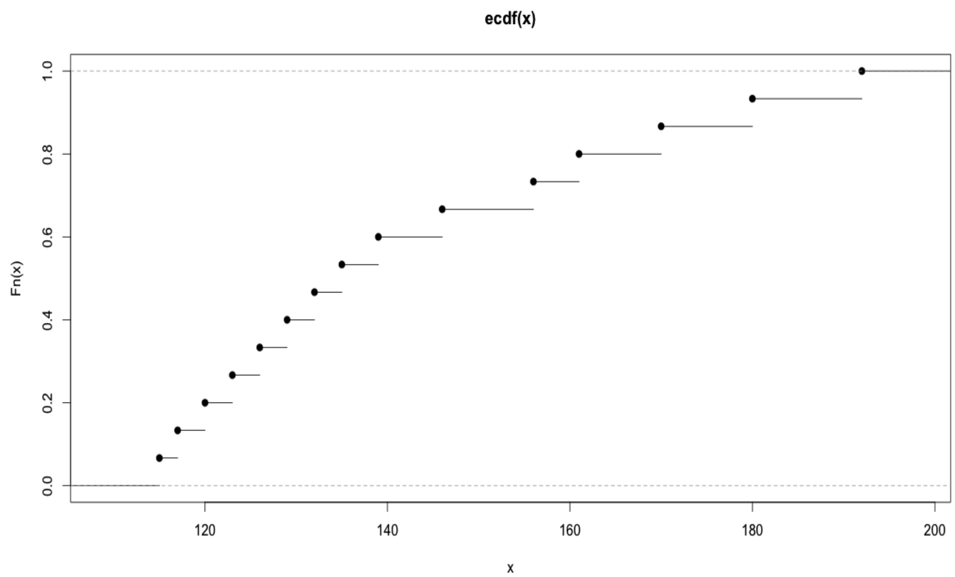
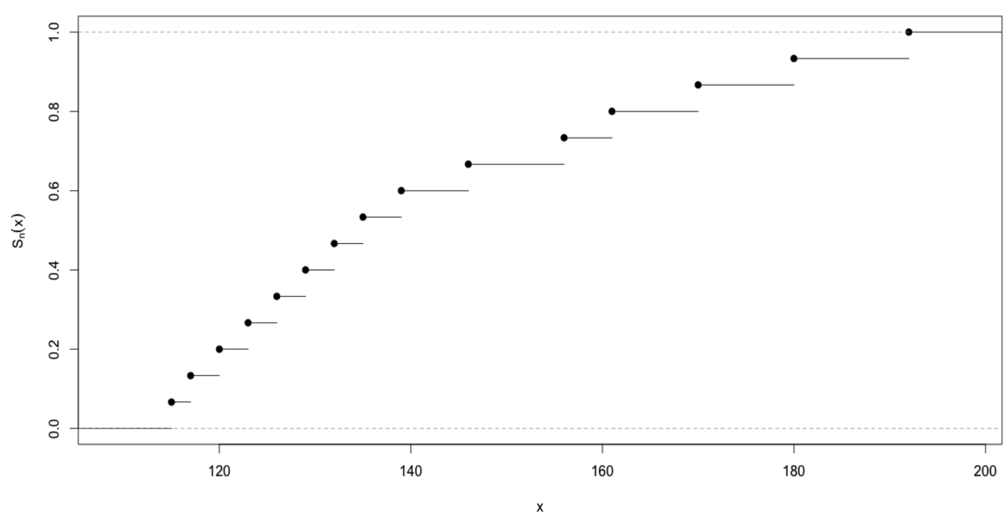
Teste de Kolmogorov-Smirnov

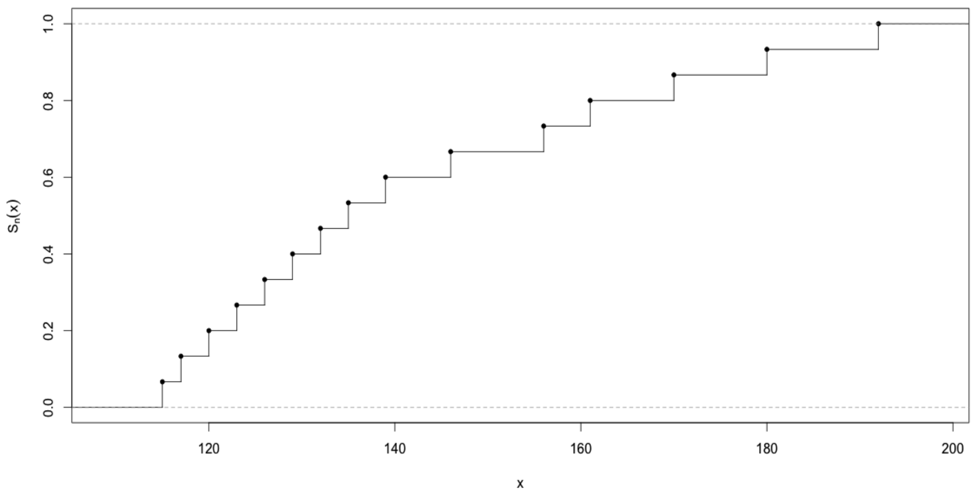
```
### Teste de Kolmogorov-Smirnov (KS)

##  1. Função distribuição empírica
# Dados
x <- c(126, 120, 117, 132, 146, 192, 180, 161, 156, 135, 129, 115, 170,
       139, 123)
cat("n = ", length(x))

n =  15
# Forma mais simples
plot(ecdf(x))
```




```
# Mudando alguns argumentos
plot(ecdf(x) , main = "", ylab = expression(S[n](x)))
```



```
plot(ecdf(x) , main = "", ylab = expression(S[n](x)), pch = 20,
     verticals = TRUE)
```



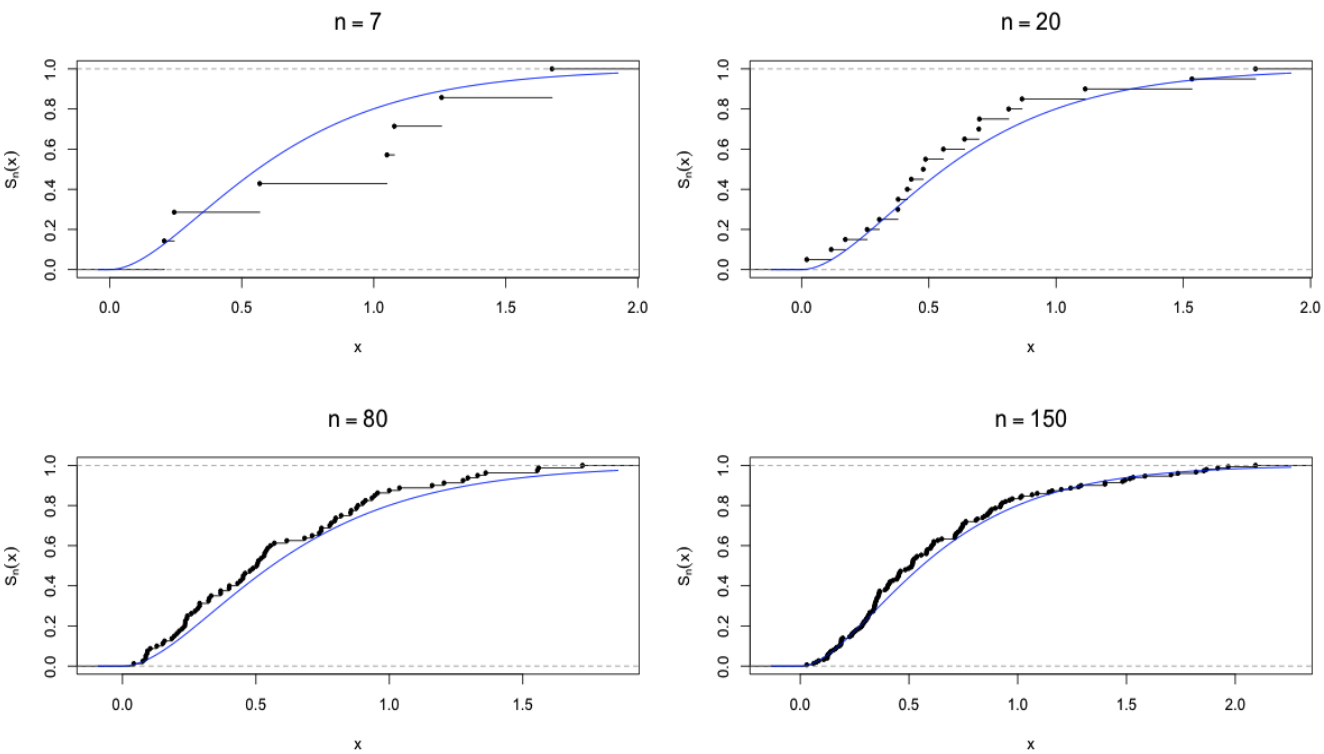

```
# Função Sn(x)
Sn <- ecdf(x)
plot(Sn , main = "", ylab = expression(S[n](x)), pch = 20)

knots(Sn)  # Valores de x sem repetições em ordem crescente

115 117 120 123 126 129 132 135 139 146 156 161 170 180 192
```

```
# Função Sn calculda em x
cbind(x, Sn(x))
 [1,] 126 0.33333333
 [2,] 120 0.20000000
 [3,] 117 0.13333333
 [4,] 132 0.46666667
 [5,] 146 0.66666667
 [6,] 192 1.00000000
 [7,] 180 0.93333333
 [8,] 161 0.80000000
 [9,] 156 0.73333333
[10,] 135 0.53333333
[11,] 129 0.40000000
[12,] 115 0.06666667
[13,] 170 0.86666667
[14,] 139 0.60000000
[15,] 123 0.26666667
```

```
# Função Sn calculda em x após ordenação
cbind(sort(unique(x)),Sn(sort(unique(x))))
 [1,]  115 0.06666667
 [2,]  117 0.13333333
 [3,]  120 0.20000000
 [4,]  123 0.26666667
 [5,]  126 0.33333333
 [6,]  129 0.40000000
 [7,]  132 0.46666667
 [8,]  135 0.53333333
 [9,]  139 0.60000000
[10,]  146 0.66666667
[11,]  156 0.73333333
[12,]  161 0.80000000
[13,]  170 0.86666667
[14,]  180 0.93333333
[15,]  192 1.00000000
```

```
Sn(c(-5, 118, 160, 200)) # Função Sn calculda em alguns pontos

0.0000000 0.1333333 0.7333333 1.0000000
```

```
# Exemplo com diferentes tamanhos de amostra
n <- c(7, 20, 80, 150)
par(mfrow = c(2, 2))
for (tamanho in n) {
   dados <- rgamma(tamanho, shape = 2, rate = 3)
   plot(ecdf(dados) , main = bquote(n == .(tamanho)),
        ylab = expression(S[n](x)), pch = 20, cex.main = 1.5)
   curve(pgamma(x, shape = 2, rate = 3), add = TRUE, col = "blue")
}
```



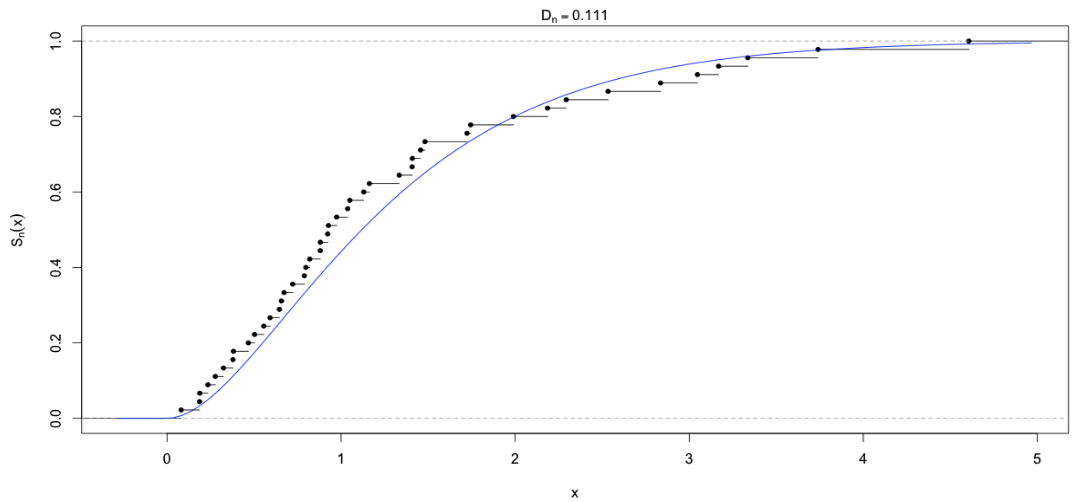
2. Teste KS

```
# Dados
# X ~ gama(forma = f0, taxa = t0)
f0 <- 2
t0 <- 1.5
n <- 45
dados <- rgamma(n, shape = f0, rate = t0)
```

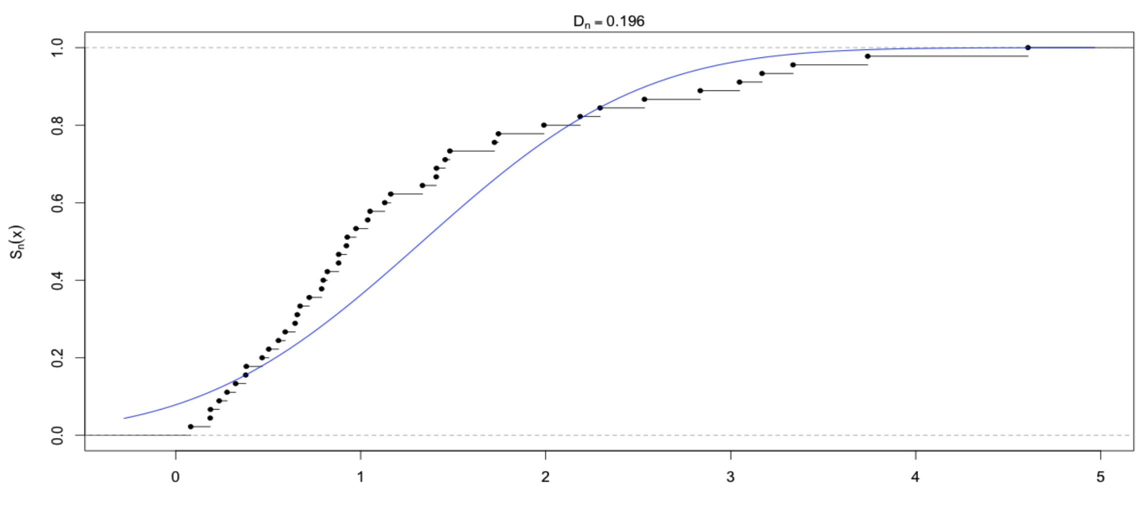
```
# H0: X ~ gama(forma = f0, taxa = t0)
# Default: H1 bilateral e valor-p exato
(tks <- ks.test(dados, "pgamma", shape = f0,
rate = t0))

     One-sample Kolmogorov-Smirnov test
data:  dados
D = 0.1106, p-value = 0.6019
alternative hypothesis: two-sided
```

```
# valor-p aproximado
ks.test(dados,      "pgamma",
shape = f0, rate = t0, exact
= FALSE)

D = 0.1106, p-value = 0.6411
alternative hypothesis: two-
sided
```

```
# Gráficos
plot(ecdf(dados) , main = "", ylab = expression(S[n](x)), pch = 20)
curve(pgamma(x, shape = f0, rate = t0), add = TRUE, col = "blue")
mtext(bquote(D[n] == .(round(tks$statistic, digits = 3))))
```




```
# H0: X ~ normal(média = f0 / t0, variância = f0 / t0^2)
# Normal com mesma média e mesma variância da dist. gama
# Default: H1 bilateral e valor-p exato
(tksn <- ks.test(dados, "pnorm", mean = f0 / t0, sd = sqrt(f0 / t0^2)))

      One-sample Kolmogorov-Smirnov test
data:  dados
D = 0.1957, p-value = 0.05517
alternative hypothesis: two-sided

plot(ecdf(dados) , main = "", ylab = expression(S[n](x)), pch = 20)
curve(pnorm(x, mean = f0 / t0, sd = sqrt(f0 / t0^2)), add = TRUE,
  col = "blue")
mtext(bquote(D[n] == .(round(tksn$statistic, digits = 3))))
```



Nota. Refaça o teste da hipótese de normalidade aumentando o valor do parâmetro de forma (f_0). Surpresa?