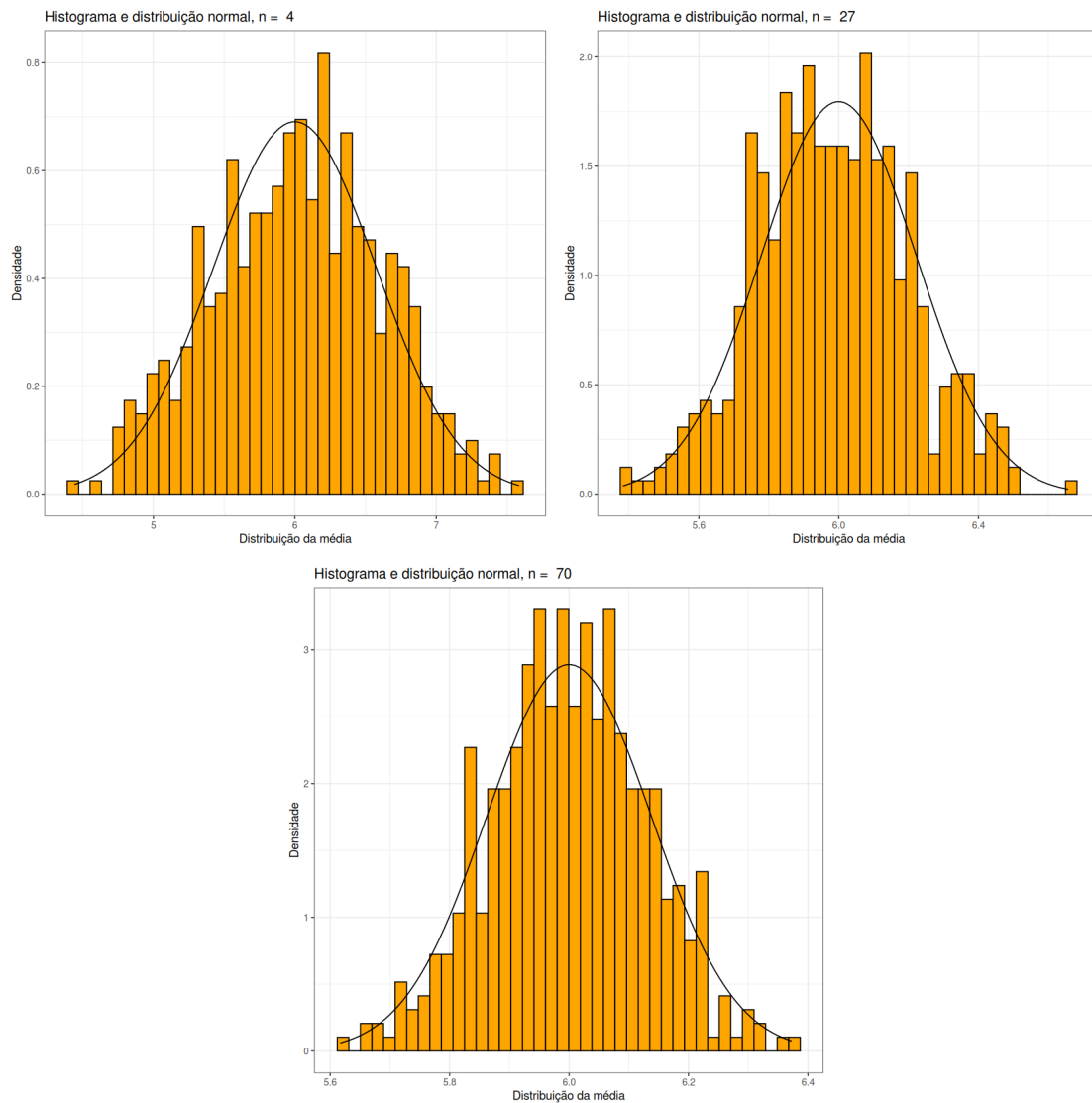


Exercício 6 - Projeto Computacional PE 2022

```
set.seed(1896)
n_values <- c(4, 27, 70)
sample_amount <- 500
a <- 4
b <- 8
expected_value <- (b + a) / 2

calc_sd <- function(n) {
  return (sqrt(((b - a)**2) / (n * 12)))
}
calc_means <- function(n) {
  means <- c()
  for (i in 1:sample_amount) {
    means <- c(means, mean(runif(n, 4, 8)))
  }
  return(means)
}

for (n in n_values) {
  means <- calc_means(n)
  df <- data.frame(means)
  print(
    ggplot(df, aes(x = means)) +
    geom_histogram(aes(y=..density..), color="black", fill="orange",
bins=40) +
    stat_function(fun=dnorm, args=list(mean=expected_value, sd=calc_sd(n)))
+
    theme_bw() +
    labs(x = "Distribuição da média", y = "Densidade") +
    ggtitle(paste("Histograma e distribuição normal, n = ", n))
  )
}
```



Observando o gráfico produzido pela chamada a `ggplot`, podemos observar que aparenta haver uma concentração algo forte de utentes com valores de colesterol entre os 200 e 240, começando, para fora desse intervalo, a haver um número mais esparsos de utentes. Mais ainda, olhando para a curva ilustrada no gráfico, aparenta haver uma relação (ainda que não muito forte) entre a idade e os valores de colesterol dos utentes: à medida que a idade avança, os níveis de colesterol vão, em média, aumentando gradualmente.