

Hugo Manuel Alves Henriques e Silva - 99235

Exercício 10

```
seed = 558; set.seed(seed); nAmostras = 550; lambda = 1.8; lambda2 = 0.12; nN = 25; alpha = 0.09; start = 1; p = 0.85
df1 = data.frame(matrix(nrow = nN, ncol = 2)) ; names(df1)[1] = "Ns"; names(df1)[2] = "Medias"
df3 = data.frame(matrix(nrow = nN, ncol = 2)) ; names(df3)[1] = "Ns"; names(df3)[2] = "Medias"
aux = 0
for(num in start : nN) {
  aux = aux + 100
  df1[num, 1] = aux
  df3[num, 1] = aux
}
df2 = data.frame(c(1 : nAmostras)) ; names(df2)[1] = "Errors"
df4 = data.frame(c(1 : nAmostras)) ; names(df4)[1] = "Errors"
dim = 0
for (i in start : nN) {
  dim = dim + 100
  for(num in start : nAmostras) {
    amostra = rexp(dim, lambda)
    minA = (1 - qnorm(1 - alpha/2) / sqrt(dim)) / mean(amostra); maxA = (1 + qnorm(1 - alpha/2) / sqrt(dim)) /
mean(amostra)
    errorAmp = maxA - minA; df2[num, 1] = errorAmp
    amostraAux = rexp((1-p) * dim, lambda2)
    amostra2 = c(amostra[1:(p*dim)], amostraAux)
    minA2 = (1 - qnorm(1 - alpha/2) / sqrt(dim)) / mean(amostra2); maxA2 = (1 + qnorm(1 - alpha/2) / sqrt(dim)) /
mean(amostra2)
    errorAmp2 = maxA2 - minA2; df4[num, 1] = errorAmp2
  }
  df1[i, 2] = mean(df2$Errors); df3[i, 2] = mean(df4$Errors)
}
ggplot() + geom_point(aes(x = df1$Ns, y = df1$Medias, colour = "Sem")) + geom_point(aes(x = df3$Ns, y = df3$Medias, colour = "Com")) +
  geom_line(aes(x = df1$Ns, y = df1$Medias, colour = "Sem"), color = "sky blue") + geom_line(aes(x = df3$Ns, y = df3$Medias, colour = "Com"), color = "orange") +
  ylab("Médias das Amplitudes dos Intervalos de Confiança") + xlab("Dimensão da Amostra") + labs(colour = "Contaminação:") +
  ggtitle("Médias das Amplitudes dos 550 intervalos de confiança com e sem contaminação para um nível de confiança de 91%")
```

Semente = 558

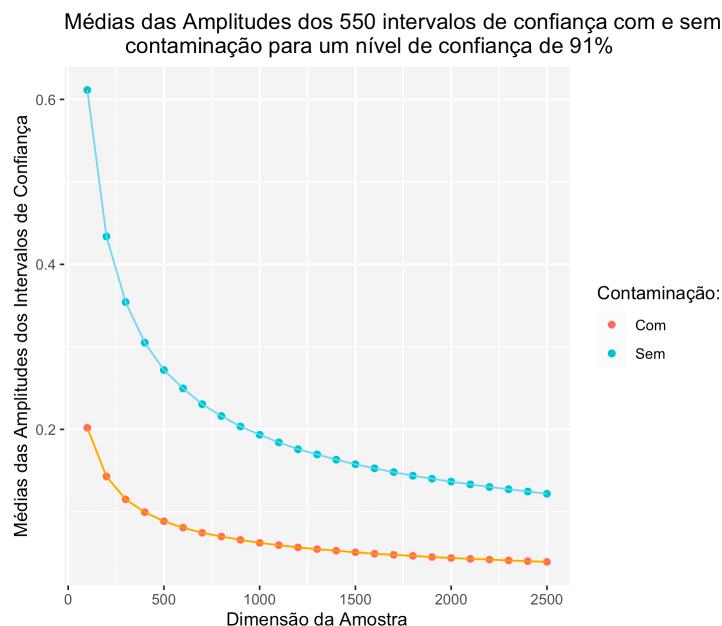
Número de amostras (m) = 550

$\lambda = 1.8$

$\lambda c = 0.12$

$\epsilon = 15\%$

Nível de confiança ($1-\alpha$) = 0.91



A partir da observação direta do gráfico, concluímos que as médias das amplitudes dos intervalos de confiança para o inverso do valor esperado diminuem com o aumento da dimensão da amostra, tal como acontece no exercício 9. Isto deve-se ao facto de a partir do TLC se concluir que esta amplitude é inversamente proporcional ao produto da média da amostra pela raiz quadrada do tamanho dela. Além disso, para a experiência com contaminação é possível observar que as amplitudes são menores em todos os casos, devido ao facto de, apesar da contaminação ser de apenas 15%, as discrepâncias entre o parâmetro lambda (inverso do valor esperado) da amostra original e da amostra contaminada serem muito grandes. Uma vez que a amostra contaminada apresenta um maior valor esperado (inverso de lambda), a média da amostra aumenta e, para uma dimensão fixa da amostra, a amplitude do intervalo de confiança diminui por ser inversamente proporcional à média da amostra. Além disso, salienta-se o facto de que para um nível de confiança de 91%, 91% dos intervalos estimados devem conter o verdadeiro valor do inverso do valor esperado.