

## Testes de normalidade

```
## Testes de normalidade em R
```

```
library(MASS)
library(nortest)
```

O pacote MASS contém a função `fitdistr`, que será usada para obter as estimativas dos parâmetros da distribuição normal (e seus respectivos erros padrão). Testes de normalidade estão implementados no pacote `nortest`.

```
## Dados
```

```
# Densidade de amostras de concreto, em kg/m^3
# Kottegoda & Rosso (1997, p. 292)
```

```
dados <- c(2429, 2469, 2448, 2436, 2454, 2411, 2441, 2445, 2444, 2447,
           2448, 2437, 2488, 2445, 2435, 2428, 2455, 2471, 2456, 2441,
           2456, 2425, 2446, 2415, 2472, 2435, 2447, 2458, 2427, 2436,
           2450, 2449, 2457, 2437, 2433, 2436, 2427, 2454, 2473, 2436)
```

```
cat("\n Número de observações (n):", length(dados))
```

```
    Número de observações (n): 40
```

```
## Estimativas dos parâmetros
```

```
(estpar <- fitdistr(dados, "normal"))
```

```
      mean      sd
2444.925000 15.793017
(  2.497095) (  1.765713)
```

No resultado acima, `mean` e `sd` são as estimativas dos parâmetros da distribuição normal, que são a média e o desvio padrão, sendo que os valores entre parênteses são os erros padrão.

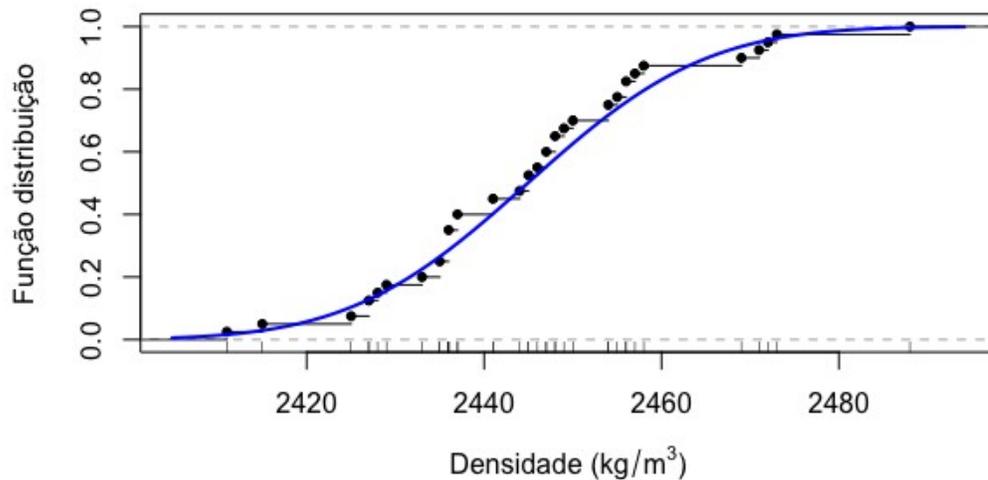
A função `names` mostra os componentes do objeto `estpar`.

```
names(estpar)
```

```
"estimate" "sd"      "vcov"      "n"      "loglik"
```

Nota 1. Procure entender o significado de cada um dos componentes acima.

```
## Função distribuição empírica e distribuição normal
plot(ecdf(dados), main = "", pch = 20,
     xlab = expression(paste("Densidade (", kg/m^3, ")")),
     ylab = "Função distribuição")
rug(dados)
curve(pnorm(x, mean = estpar$estimate[1], sd = estpar$estimate[2]),
     col = "blue", add = TRUE, lwd = 2)
```



```
## Teste de Anderson-Darling
ad.test(dados)
```

Anderson-Darling normality test

```
data: dados
A = 0.3877, p-value = 0.3714
```

```
## Teste de Cramér-von Mises
cvm.test(dados)
```

Cramer-von Mises normality test

```
data: dados
W = 0.058591, p-value = 0.3862
```

```
## Teste de Lilliefors
lillie.test(dados)
```

Lilliefors (Kolmogorov-Smirnov) normality test

```
data: dados
D = 0.089874, p-value = 0.5734
```

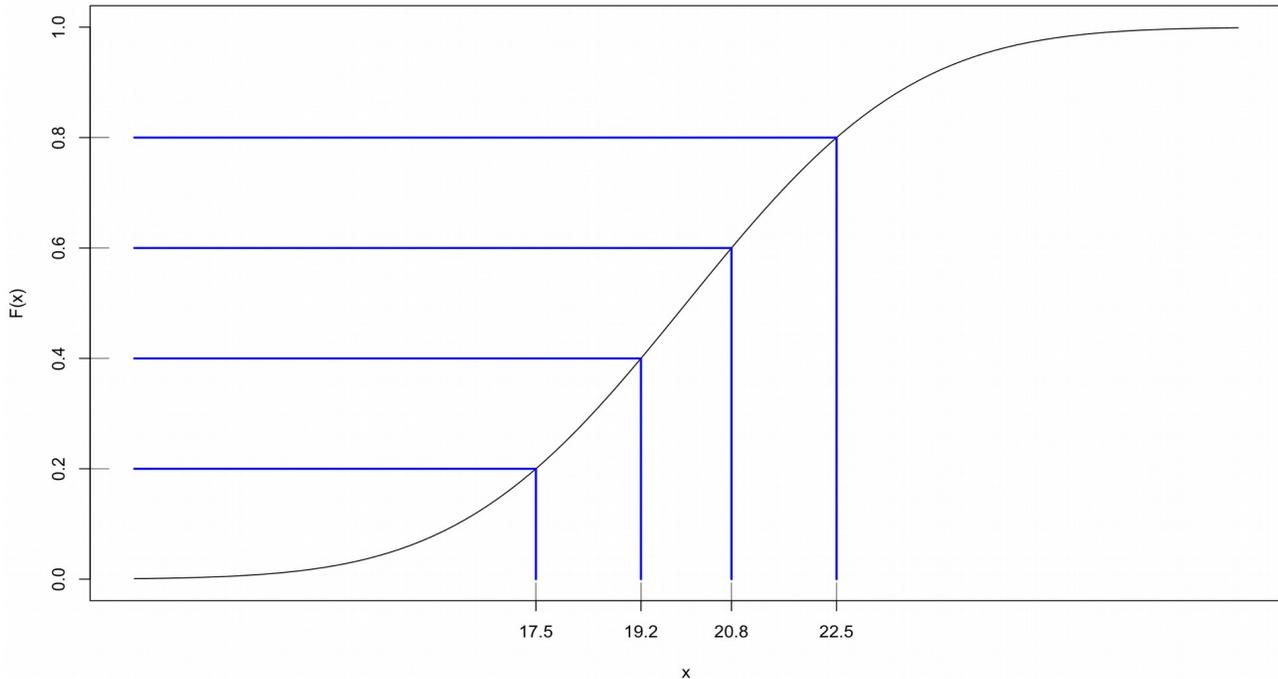
```
## Teste qui-quadrado de Pearson
pearson.test(dados)
```

Pearson chi-square normality test

```
data: dados
P = 10.4, p-value = 0.1088
```

Nota 2. O teste qui-quadrado de Pearson utiliza  $k = \lceil 2n^{2/5} \rceil$  classes equiprováveis, sendo que  $\lceil b \rceil$  denota o menor inteiro maior do que ou igual a  $b$  (função `ceiling` em R).

O gráfico abaixo mostra os cinco intervalos (probabilidade = 1/5) para uma distribuição normal com média 20 e desvio padrão 3.



Nota 3. Desenvolva um código em R para gerar o gráfico acima.

```
## Teste de Shapiro-Francia
sf.test(dados)

Shapiro-Francia normality test

data: dados
W = 0.97249, p-value = 0.3636
```

Nota 4. A estatística do teste de Shapiro-Francia é dada pelo quadrado do coeficiente de correlação entre as estatísticas de ordem da amostra e uma aproximação das estatísticas de ordem das distribuição normal padrão.

De acordo com os cinco testes realizados e com um nível de significância de 5%, não se rejeita a hipótese de que os dados de densidade de concreto seguem uma distribuição normal.

Nota 5. O comando `options(OutDec = ", ")` faz com que seja usada “,” no lugar de “.” como separador decimal nos resultados.