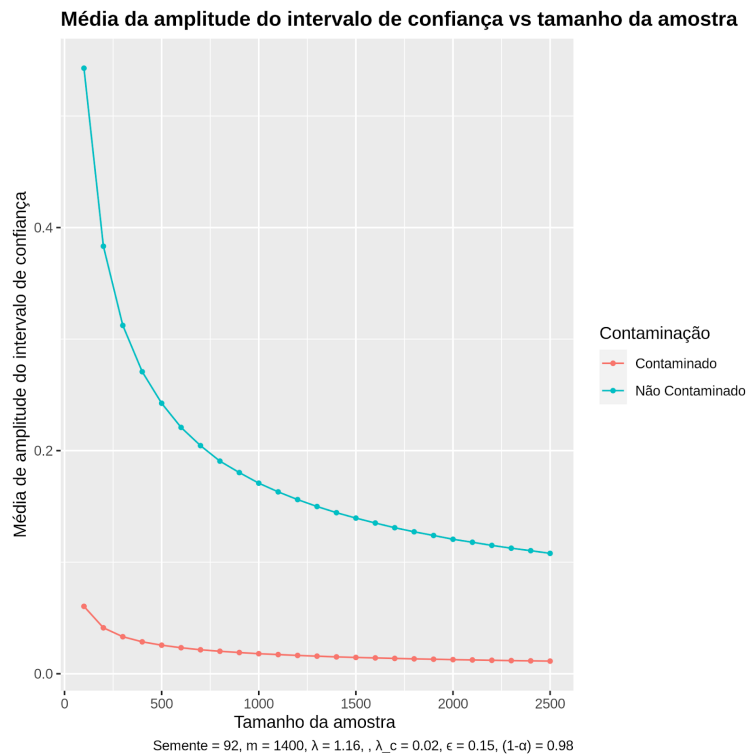


```

library(ggplot2)
library("tidyr")
set.seed(346)
df1 <- data.frame(value = c(), mean = c())
confidence_interval <- function(vector, confidence_level) {
  a <- qnorm(1 - ((1 - confidence_level) / 2))
  erro <- a / sqrt(length(vector))
  result <- c((1 - erro) / mean(vector), (1 + erro) / mean(vector))
  return(result)
}
for (num in seq(100, 2500, 100)) {
  all_values <- c()
  all_values_c <- c()
  for (i in 1:100) {
    observations <- rexp(num, 0.45)
    outliers <- rexp(num * 0.2, 0.01)
    observations_c <- c(outliers, observations[(num * 0.2):length(observations)])
    interval <- confidence_interval(observations, 0.999)
    interval_c <- confidence_interval(observations_c, 0.999)
    all_values <- c(all_values, abs(interval[1] - interval[2]))
    all_values_c <- c(all_values_c, abs(interval_c[1] - interval_c[2]))
  }
  mean <- mean(all_values)
  mean_c <- mean(all_values_c)
  df1 <- rbind(df1, data.frame(value = num, mean = mean, mean_c = mean_c))
}
names(df1)[2] <- "Não Contaminado"
names(df1)[3] <- "Contaminado"
df1_tidy <- df1 %>%
  pivot_longer(c("Não Contaminado", "Contaminado"), names_to = "Contaminação", values_to = "Intervalo")
df1_plot <- ggplot(df1_tidy, aes(x = value, y = Intervalo, color = Contaminação)) +
  labs(x = "Tamanho da amostra", y = "Média de amplitude do intervalo de confiança", title = "Média
da amplitude do intervalo de confiança vs tamanho da amostra",
caption = "Semente = 92, m = 1400, λ = 1.16, λ_c = 0.02, ε = 0.15, (1-α) = 0.98") +
  theme(plot.title = element_text(size = 13, face = "bold")) +
  geom_line() +
  geom_point(size = 1)
ggsave("amplitudevs sample.png", df1_plot)

```



Infere-se que as médias de amplitudes das amostras contaminadas são inferiores às médias das amostras não contaminadas, dada a diferença nos lambdas, 1.16 e 0.02 respetivamente. Ao contaminar uma amostra, baixa o valor da sua média, pelo que a amplitude do intervalo também diminui. Continua a notar-se que a média diminui à medida que aumenta o tamanho da amostra.