

Questão 10

João Barreiros C. Rodrigues

Junho 2022

1 Parâmetros

seed=341, m=800, $\lambda=5.54$, $\lambda_c=0.51$, $\epsilon=0.15$, $(1-\alpha)=0.999$

2 Código

```
#install.packages("ggplot2") #Install packages, using Netherlands CRAN mirror
#install.packages("dplyr")
library("dplyr")
library("ggplot2") #Link library
set.seed(341)
m = 800
minn=100
maxnn=2500
lambda = 5.54
lambdac= 0.51
alpha=0.001
alphapercentile=alpha/2 #1-alpha=0.999 therefore we can be assure we are in the 99.5 th percentile
mean0=rep(0,m)
amplitude=rep(0,m)
amplitudemmean=rep(0, 25)
cmean0=rep(0,m)
camplitude=rep(0,m)
camplitudemmean=rep(0, 25)
ninterval=seq(from = minn, to= maxnn, by =100)
for(n in ninterval){
  for (i in 1:m){ #generate regular observations
    normal=rexp(n,lambda)
    fullycontaminated=rexp(n,lambdac)
    partialcontaminated=normal
    for(a in 1:n*0.15){
      partialcontaminated[a]=fullycontaminated[a]
    }

    mean0[i]=mean(normal)
    cmean0[i]=mean(partialcontaminated)
  }
  s=mean0
  cs=cmean0
  amplitude <- qt(1-alphapercentile,df=n-1)/(s*sqrt(n))
  camplitude <- qt(1-alphapercentile,df=n-1)/(cs*sqrt(n))
  amplitudemmean[n/100]=mean(amplitude)
  camplitudemmean[n/100]=mean(camplitude)
  print(n) #Added in order to be sure the script was indeed running
  print(amplitudemmean[n/100])
}

df.amplitudemmean <- as.data.frame(amplitudemmean)
print(df.amplitudemmean)
df.amplitudemmean <- cbind(df.amplitudemmean, Contaminação="Não Contaminado") #label dataframe
df.amplitudemmean <- cbind(df.amplitudemmean, n=ninterval)
df.camplitudemmean <- as.data.frame(camplitudemmean)
print(df.camplitudemmean)
df.camplitudemmean <- cbind(df.camplitudemmean, Contaminação="Contaminado")
df.camplitudemmean <- cbind(df.camplitudemmean, n=ninterval)
print(df.camplitudemmean)
colnames(df.camplitudemmean)=c("amplitudemmean", "Contaminação", "n")
print(df.camplitudemmean)
print(df.amplitudemmean)
df.compact <- union(df.amplitudemmean, df.camplitudemmean)
print(df.compact)
plot <- ggplot(df.compact, aes(x=n, y=amplitudemmean, colour=Contaminação))
plot <- plot + geom_point()+stat_smooth(method="auto", se=FALSE, fullrange=TRUE)
plot <- plot + ggtitle("Média das amplitudes dos intervalos de confiança das 800 observações em função de n") + ylab("Média das
Amplitudes dos intervalos de confiança obtidos") + xlab("Dimensão das amostras (n)") #add title and rename y and x axis
print(plot)
ggsave("module10.png", plot = plot, device=png, width=14) #save plot
```

3 Outputs Gráficos

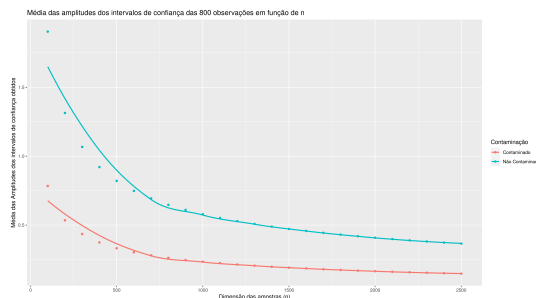


Figura 1: Amplitude do intervalo de confiança em função da dimensão da amostra (n)

4 Comentários

À semelhança do exercício anterior verifica-se uma relação aproximadamente inversamente proporcional entre a dimensão da amostra n e a amplitude do intervalo de confiança em ambas as amostras.

Deve salientar-se que embora a amostra contaminada obedea à relação anteriormente estipulada, em comparação para igual n , a amostra contaminada tem sempre menor média dos intervalos de confiança. Tal pode ser justificado teoricamente por $\lambda > \lambda_c$