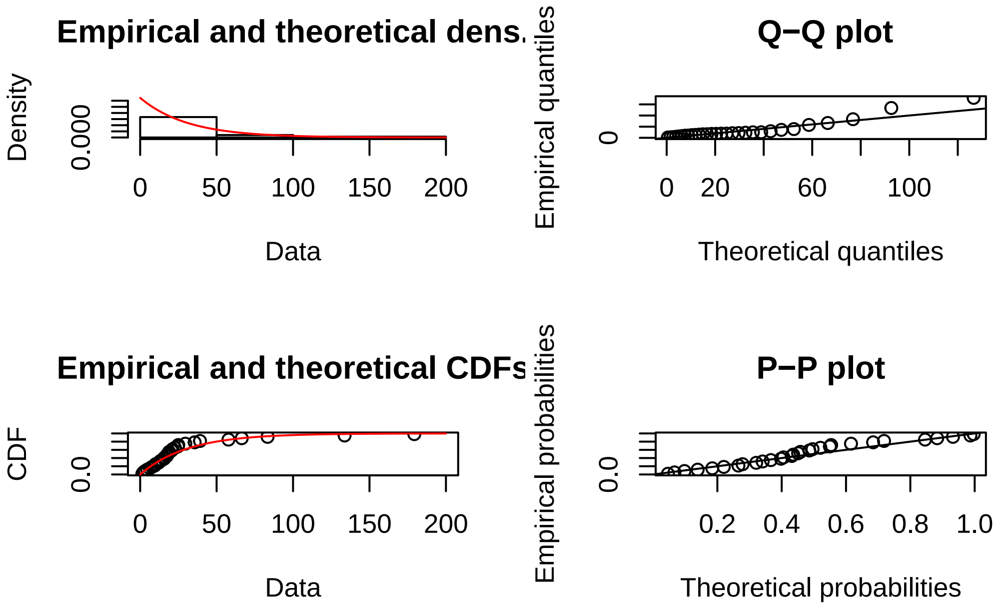
Beta:



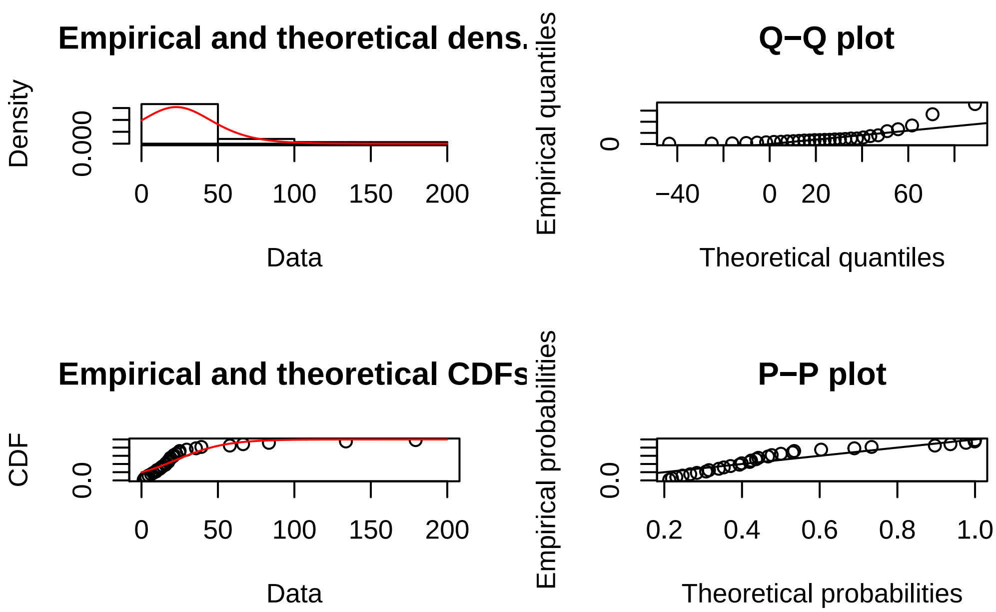
Uniforme:



Exponencial:



Logística:




```
Critério de Informação de Akaike
Weibull:  269.726392071423
Gamma:  269.854267184135
Lognormal:  265.910553623527
```

```
Normal:  308.549444866565
Beta:  NaN
Uniforme:  314.856917728505
Exponencial:  267.86587199743
Logística:  297.626998043809
```

A distribuição que apresenta menor Critério de Informação de Akaike é a **Lognormal**. Portanto, realiza-se o teste de Kolmogorov-Smirnov e não se rejeita a hipótese de que os dados seguem a distribuição **Lognormal**, com um nível de significância de 5%.

```r
fitdist(dados, "lnorm", method="mle")
```

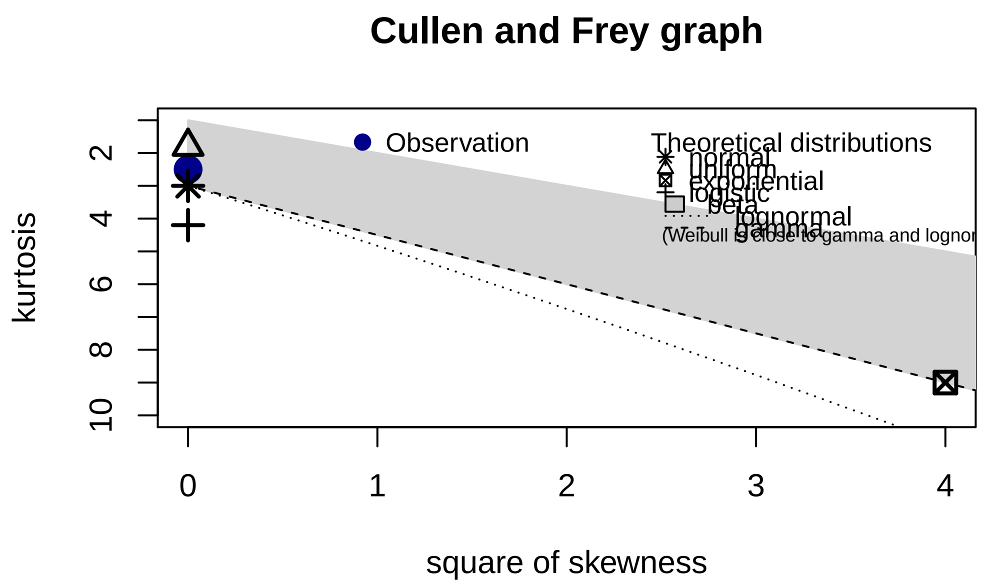
```
Fitting of the distribution ' lnorm ' by maximum likelihood
Parameters:
        estimate Std. Error
meanlog 2.869537  0.1971287
sdlog   1.079718  0.1393905
```

```r
ks.test(dados, "plnorm", 2.869537, 1.079718, exact=FALSE)
```

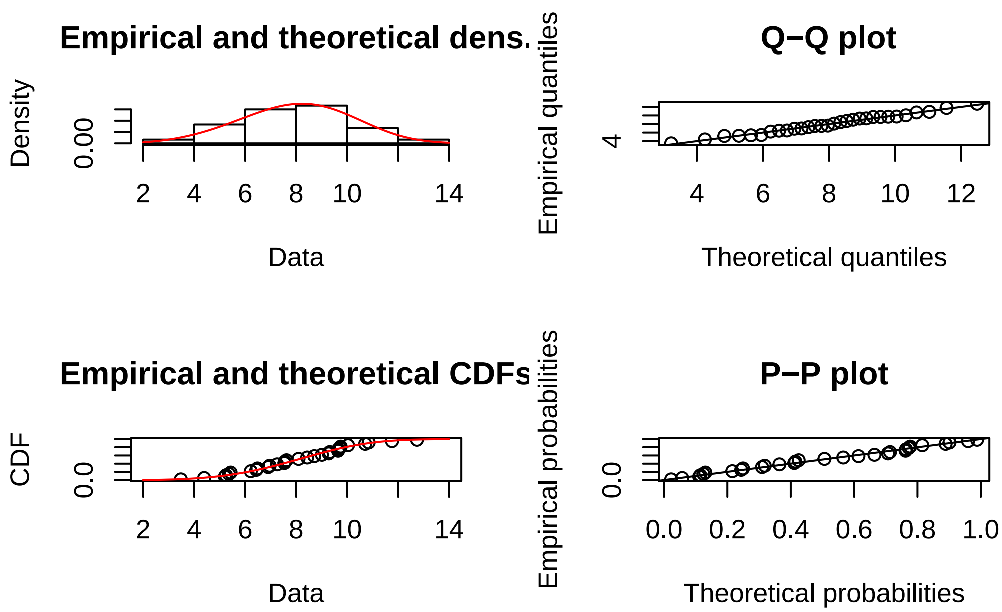
```
    One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.10729, p-value = 0.8802
alternative hypothesis: two-sided
```

d)

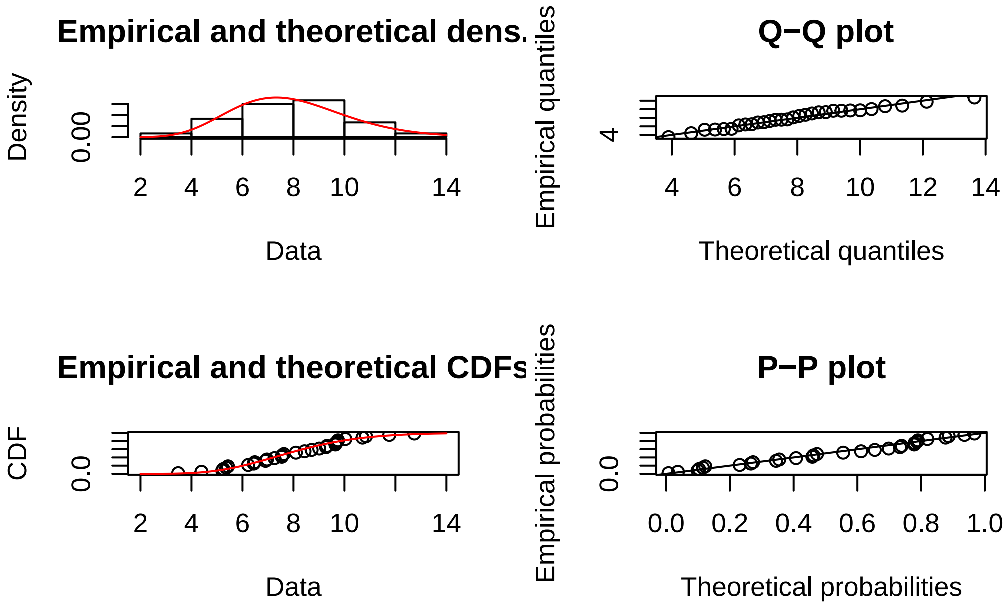
```r
dados <- c(4.391658, 5.364267, 10.707930, 5.431008, 6.904122, 6.960462,
 ↪  12.741468, 8.094473, 7.255829, 8.434530, 9.747057, 6.440681,
 ↪  7.623020, 9.276933, 8.711818, 5.250229, 6.482474, 3.478216,
 ↪  9.717008, 9.317296, 9.011653, 11.758927, 10.844472, 9.644711,
 ↪  7.541715, 7.561009, 10.034726, 9.654606, 6.222452, 5.207637)
id_dist(dados);
```



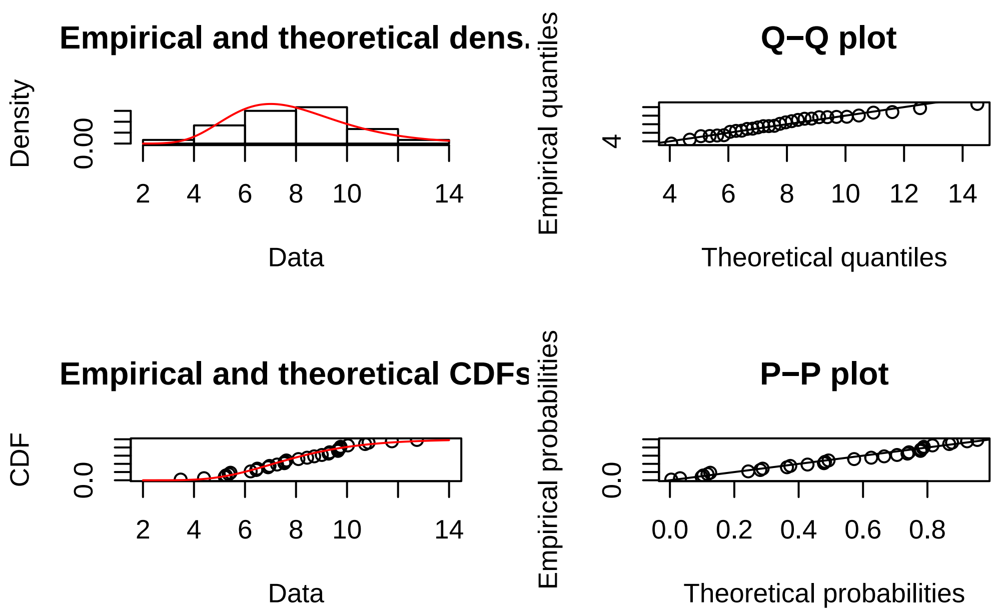
Weibull:



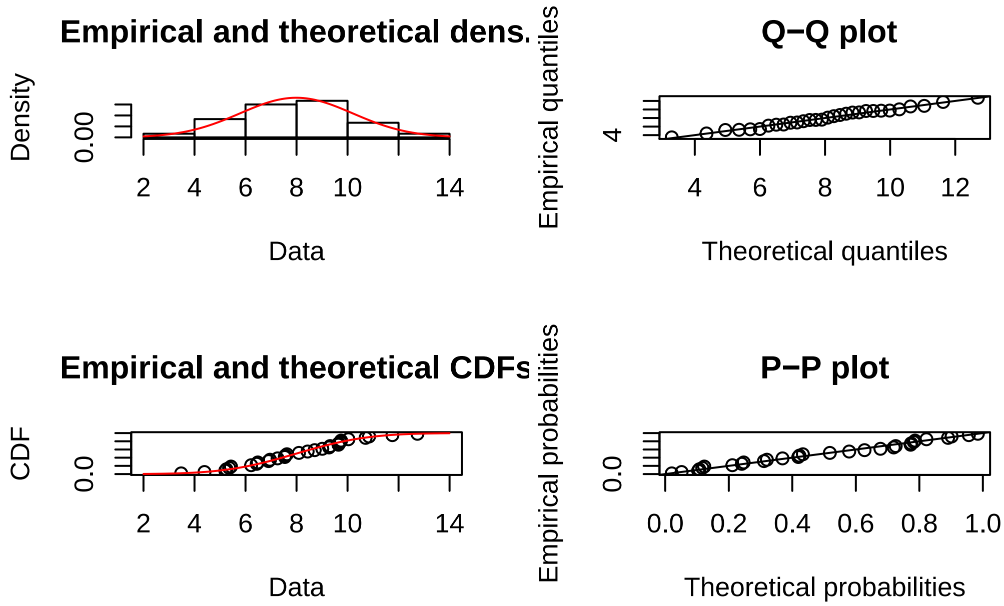
Gamma:



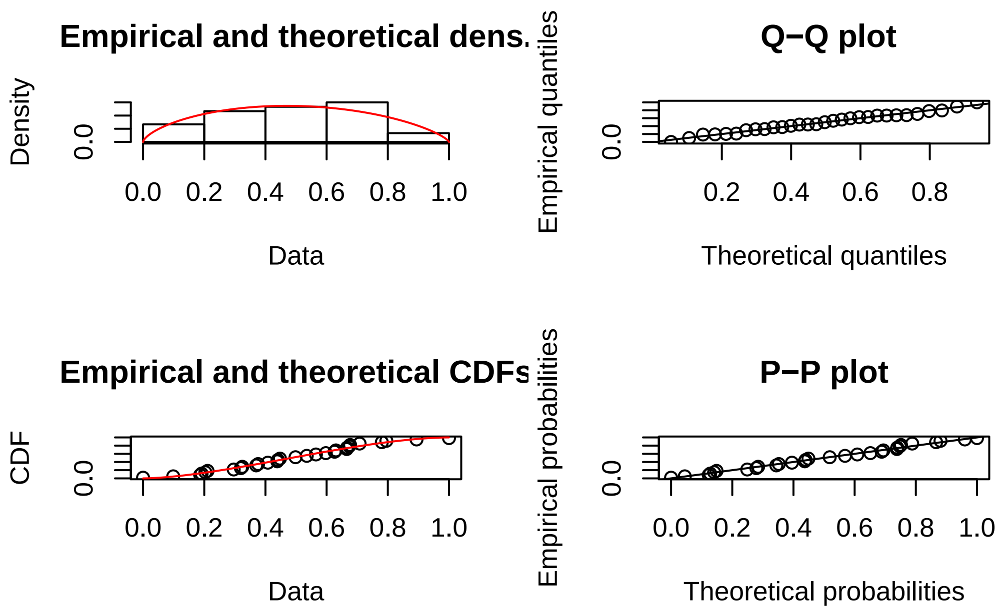
`Lognormal:`



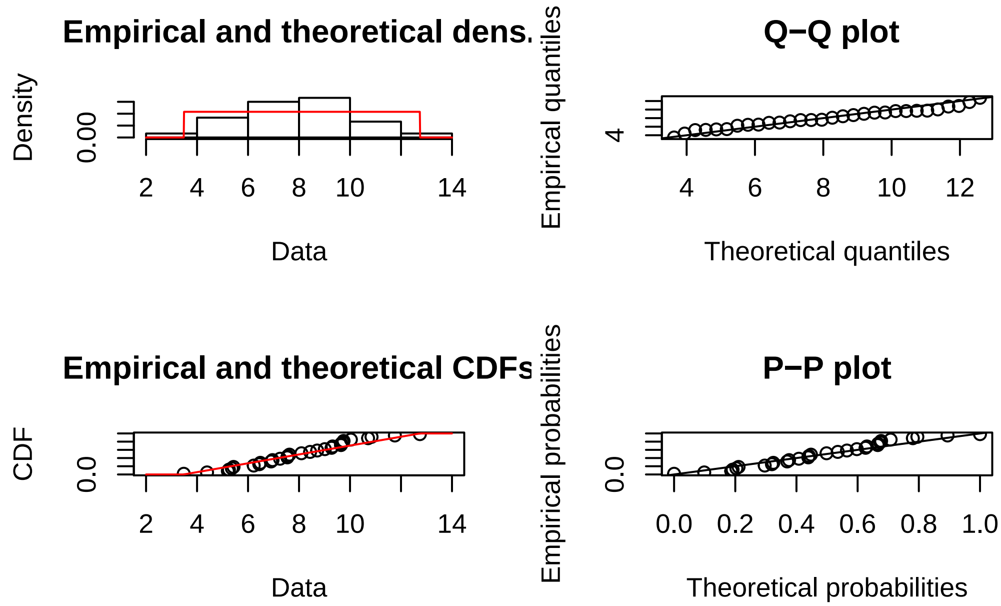
`Normal:`



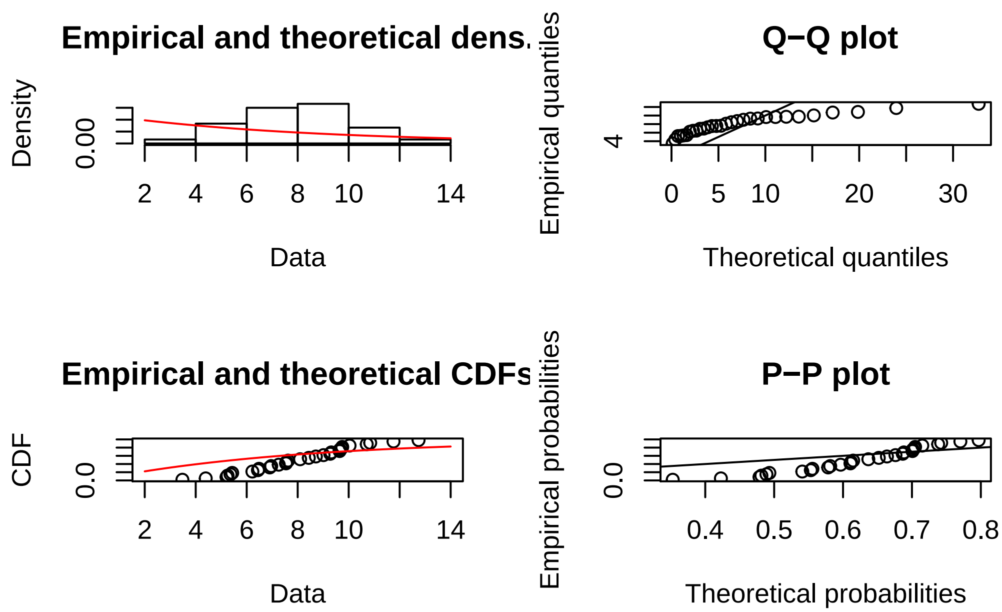

```
Beta:
```



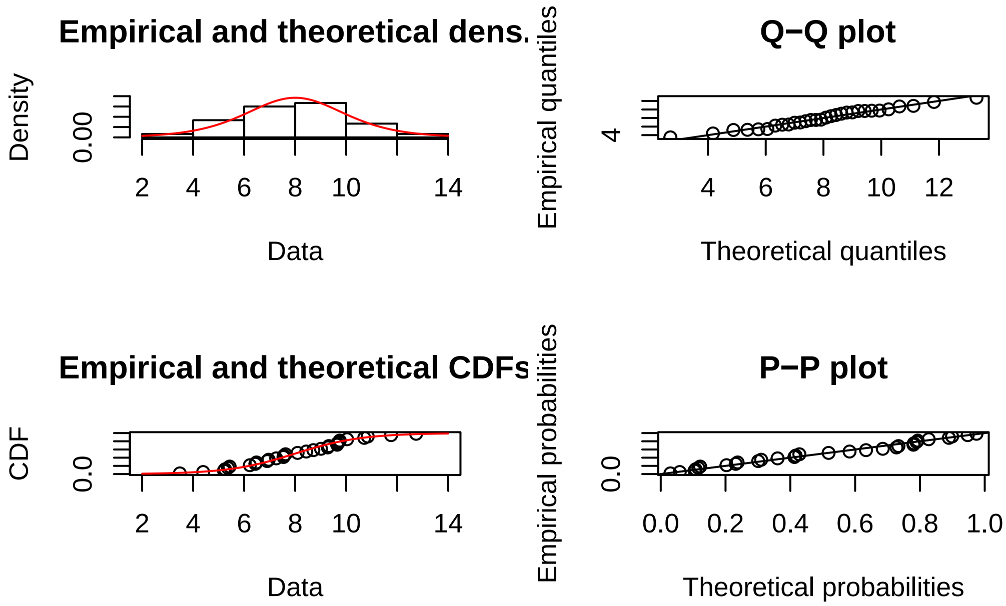

```
Uniforme:
```



Exponencial:



Logística:




```
Critério de Informação de Akaike
Weibull:  136.371061137677
Gamma:  137.524488802914
Lognormal:  139.025879465118
Normal:  136.6818515043
Beta:  Inf
Uniforme:  137.563310494661
Exponencial:  186.719570908607
Logística:  138.263602150104
```

A distribuição que apresenta menor Critério de Informação de Akaike é a **Weibull**. Portanto, realiza-se o teste de Kolmogorov-Smirnov e não se rejeita a hipótese de que os dados seguem a distribuição **Weibull**, com um nível de significância de 5%.

```
fitdist(dados, "weibull", method="mle")
```

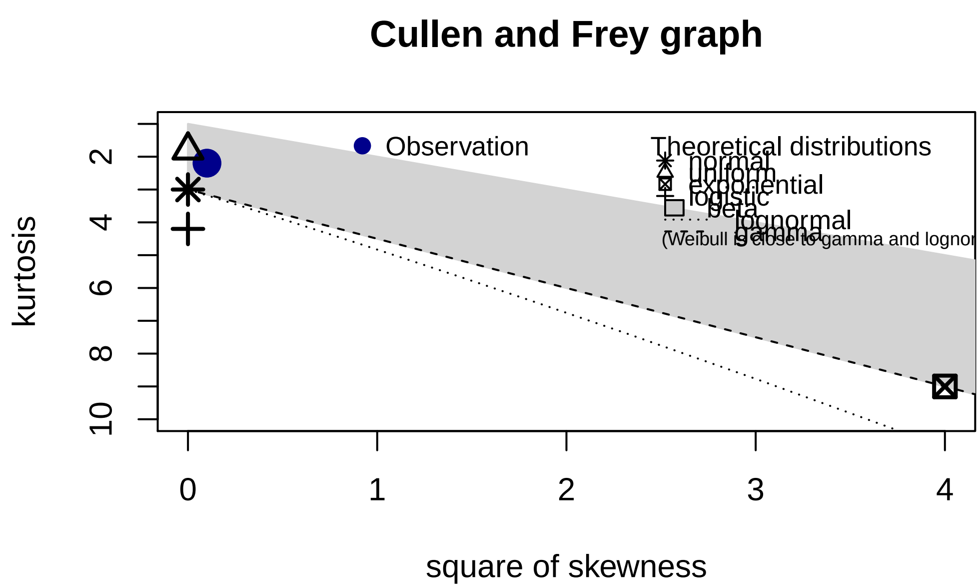
```
Fitting of the distribution ' weibull ' by maximum likelihood
Parameters:
      estimate Std. Error
shape 4.057284  0.5782002
scale 8.819386  0.4185798
```

```
ks.test(dados, "pweibull", 4.057284, 8.819386, exact=FALSE)
```

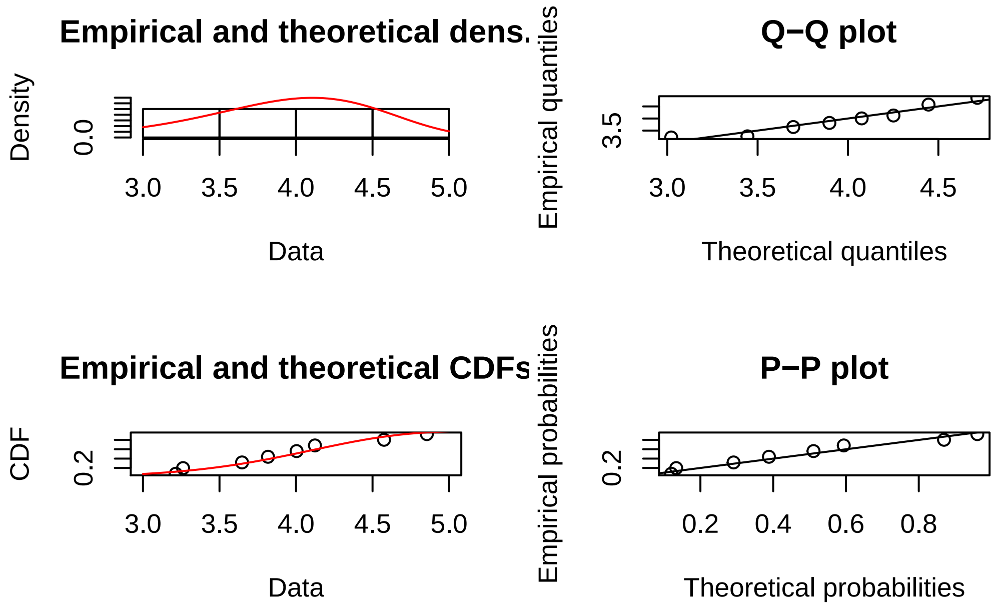
```
	One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.074926, p-value = 0.996
alternative hypothesis: two-sided
```

e)

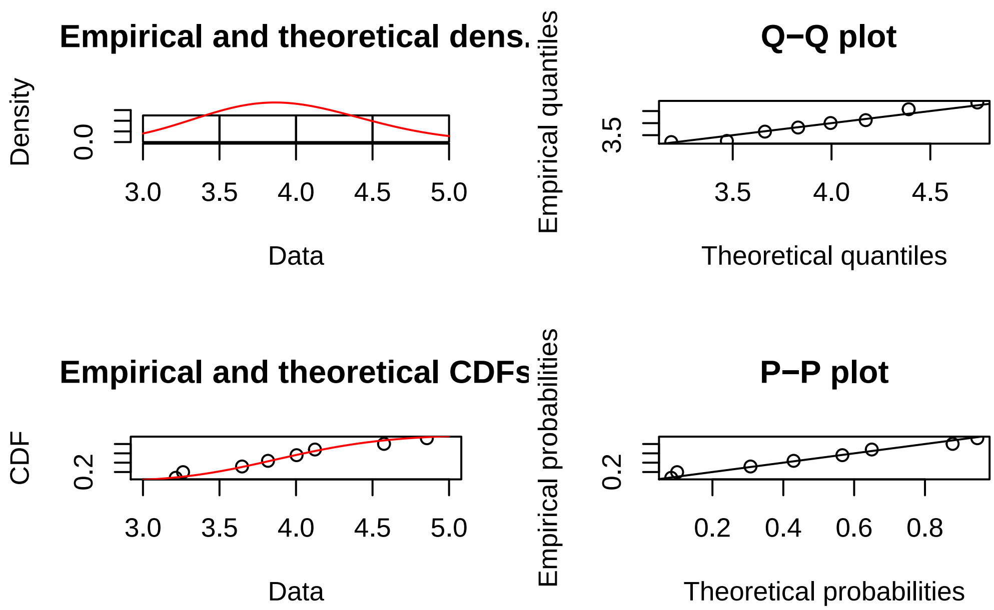
```
dados <- c(3.816942, 4.123619, 4.575150, 3.214129, 4.854917, 3.647232,
 ↪  4.003734, 3.261923)
id_dist(dados);
```



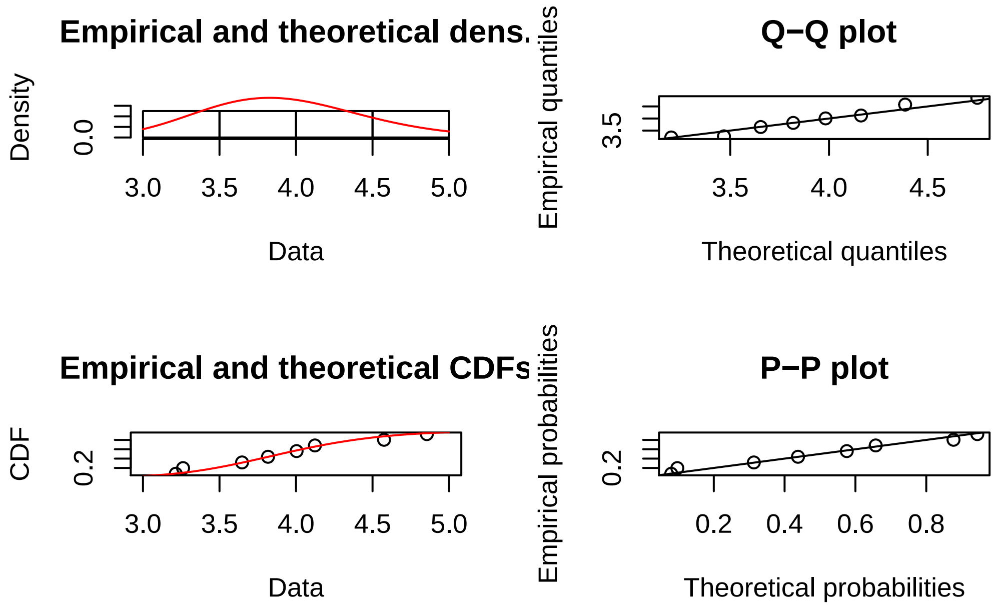

```
Weibull:
```



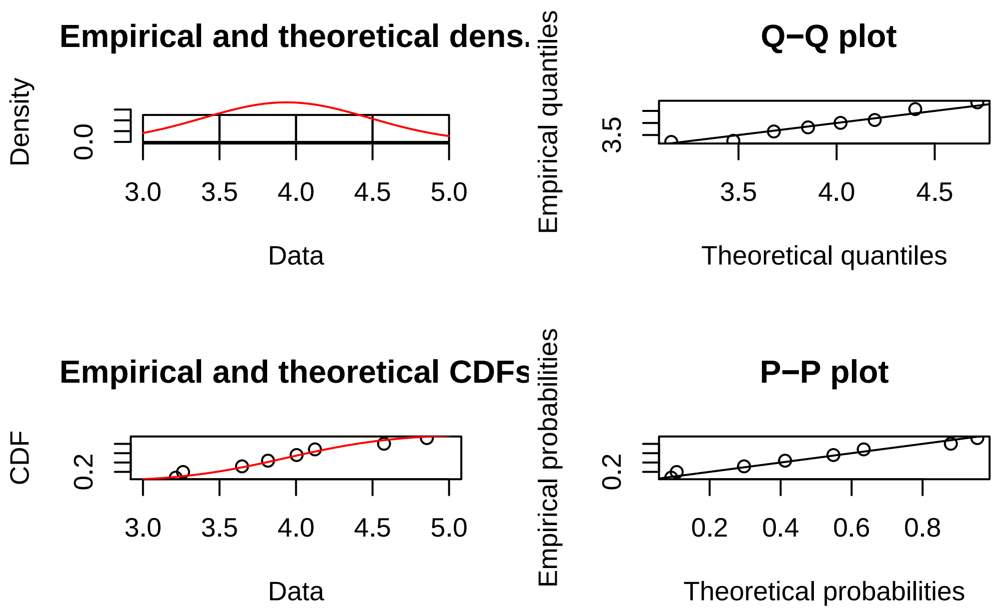
Gamma:



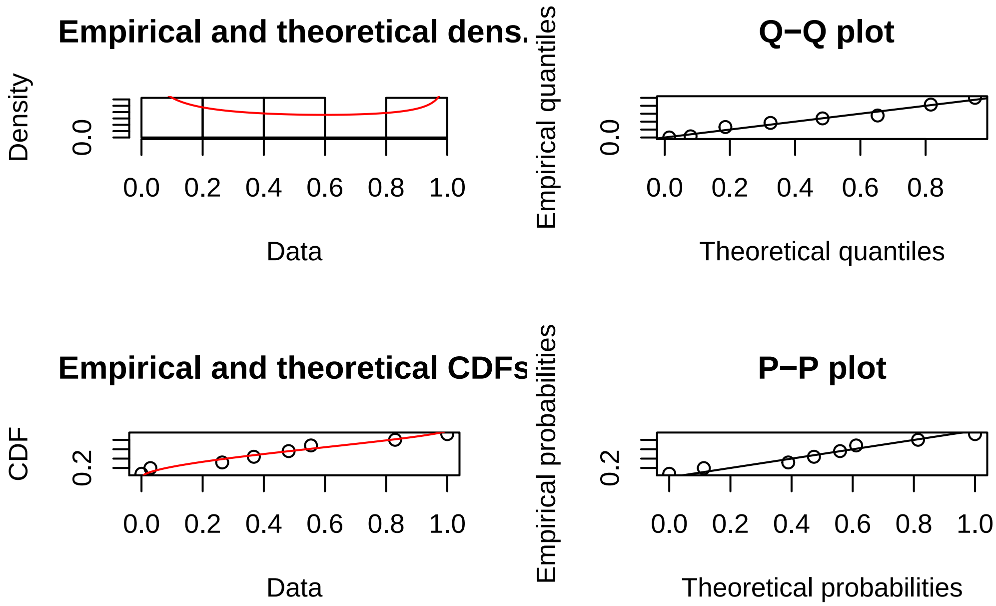
Lognormal:


Empirical and theoretical dens.
Density
0.0
3.0
3.5
4.0
4.5
5.0
Data
Q−Q plot
Empirical quantiles
3.5
3.5
4.0
4.5
Theoretical quantiles
Empirical and theoretical CDFs
CDF
0.2
3.0
3.5
4.0
4.5
5.0
Data
P−P plot
Empirical probabilities
0.2
0.2
0.4
0.6
0.8
Theoretical probabilities

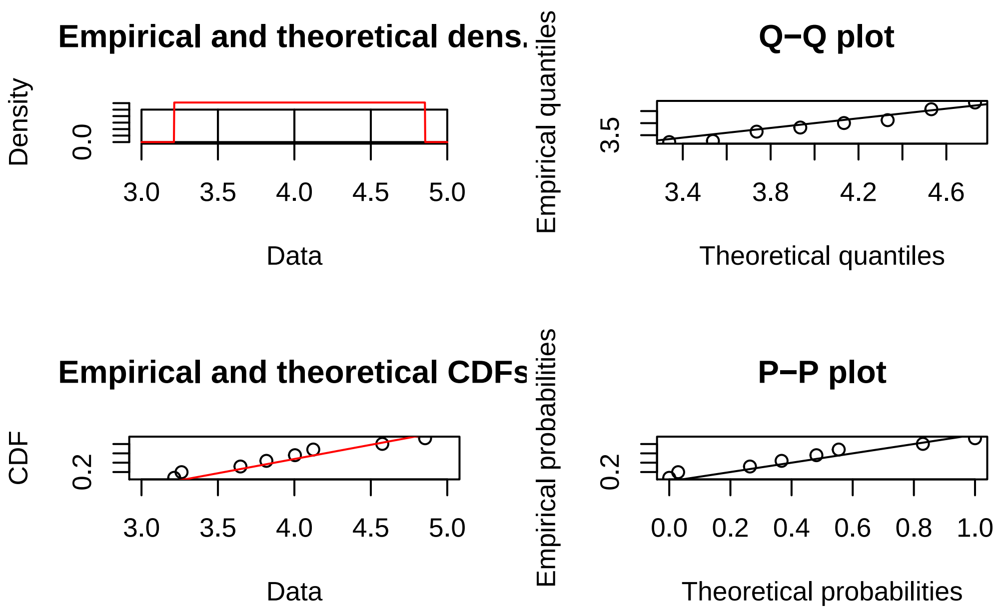
Normal:


Empirical and theoretical dens.
Density
0.0
3.0
3.5
4.0
4.5
5.0
Data
Q−Q plot
Empirical quantiles
3.5
3.5
4.0
4.5
Theoretical quantiles
Empirical and theoretical CDFs
CDF
0.2
3.0
3.5
4.0
4.5
5.0
Data
P−P plot
Empirical probabilities
0.2
0.2
0.4
0.6
0.8
Theoretical probabilities

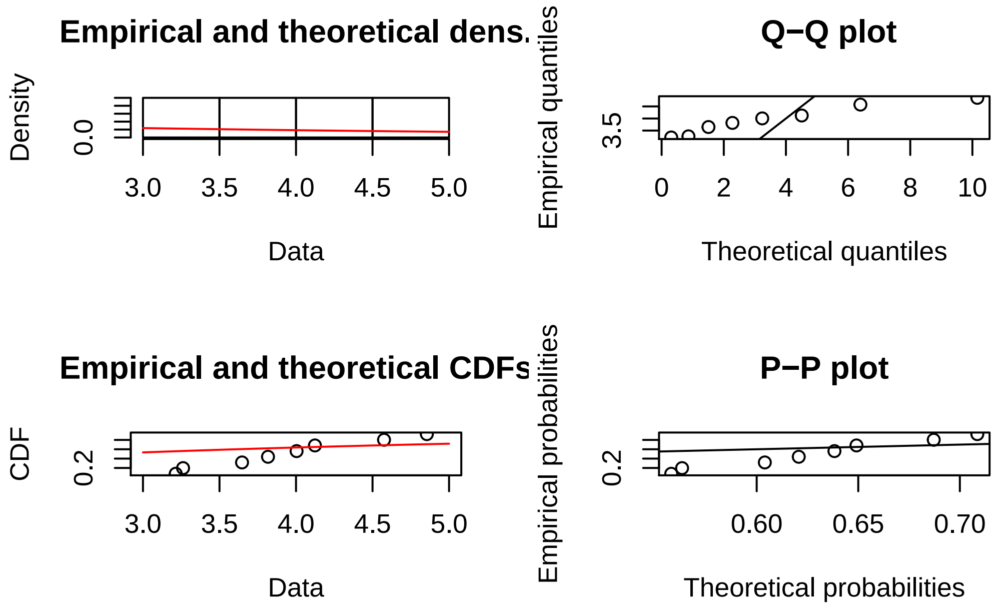
Beta:



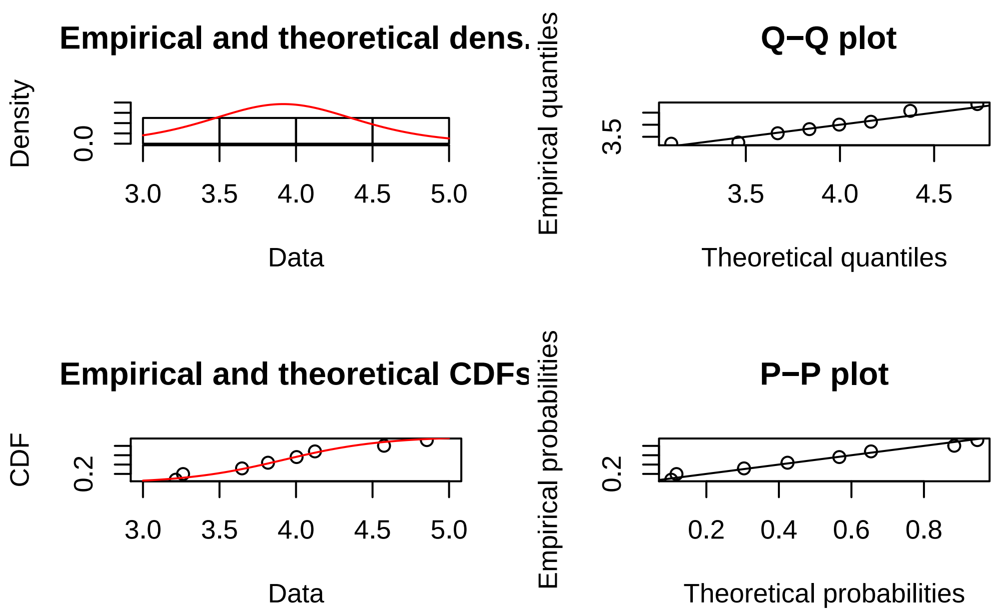
`Uniforme:`



`Exponencial:`




```
Logística:
```




```
Critério de Informação de Akaike
Weibull:  17.4848067022608
Gamma:  16.7857780075344
Lognormal:  16.7480979226041
```

```
Normal:  16.9562031603306
Beta:  -Inf
Uniforme:  11.9228258278989
Exponencial:  39.9275403392093
Logística:  17.5514378926189
```

A distribuição que apresenta menor Critério de Informação de Akaike é a **Uniforme**. Portanto, realiza-se o teste de Kolmogorov-Smirnov e não se rejeita a hipótese de que os dados seguem a distribuição **Uniforme**, com um nível de significância de 5%.

```r
fitdist(dados, "unif", method="mle")
```

```
Fitting of the distribution ' unif ' by maximum likelihood
Parameters:
    estimate Std. Error
min 3.214129         NA
max 4.854917         NA
```

```r
ks.test(dados, "punif", 3.214129, 4.854917, exact=FALSE)
```

```

    One-sample Kolmogorov-Smirnov test

data:  dados
D = 0.22087, p-value = 0.83
alternative hypothesis: two-sided
```