

# Planejamento e Análise de Experimentos (EEE933)

## Estudo de Caso 1

Pedro Vinícius, Samara Silva e Savio Vieira

10 de Agosto de 2020

```
## Registered S3 method overwritten by 'GGally':  
##   method from  
##   +.gg      ggplot2
```

## Introdução

Lindsey (1996) define inferência estatística como o campo que se preocupa com o processo probabilístico relacionado a eventos já ocorridos e às ilações sobre a população que pode-se chegar com base nesse conhecimento, ou seja, trata-se de extrair conclusões para o universo de estudo a partir de observações empíricas ou processos estocásticos.

Neste trabalho, esse conceito será utilizado mediante a geração de dados experimentais da nova versão de um determinado software, seguida da análise comparativa quanto a variante já conhecida. Nessa perspectiva, é investigado se houve redução do custo médio de execução e/ou da variância fornecida. Para isso, em ambos os casos as hipóteses nulas foram definidas de maneira conservadora, partindo-se do pressuposto de que os parâmetros populacionais conhecidos foram mantidos na nova versão. Posteriormente foram estabelecidas condições iniciais quanto à quantidade de observações, junto à coleta de dados, análise da distribuição amostral e aplicação dos testes.

Nas próximas seções é feito o detalhamento de cada uma dessas etapas.

## Descrição do Problema

### Parte 1: Teste Sobre o Custo Médio

#### Planejamento dos Experimentos

A primeira fase desse experimento consiste em gerar uma amostra representativa do desempenho da nova versão do software. Para isso, é necessário especificar o tamanho dessa amostra, considerando os critérios de análise pré-estabelecidos. A saber, nível de significância  $\alpha = 0.01$ , efeito relevante de  $\delta^* = 4$ , e poder de  $\pi = 1 - \beta = 0.8$ .

Neste âmbito, o Poder do Teste pode ser considerado. Esse teste é originalmente usado para avaliar a hipótese nula, e rejeitá-la quando for falsa (...), porém também pode ser utilizado para estimar outros parâmetros amostrais, como efeito relevante, significância, poder e tamanho da amostra. Este último, é tido como critério de fundamental importância no planejamento da pesquisa, para reduzir o custo operacional, que é aproximadamente linear (KOCHER, JAFFE, JUN, 1999).

Contudo, para executar esse teste é necessário fornecer um valor inicial para variância, que ainda não é conhecida. A bibliografia sugere três maneiras de fazê-lo, ...

Considerando as vantagens e desvantagens de cada uma, optou-se por utilizar...

$$\begin{cases} H_0 : \mu = 50 \\ H_1 : \mu < 50 \end{cases}$$

```
# Define the sample size to be used in this experiment
(params <- power.t.test(delta = delta_star,
  sd = sigma_n,
  sig.level = alpha,
  power = pi,
  type = "one.sample",
  alternative = "one.sided"))
```

```
##
##      One-sample t test power calculation
##
##              n = 65.45847
##              delta = 4
##              sd = 10
##              sig.level = 0.01
##              power = 0.8
##      alternative = one.sided
# Number of observations
n <- ceiling(params$n)
```

## Coleta dos Dados

```
data_generation <- function(n){

  mre <- list(name = "recombination_bin", cr = 0.9)
  mmu <- list(name = "mutation_rand", f = 2)
  mpo <- 100
  mse <- list(name = "selection_standard")
  mst <- list(names = "stop_maxeval", maxevals = 10000)
  mpr <- list(name = "sphere", xmin = -seq(1, 20), xmax = 20 + 5 * seq(5, 24))

  sample <- c()
  # Generate n observations
  for (i in 1:n){
    observation <- ExpDE(mpo, mmu, mre, mse, mst, mpr,
      showpars = list(show.iters = "none"))$Fbest
    sample <- c(sample, observation)
  }

  return(sample)
}
```

```
# Random seed
set.seed(1007)

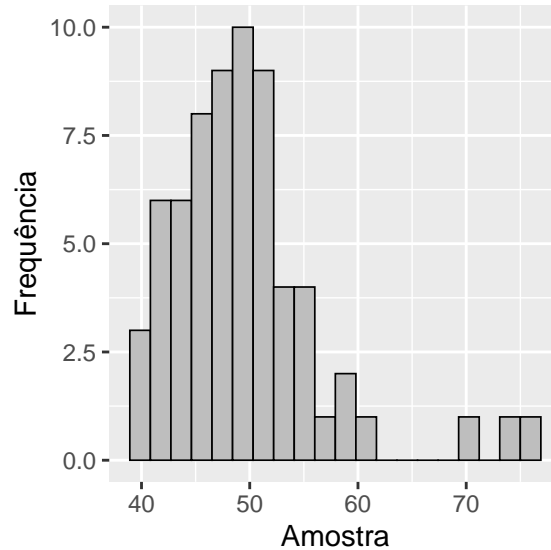
# Collect the sample with n observations
sample <- data_generation(n = n)

# Saves data to the csv file
```

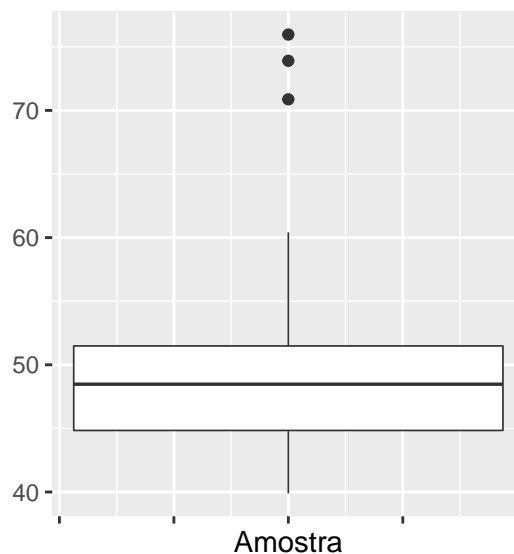
```
write.table(sample, file = 'sample.csv', row.names = FALSE, col.names = FALSE)
```

## Análise Exploratória de Dados

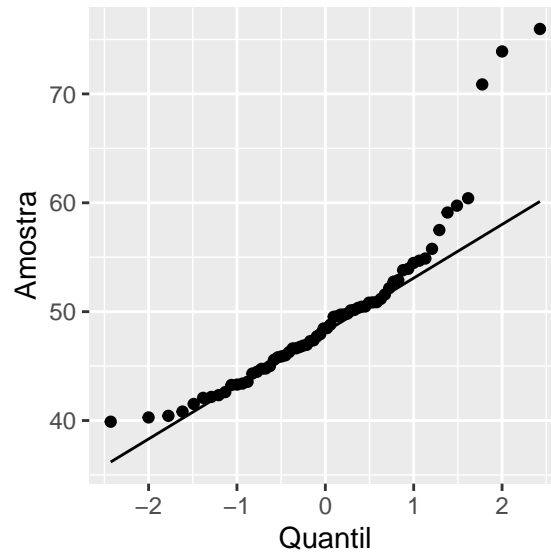
```
# Histogram
histogram <- ggplot(data = as.data.frame(sample), mapping = aes(x = sample))
histogram + geom_histogram(lwd = 0.3, bins = 20, color = 'black', fill = 'gray') +
  scale_x_continuous(name = 'Amostra') +
  scale_y_continuous(name = 'Frequência')
```



```
# Boxplot
boxplot <- ggplot(data = as.data.frame(sample), mapping = aes(y = sample))
boxplot + geom_boxplot(lwd = 0.3) +
  scale_x_continuous(name = 'Amostra') +
  scale_y_continuous(name = '') +
  theme(axis.text.x = element_blank())
```



```
# QQ-Plot
qqplot <- ggplot(data = as.data.frame(sample), mapping = aes(sample = sample))
qqplot + geom_qq_line() +
  geom_qq() +
  scale_y_continuous(name = 'Amostra') +
  scale_x_continuous(name = 'Quantil')
```



## Análise Estatística

```
# ----- Hypothesis Testing ----- #
(t_test <- t.test(x = sample,
  mu = mu_c,
  alternative = "less",
  conf.level = conf_level))

##
## One Sample t-test
##
## data: sample
## t = -0.5742, df = 65, p-value = 0.2839
## alternative hypothesis: true mean is less than 50
## 99 percent confidence interval:
## -Inf 51.5952
## sample estimates:
## mean of x
## 49.49419

# Confidence Interval
CI <- t_test$conf.int[1:2]
```

## Validação de Premissas

```
# Wilcoxon Signed-Ranks Test, because it does not assume normality
(wilcoxon_test <- wilcox.test(x = sample,
  alternative = "less",
```

```
mu = mu_c,  
paired = FALSE,  
exact = NULL,  
correct = TRUE,  
conf.int = FALSE,  
conf.level = conf_level))
```

```
##  
## Wilcoxon signed rank test with continuity correction  
##  
## data: sample  
## V = 820, p-value = 0.03433  
## alternative hypothesis: true location is less than 50
```

## Parte 2: Teste Sobre a Variância do Custo

### Planejamento dos Experimentos

$$\begin{cases} H_0 : \sigma^2 = 100 \\ H_1 : \sigma^2 < 100 \end{cases}$$

### Conclusões

## Referências

KOCHER, Paul; JAFFE, Joshua; JUN, Benjamin. Differential power analysis. In: Annual international cryptology conference. Springer, Berlin, Heidelberg, 1999. p. 388-397.

LINDSEY, James K. Parametric statistical inference. Oxford University Press, 1996.