

Exercício 9 - Projeto Computacional PE 2022

Diogo Gaspar, 99207

Consideremos como premissas que foram fixadas uma semente em 139, um conjunto de tamanhos de amostras {100, 200, ..., 5000}. O objetivo deste exercício passa por gerar 650 amostras com distribuição exponencial de valor esperado $\frac{1}{\lambda} = \frac{1}{1.32}$ para cada tamanho supra-mencionado. De seguida, construir para cada uma das amostras geradas um intervalo de confiança aproximado para λ (com intervalo de confiança $1 - \alpha = 0.93$), e para cada tamanho de amostra calcular a média da amplitude de todos os intervalos de confiança obtidos.

Para tal, recorreu-se ao seguinte trecho de código R (utilizando as bibliotecas `ggplot2`, `lattice`, `plyr` e `Rmisc`):

```
1  set.seed(139)
2  m <- 650
3  lambda <- 1.32
4  theoric_confidence_interval <- 0.93
5  dimensions <- seq(100, 5000, 100)
6
7  calculate_mean_widths <- function(n) {
8    widths <- c()
9    for (i in 1:m) {
10      exp <- rexp(n, rate=lambda)
11      confidence_interval <- CI(exp, ci=theoric_confidence_interval)
12      widths <- c(widths, c(abs(confidence_interval[["upper"]] - confidence_interval[["lower"]]))))
13    }
14    return (mean(widths))
15  }
16
17  mean_widths <- c()
18  for (n in dimensions) {
19    mean_widths <- c(mean_widths, calculate_mean_widths(n))
20  }
21
22  ggplot(data.frame(dimensions, mean_widths), aes(x=dimensions, y=mean_widths)) +
23    geom_line(color="blue") +
24    geom_point(color="blue") +
25    theme_bw() +
26    theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
27    xlab("Dimensão da amostra") + ylab("Amplitude média para 650 amostras") +
28    ggtitle("Amplitude média dos intervalos de confiança da distribuição exponencial") +
29    scale_x_continuous(breaks = round(seq(0, 5000, 500)))
```

Note-se que à medida que o tamanho da amostra aumenta, a amplitude média entre intervalos de confiança torna-se rapidamente mais pequena à medida que nos aproximamos de 5000. Podemos, portanto, retirar deste gráfico que quanto maior o tamanho da população, mais podemos **confiar** nos resultados obtidos, visto que vemos a amplitude média dos intervalos de confiança a ser reduzida.

