

Intervalos de confiança *bootstrap*

Apresentamos um exemplo com o coeficiente de correlação linear de Pearson. Serão obtidos intervalos de confiança baseados nos percentis e no método *t bootstrap*.

```
# Coeficiente de confiança e probabilidades
conf <- 0.95
a <- 1 - conf
limi <- a / 2
lims <- 1 - limi

# Dados de medições angulares por dois métodos
X <- c(100, 58, 95, 55, 79, 95, 60, 88, 68, 94, 60, 64, 88, 57, 66, 67,
      76, 95, 85, 105, 80, 85, 82, 102, 100, 75, 40, 70, 63, 103, 95,
      80, 72, 68, 48, 70, 90, 60, 80, 96, 54, 80, 88, 70, 90, 79, 100,
      85, 108, 53, 58, 49)
Y <- c(97, 77, 74, 59, 79, 85, 78, 78, 68, 96, 74, 64, 76, 60, 78, 71,
      67, 103, 95, 78, 70, 80, 78, 102, 102, 77, 45, 60, 50, 94, 91,
      66, 63, 65, 58, 75, 105, 65, 80, 90, 58, 75, 83, 78, 85, 65, 90,
      76, 100, 65, 40, 53)
```

```
n <- length(X)
# Identificação das observações
indices <- 1:n

cat("\n Tamanho da amostra:", n, "\n")
```

Tamanho da amostra: 52

```
# Estimativa
tetac <- cor(X, Y)
```

Estimativa do coef. de correlação de Pearson: 0.8196273

Nota 1. Represente graficamente os dados.

```
# Número de reamostras (1o. e 2o. níveis)
B <- 1000
B2 <- 500

# Índices das amostras bootstrap (1o. nível)
ind1 <- matrix(sample(indices, n * B, replace = TRUE), ncol = n)

# Estimativas bootstrap
tetace <- c()
for (b in 1:B) {
  tetace[b] <- cor(X[ind1[b,]], Y[ind1[b,]])
}
epe <- sd(tetace)
cat("\n Erro padrão bootstrap:", epe)
```

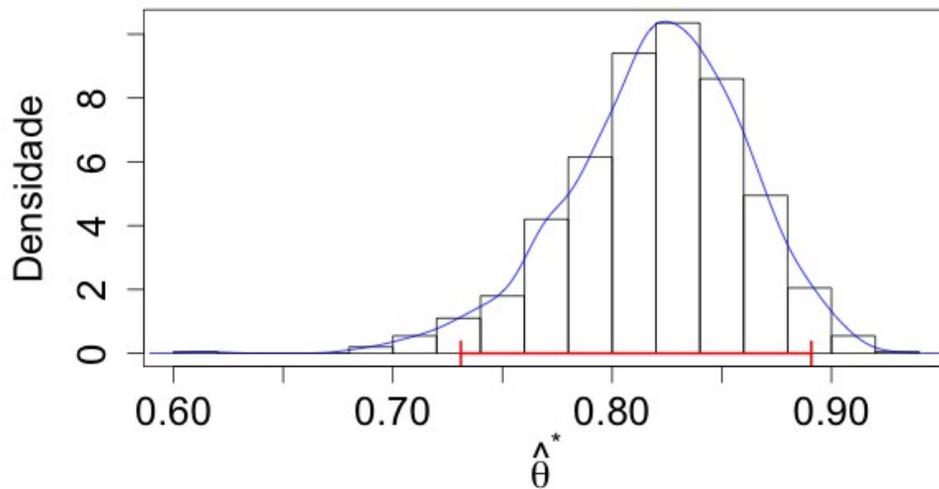
Erro padrão bootstrap: 0.04043442

Primeiro apresentamos o intervalo de confiança (IC) percentil.

```
## Intervalo de confiança percentil
ICperc <- quantile(tetace, probs = c(limi, lims), type = 6)
cat("\n IC percentil de", 100 * conf, "%:", ICperc)
```

```
IC percentil de 95 %: 0.7310086 0.8908386
```

```
par(mai = c(1.2, 1.2, 0.1, 0.3))
hist(tetace, freq = FALSE, main = "", xlab = expression(hat(theta)^"*"),
     ylab = "Densidade", cex.axis = 1.5, cex.lab = 1.5)
lines(density(tetace), col = "blue")
arrows(ICperc[1], 0, ICperc[2], 0, col = "red", code = 3, angle = 90,
       lwd = 2, length = 0.1)
box()
```



Nota 2. Apresente uma estimativa do viés do estimador para este exemplo.

Para construir o intervalo *t bootstrap* primeiro calculamos t^*_b , $b = 1, \dots, B$. A função `system.time` é utilizada para obter o tempo envolvido (em segundos) neste cálculo.

```
# Iteração (2o. nível)
te <- c()
system.time(
for (b in 1:B) {
  tetacee <- c()
  for (j in 1:B2) {
    ind2 <- sample(ind1[b,], n, replace = TRUE)
    tetacee[j] <- cor(X[ind2], Y[ind2])
  }
  epeb <- sd(tetacee)
  te[b] <- (tetace[b] - tetac) / epeb
}
)
```

```

user  system elapsed
42.444  0.615  43.214

```

Em seguida apresentamos uma solução em que o laço `for` mais interno (vide acima) é evitado. A função abaixo efetua o cálculo do coeficiente de correlação de Pearson entre X e Y em B_0 reamostras obtidas do vetor de índices `ind0`.

```

estimates <- function(B0, ind0) {
  corXY <- function(ind10) {
    cor(X[ind10], Y[ind10])
  }
  ind1 <- matrix(sample(ind0, n * B0, replace = TRUE), ncol = n)
  apply(ind1, MARGIN = 1, FUN = corXY)
}

```

Chamando a função `estimates` com $B_0 = B$ e `ind0 = indices` (ou seja, identificação das observações na amostra), obtemos o vetor de estimativas bootstrap (`tetace`, pag. 1).

```

# Estimativas bootstrap
tetace <- estimates(B, indices)
epe <- sd(tetace)

# Cálculo de  $t^*$ 
te <- c()
system.time(
  for (b in 1:B) {
    tetacee <- estimates(B2, ind1[b,])
    epeb <- sd(tetacee)
    te[b] <- (tetace[b] - tetac) / epeb
  }
)

```

```

user  system elapsed
36.755  0.452  37.393

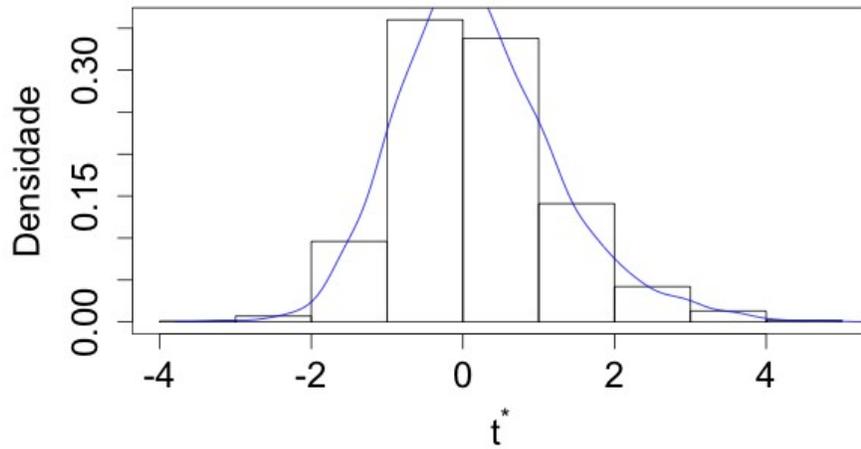
```

A distribuição de t^* é usada para obter os limites do intervalo de confiança t bootstrap.

```

hist(te, freq = FALSE, main = "", xlab = expression(t^"*"),
     ylab = "Densidade", cex.axis = 1.5, cex.lab = 1.5)
lines(density(te), col = "blue")
box()

```



Nota 3. Modifique o código para que a função densidade estimada (linha azul) seja visível por completo.

```
# IC t bootstrap
qte <- quantile(te, probs = c(lims, limi), type = 6)
ICt <- tetac - qte * epe
cat("\n IC t bootstrap de", 100 * conf, "%:", ICt)
```

IC t bootstrap de 95 %: 0.709899 0.8840748

