

# Avaliação Continuada 3

Matheus de Andrade Lourenço

2023-11-02

## Atividade Continuada 3

### Bases de Dados:

*# Base de dados aleatória:*

```
set.seed(23)
```

```
base1 <- sample(x <- 1L:100L, size <- 15, replace = TRUE)
```

```
cat("BASE DE DADOS SIMULADA:", base1) # Impressão com formatação.
```

```
## BASE DE DADOS SIMULADA: 29 28 72 43 45 34 48 17 21 17 40 36 70 70 31
```

*# Base de dados real:*

```
base2 <- read.csv("diabetes.csv", sep=";", dec=".")
```

```
base2 <- base2$DiabetesPedigreeFunction
```

```
cat("\n\nBASE DE DADOS REAL:", head(base2, n = 10), "...")
```

```
##
```

```
##
```

```
## BASE DE DADOS REAL: 0.627 0.351 0.672 0.167 2.288 0.201 0.248 0.134  
0.158 0.232 ...
```

### Exercício 1:

#### Média:

### Funções Programadas

```
media = function(base) {  
  soma <- 0  
  for(i in base) {  
    soma <- soma + i  
  }  
  return(soma / length(base))  
}
```

```
cat("A média da base de dados simulada é:", sprintf("%.2f",  
media(base1)))
```

```
## A média da base de dados simulada é: 40.07
```

## Funções do R

```
cat("A média da base de dados simulada é:", sprintf("%.2f",  
mean(base1)))
```

```
## A média da base de dados simulada é: 40.07
```

## Mediana:

### Funções Programadas

```
mediana <- function(base) {  
  aux <- sort(base)  
  n <- length(base)  
  if (n %% 2 == 0) {  
    return((aux[n / 2] + aux[n / 2 + 1]) / 2)  
  } else {  
    return(aux[ n %% 2 + 1])  
  }  
}
```

```
cat("\nA mediana da base de dados simulada é:", mediana(base1))
```

```
##
```

```
## A mediana da base de dados simulada é: 36
```

## Funções do R

```
cat("A mediana da base de dados simulada é:", median(base1))
```

```
## A mediana da base de dados simulada é: 36
```

## Variância

### Amostrai

### Funções Programadas

```
varAmostrai = function(base) {  
  soma <- 0  
  n <- length(base)  
  mediaB <- media(base)  
  for (i in 1:n) {  
    soma <- soma + (base[i] - mediaB)^2  
  }  
  
  return(soma / (n - 1))  
}
```

```
cat("A Variância Amostral da base de dados simulada é: ", sprintf("%.2f",  
varAmostral(base1)))
```

```
## A Variância Amostral da base de dados simulada é: 338.50
```

## Funções do R

```
cat("A Variância Amostral da base de dados simulada é:", sprintf("%.2f",  
var(base1)))
```

```
## A Variância Amostral da base de dados simulada é: 338.50
```

## Populacional

### Funções Programadas

```
varPopul <- function(base) {  
  soma <- 0  
  n <- length(base)  
  mediaB <- media(base)  
  for (i in 1:n) {  
    soma <- soma + (base[i] - mediaB)^2  
  }  
  
  return(soma / n)  
}
```

```
cat("A Variância Populacional da base de dados simulada é: ",  
sprintf("%.2f", varPopul(base1)))
```

```
## A Variância Populacional da base de dados simulada é: 315.93
```

## Funções do R

```
cat("A Variância Populacional da base de dados simulada é:",  
sprintf("%.2f", var(base1) * (length(base1) - 1) / length(base1)))
```

```
## A Variância Populacional da base de dados simulada é: 315.93
```

*# A linguagem R não possui a função voltada para a variância populacional.*

## Desvio Padrão

### Amostral

### Funções Programadas

```
desvAmostral <- function(base) {  
  return(sqrt(varAmostral(base1)))  
}
```

```
cat("O desvio padrão amostral da base de dados simulada é: ",
sprintf("%.2f", desvAmostral(base1)))

## O desvio padrão amostral da base de dados simulada é: 18.40
```

## Funções do R

```
cat("O desvio padrão amostral da base de dados simulada é: ",
sprintf("%.2f", sd(base1)))

## O desvio padrão amostral da base de dados simulada é: 18.40
```

## Populacional

### Funções Programadas

```
desvPopul <- function(base) {
  return(sqrt(varPopul(base1)))
}

cat("O desvio padrão populacional da base de dados simulada é: ",
sprintf("%.2f", desvPopul(base1)))

## O desvio padrão populacional da base de dados simulada é: 17.77
```

## Funções do R

```
cat("O desvio padrão populacional da base de dados simulada é: ",
sprintf("%.2f", sd(base1) * sqrt(length(base1) - 1) / sqrt(length(base1))
))

## O desvio padrão populacional da base de dados simulada é: 17.77

# O R não possui nenhuma função pronta para desvio padrão populacional
```

## Quartis

### Funções Programadas

```
quartis <- function(base) {
  aux <- base

  if(length(base) %% 2 == 0) {
    aux <- c(aux, median(base))
  }

  aux <- sort(aux)
  n <- length(aux)

  q2 <- median(aux)
  q1 <- median(aux[1:which(aux == q2)])
  q3 <- median(aux[which(aux == q2):n])
}
```

```
qts <- c(q1, q2, q3)

return(qts)
}

quartis1 <- quartis(base1)

cat("Primeiro quartil:", quartis1[1], "\n")
## Primeiro quartil: 28.5

cat("Segundo quartil (mediana): ", quartis1[2], "\n")
## Segundo quartil (mediana): 36

cat("Terceiro quartil: ", quartis1[3], "\n")
## Terceiro quartil: 46.5
```

## Funções do R

```
cat("Primeiro quartil: ", quantile(base1, probs = 0.25), "\n")
## Primeiro quartil: 28.5

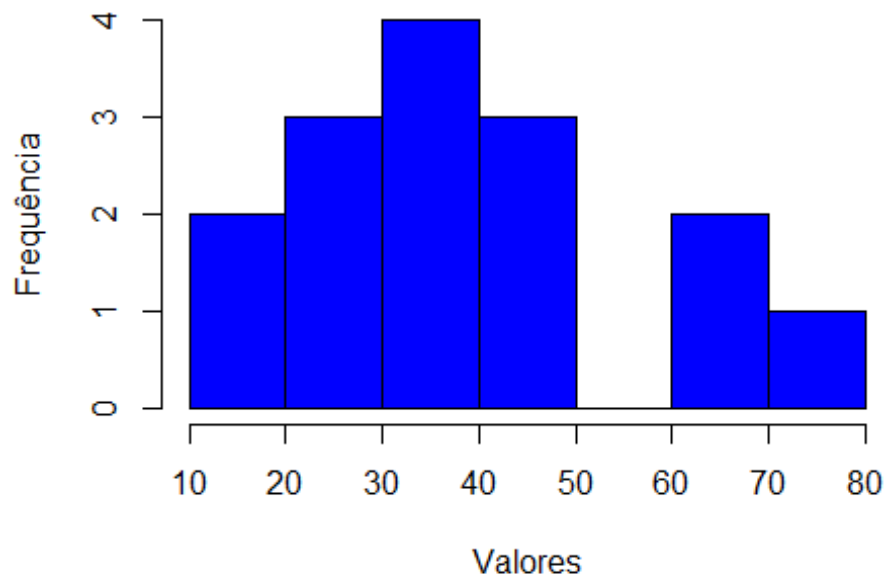
cat("Segundo quartil: ", quantile(base1, probs = 0.50), "\n")
## Segundo quartil: 36

cat("Terceiro quartil: ", quantile(base1, probs = 0.75), "\n")
## Terceiro quartil: 46.5
```

## Histograma

```
hist(base1,col = "blue", border = "black", main = "Histograma da base de
dados simulada", xlab = "Valores", ylab = "Frequência")
```

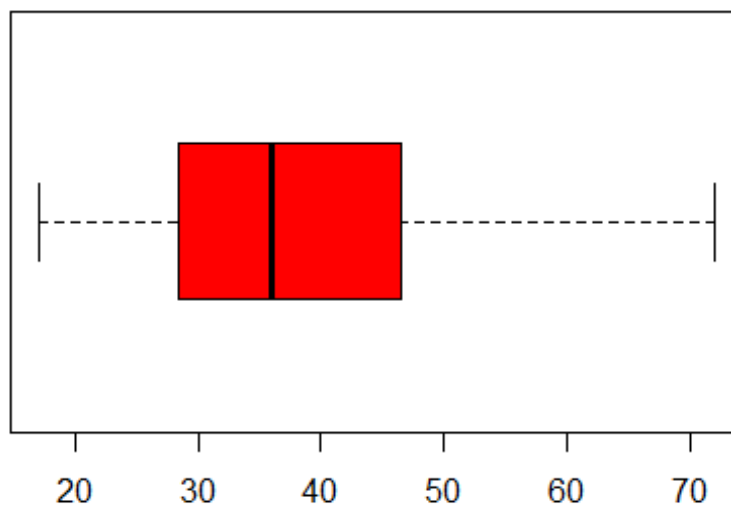
**Histograma da base de dados simulada**



**Boxplot**

```
boxplot(base1, main = "Boxplot da base simulada", horizontal = TRUE, col = "red")
```

**Boxplot da base simulada**



## Exercício 2:

### Média:

#### Funções Programadas

```
cat("A média da base de dados real é: ", sprintf("%.2f", media(base2)))  
## A média da base de dados real é:  0.47
```

#### Funções do R

```
cat("A média da base de dados simulada é: ", sprintf("%.2f",  
mean(base2)))  
## A média da base de dados simulada é:  0.47
```

### Mediana:

#### Funções Programadas

```
cat("\nA mediana da base de dados real é: ", mediana(base2))  
##  
## A mediana da base de dados real é:  0.3725
```

#### Funções do R

```
cat("A mediana da base de dados real é: ", median(base2))  
## A mediana da base de dados real é:  0.3725
```

### Variância

#### Amostral

#### Funções Programadas

```
cat("A Variância Amostral da base de dados real é: ", sprintf("%.2f",  
varAmostral(base2)))  
## A Variância Amostral da base de dados real é:  0.11
```

#### Funções do R

```
cat("A Variância Amostral da base de dados real é: ", sprintf("%.2f",  
var(base2)))  
## A Variância Amostral da base de dados real é:  0.11
```

#### Populacional

#### Funções Programadas

```
cat("A Variância Populacional da base de dados real é: ", sprintf("%.2f",  
varPopul(base2)))
```

```
## A Variância Populacional da base de dados real é:  0.11
```

## Funções do R

```
cat("A Variância Populacional da base de dados real é: ", sprintf("%.2f",  
var(base2) * (length(base2) - 1) / length(base2)))
```

```
## A Variância Populacional da base de dados real é:  0.11
```

*# A linguagem R não possui a função voltada para a variância populacional.*

## Desvio Padrão

### Amostrai

## Funções Programadas

```
cat("O desvio padrão amostral da base de dados real é: ", sprintf("%.2f",  
desvAmostrai(base2)))
```

```
## O desvio padrão amostral da base de dados real é:  18.40
```

## Funções do R

```
cat("O desvio padrão amostral da base de dados real é: ", sprintf("%.2f",  
sd(base2)))
```

```
## O desvio padrão amostral da base de dados real é:  0.33
```

### Populacional

## Funções Programadas

```
cat("O desvio padrão populacional da base de dados real é: ",  
sprintf("%.2f", desvPopul(base2)))
```

```
## O desvio padrão populacional da base de dados real é:  17.77
```

## Funções do R

```
cat("O desvio padrão populacional da base de dados real é: ",  
sprintf("%.2f", sd(base2) * sqrt(length(base2) - 1) / sqrt(length(base2))  
))
```

```
## O desvio padrão populacional da base de dados real é:  0.33
```

*# O R não possui nenhuma função pronta para desvio padrão populacional*

## Quartis

## Funções Programadas



```
quartis2 <- quantis(base2)

cat("Primeiro quartil:", quartis2[1], "\n")
## Primeiro quartil: 0.244

cat("Segundo quartil (mediana): ", quartis2[2], "\n")
## Segundo quartil (mediana): 0.3725

cat("Terceiro quartil: ", quartis2[3], "\n")
## Terceiro quartil: 0.626
```

### Funções do R

```
cat("Primeiro quartil: ", quantile(base2, probs = 0.25), "\n")
## Primeiro quartil: 0.24375

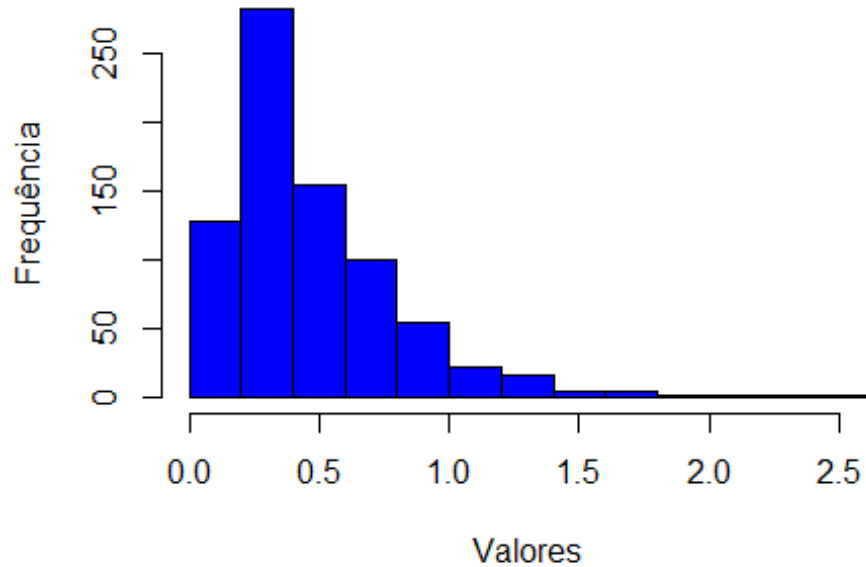
cat("Segundo quartil: ", quantile(base2, probs = 0.50), "\n")
## Segundo quartil: 0.3725

cat("Terceiro quartil: ", quantile(base2, probs = 0.75), "\n")
## Terceiro quartil: 0.62625
```

### Histograma

```
hist(base2,col = "blue", border = "black", main = "Histograma da base de  
dados real", xlab = "Valores", ylab = "Frequência")
```

**Histograma da base de dados real**



**Boxplot**

```
boxplot(base1, main = "Boxplot da base real", horizontal = TRUE, col = "red")
```

**Boxplot da base real**

