

Ex 12

Natanael Magalhães Cardoso

24/05/2021

Importação das bibliotecas

```
library('tidyverse')
library('ggplot2')
library('readr')
```

```
df <- read.csv("characters.csv")
```

Variáveis aleatórias:

- Quantitativas
 - Discretas
 - Contínuas
- Qualitativas
 - Nominais
 - Ordinais

Tratamento de dados

```
df <- df %>% filter(!is.na(mass))
df <- df %>% filter(!is.na(height))
```

```
df$mass = as.numeric(df$mass)
```

```
## Warning: NAs introduzidos por coerção
```

```
df$height=as.numeric(df$height)
```

```
df <- df %>% filter(!is.na(mass))
```

Sumário

```
summary(df)
```

```
##      name          height      mass      hair_color
## Length:58      Min.   : 66.0    Min.   : 15.00    Length:58
## Class :character 1st Qu.:168.5    1st Qu.: 55.30    Class :character
## Mode  :character Median :180.0    Median : 79.00    Mode  :character
##              Mean   :174.0    Mean   : 75.58
##              3rd Qu.:192.5    3rd Qu.: 84.00
##              Max.   :234.0    Max.   :159.00
## skin_color      eye_color      birth_year      gender
## Length:58      Length:58      Length:58      Length:58
## Class :character Class :character Class :character Class :character
```

```
## Mode :character Mode :character Mode :character Mode :character
##
##
##
## homeworld species
## Length:58 Length:58
## Class :character Class :character
## Mode :character Mode :character
##
##
##
```

Implementação da função do Intervalo de Confiança da média com variância conhecida

```
ic <- function(dados, confianca, dp) {
  xbarra = mean(dados, na.rm=TRUE)
  z = qnorm((1 + confianca) / 2)
  n = length(dados)
  e = z * (dp / sqrt(n))
  LS = xbarra + e
  LI = xbarra - e
  return(c(LS, LI))
}
```

```
ic(df$height, 0.95, 30)
```

```
## [1] 181.7207 166.2793
```

Eu tenho 95% de chance da média estar dentro do range. O intervalo é uma variável aleatória, quando mudar amostra vou mudar \bar{x} .

Aumentar o nível de confiança diminui a amplitude do IC.

```
ic(df$height, 0.99, 30)
```

```
## [1] 184.1467 163.8533
```

Aumentar o desvio padrão aumenta a amplitude do desvio padrão.

```
ic(df$height, 0.99, 40)
```

```
## [1] 187.5289 160.4711
```