

# Hugo Manuel Alves Henriques e Silva - 99235

## Exercício 6

```
library(tidyr)
library(ggplot2)
library(dplyr)

n1 = 2; n2 = 29; n3 = 50; a = 0; b = 4; seed = 678; nAmostras = 260; start = 1
expectedVal = (a + b) / 2; variance = (b - a)^2 / 12; set.seed(678)

df1 = data.frame(c(1 : nAmostras)); df2 = data.frame(c(1 : nAmostras)); df3 = data.frame(c(1 : nAmostras))
names(df1)[1] = "means1"; names(df2)[1] = "means2"; names(df3)[1] = "means3"

for(num in start : nAmostras) {
  mean1 = mean(runif(n1,min = 0,max = 4))
  df1[num, 1] = mean1

  mean2 = mean(runif(n2, min = 0, max = 4))
  df2[num, 1] = mean2

  mean3 = mean(runif(n3, min = 0, max = 4))
  df3[num, 1] = mean3
}

hist1 = ggplot(df1, aes(means1, color = "means1"))
hist1 = hist1 + geom_histogram(aes(y = ..density..), fill="white", alpha = 0.5, binwidth = 0.145) +
  stat_function(fun = function(x) dnorm(x, mean = expectedVal, sd = sqrt(variance/n1))) +
  scale_y_continuous("Density", sec.axis = sec_axis(trans = ~ . * 0.145, name = "Freq. Relativa", labels =
scales::percent))
hist1 + xlab("valores obtidos da distribuição da média") + ggtitle("Histograma de frequência relativa para n = 2")

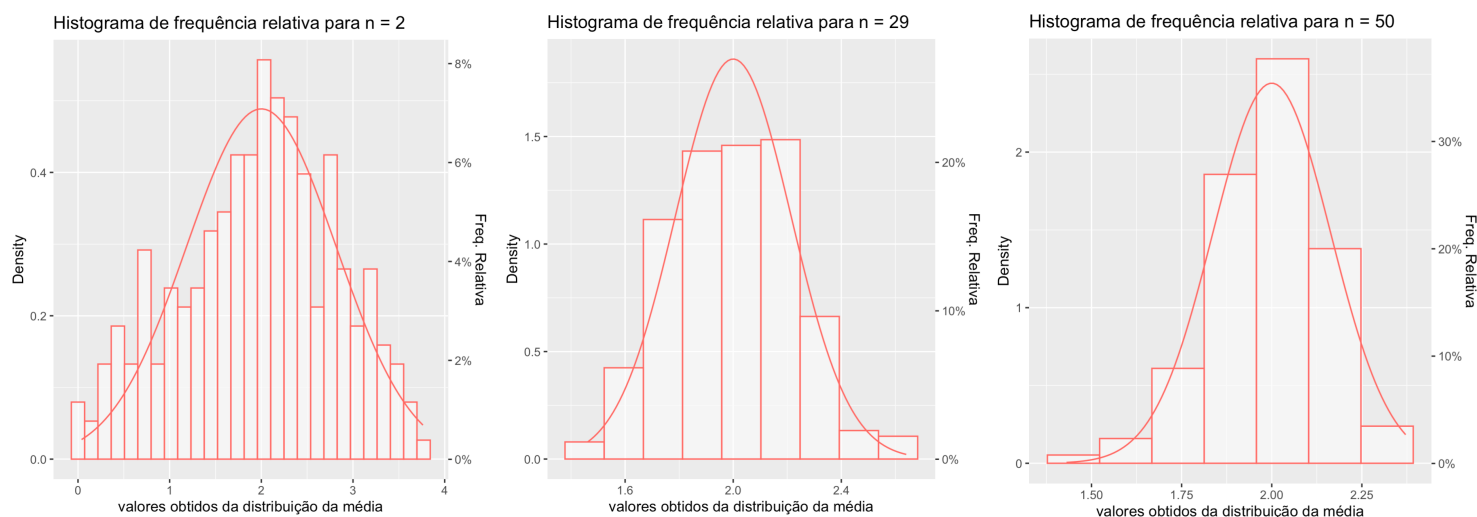
hist2 = ggplot(df2, aes(means2, color = "means2"))
hist2 = hist2 + geom_histogram(aes(y = ..density..), fill="white", alpha=0.5, binwidth = 0.145) +
  scale_y_continuous("Density", sec.axis = sec_axis(trans = ~ . * 0.145, name = "Freq. Relativa", labels =
scales::percent)) +
  stat_function(fun = function(x) dnorm(x, mean = expectedVal, sd = sqrt(variance/n2))) + ylab("Frequência
Relativa")
hist2 + xlab("valores obtidos da distribuição da média") + ggtitle("Histograma de frequência relativa para n = 29")

hist3 = ggplot(df3, aes(means3, color = "means3"))
hist3 = hist3 + geom_histogram(aes(y = ..density..), fill="white", alpha=0.5, binwidth = 0.145) +
  scale_y_continuous("Density", sec.axis = sec_axis(trans = ~ . * 0.145, name = "Freq. Relativa", labels =
scales::percent)) +
  stat_function(fun = function(x) dnorm(x, mean = expectedVal, sd = sqrt(variance/n3))) + ylab("Frequência
Relativa")
hist3 + xlab("valores obtidos da distribuição da média") + ggtitle("Histograma de frequência relativa para n = 50")
```

Semente = 678

Dimensões das amostras: n1 = 2, n2 = 29, n3 = 50.

Parâmetros da distribuição uniforme: min = 0, max = 4.



Tal como indica o TLC, as médias das amostras das diferentes dimensões seguem uma distribuição normal tal como observável pelos 3 gráficos. Tal como representa a curva da distribuição normal, a frequência relativa é maior para os valores no centro da mesma e à medida que se afastam do centro a frequência relativa diminui. Além disso, a soma total das percentagens apresentadas em cada barra dos gráficos corresponde a 100% tal como a área que se situa por baixo da curva da distribuição normal. Para o cálculo da curva da distribuição normal foi usado o valor esperado  $(b+a)/2$ , neste caso  $(4+0)/2$  e a variância  $(b-a)^2 / (12*n)$  ou seja,  $(4-0)^2 / (12*n)$  de onde se conclui que a variância é inversamente proporcional ao tamanho das amostras, daí a diminuição da dispersão dos valores com o aumento do tamanho da amostra que se verifica do primeiro gráfico para o segundo e ainda do segundo para o terceiro.