Trabalho\_BCC3EB2019.R

João

2019-04-16

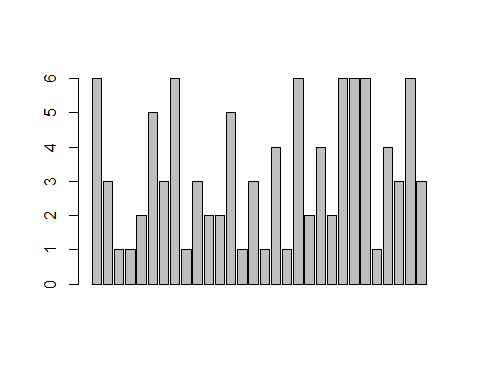
# Curso: Bacharelado em Ciência da Computação  
# Disciplina: Estatística Básica  
# Aluno: João Victor Lemes de Oliveira  
  
# Relação de comandos auxiliares ------------------------------------------  
  
# A sintaxe dos comandos poderá ser obtita utilizando   
# 'help <nome\_do\_comando>' ou '? <nome\_do\_comando>'.  
  
# Gráficos: plot, barplot, boxplot, stem.  
# Estatísticos: mean, median, var, sd, diff, range, nclass.Sturges, quantile,  
# summary, help,   
#  
  
  
# Para cada questão: execute o comando, mostre o resultado e, se necessário,   
# explique os valores obtidos.  
  
  
  
# Variável aleatória discreta ---------------------------------------------  
  
#1) Um dado foi lançado 30 vezes e os resultados obtidos foram anotados e   
# armazenados na variável 'dado'.  
  
set.seed(7)  
faces <- sample(1:6, 30, replace = T)  
  
#a) Construa a tabela de frequência.  
# (utilize o comando table)  
nclass.Sturges(faces)

## [1] 6

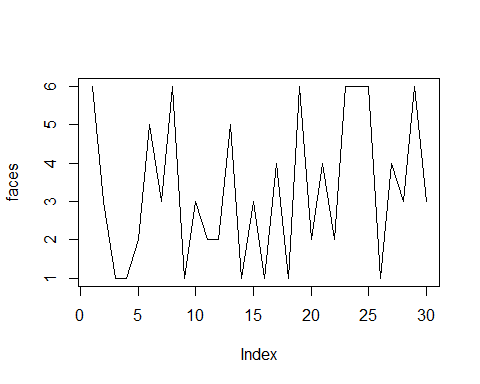
table(cut(faces, breaks = 6))

##   
## (0.995,1.83] (1.83,2.67] (2.67,3.5] (3.5,4.33] (4.33,5.17]   
## 7 5 6 3 2   
## (5.17,6]   
## 7

#b) Crie um gráfico de barras para representar os resultados obtidos nos  
# lançamentos do dado.  
# (utilize o comando barplot)  
barplot(faces)



#c) Crie um gráfico de linha para representar a frequência relativa acumulada.  
plot(faces, type = "l")



# Variável aleatória contínua ---------------------------------------------  
  
#2) Para adquirir um medicamento, realizou-se uma pesquisa de preços e   
# os dados foram armazenados na variável 'valor'.  
  
set.seed(7)  
valor <- round(runif(35, min=30, max=100), 2)  
  
# Com base nas informações obtidas, e utilizandos os comandos do R, calcule:  
  
  
#a) A amplitude da amostra (L);  
L <- max(valor) - min(valor)  
L

## [1] 69.16

#b) A quantidade de classes segundo o método de Sturges (k);  
K <- nclass.Sturges(valor)  
K

## [1] 7

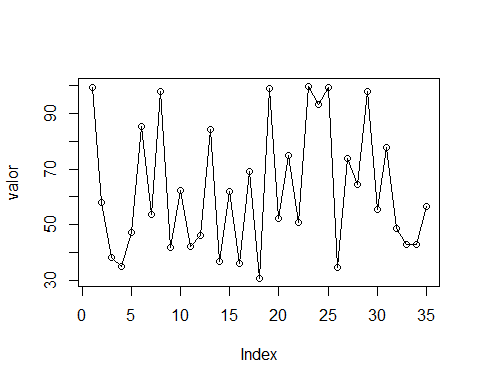
#c) A amplitude de cada classe (h);  
H <- L/K  
H

## [1] 9.88

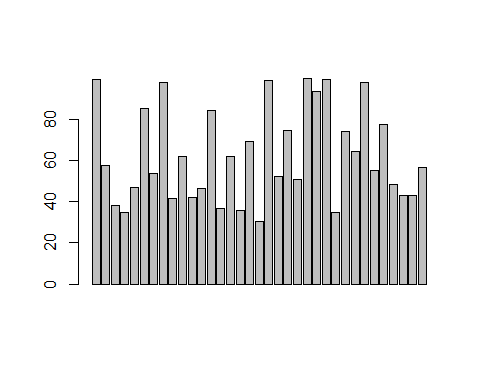
#d) A distribuição de frequências (absoluta e relativa);  
table(cut(valor, breaks = K))

##   
## (30.5,40.5] (40.5,50.4] (50.4,60.2] (60.2,70.1] (70.1,80] (80,89.9]   
## 6 7 6 4 3 2   
## (89.9,99.8]   
## 7

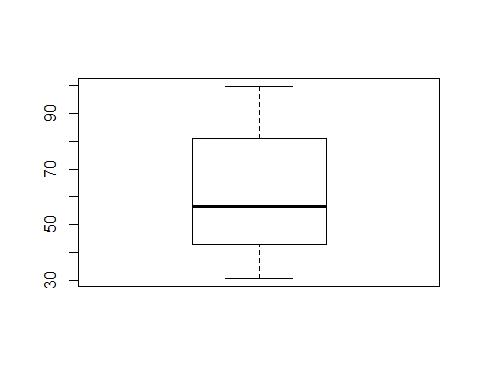
#e) Gere um diagrama de dispersão para visualizar a forma como os dados   
# estão distribuídos;  
plot(valor,type = "o")



#f) Gere um gráfico de barras para representar a distribuição de frequências;  
barplot(valor)



#g) Gere um gráfico boxplot (gráfico de bigode) e explique a variabilidade dos  
# dados representados plea variável 'valor';  
boxplot(valor)



#h) Determine os quartis e a amplitude interquartis (AIQ); Exlique a relação  
# com o boxplot criado no item g.  
quantile(valor)

## 0% 25% 50% 75% 100%   
## 30.61 42.98 56.55 80.85 99.77

Q1 <- quantile(valor, probs = 0.25)  
Q3 <- quantile(valor, probs = 0.75)  
AIQ <- Q3 - Q1  
AIQ

## 75%   
## 37.87

# EXPLICACAO:  
# Os dados dos quartis nos mostram os dados em determinado ponto, 25%,50%,75%,100% e 0%, se relacionando com  
# o bloxplot podemos dizer que um acompaha o outro, nos informando melhor onde está cada determinada porcentagem  
  
#i Cosntrua um diagrama de ramo-e-folhas.  
stem(valor)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 3 | 155678  
## 4 | 2233678  
## 5 | 124578  
## 6 | 2249  
## 7 | 458  
## 8 | 45  
## 9 | 388999  
## 10 | 0

#j) Calcule as estatísticas: média, mediana, variância, desvio padrão e   
# coeficiente de variação;  
# DADOS ORDENADOS:  
sort(valor)

## [1] 30.61 34.60 34.88 35.93 36.74 38.10 41.61 42.02 42.96 43.00 46.20  
## [12] 47.06 48.46 50.67 52.16 53.80 55.36 56.55 57.84 61.74 62.14 64.33  
## [23] 69.25 73.89 74.76 77.60 84.10 85.44 93.42 97.97 98.04 99.00 99.21  
## [34] 99.22 99.77

# MÉDIA:  
mean(valor)

## [1] 62.52657

# MEDIANA:  
median(valor)

## [1] 56.55

# VARIÂNCIA:  
var(valor)

## [1] 521.1142

# DESVIO PADRÃO:  
sd(valor)

## [1] 22.82793

# COEFICIENTE DE VARIAÇÃO  
sd(valor) / mean(valor)

## [1] 0.3650916

# Questão com análise livre -----------------------------------------------  
  
#3) Sabendo que a altura das pessoas de uma determinada região varia   
# entre 1,50m e 1,95m. Faça sua análise, de acordo com o conteúdo estudado,  
# considerando os dados seguintes:  
  
set.seed(7)  
altura <- round(runif(100, min = 1.50, max = 1.95), 2)  
  
# DADOS ORDENADOS:  
sort(altura)

## [1] 1.50 1.50 1.53 1.53 1.54 1.54 1.54 1.54 1.54 1.55 1.56 1.57 1.57 1.58  
## [15] 1.58 1.58 1.58 1.58 1.58 1.59 1.60 1.61 1.61 1.62 1.62 1.63 1.63 1.64  
## [29] 1.64 1.65 1.65 1.65 1.66 1.66 1.66 1.67 1.67 1.67 1.67 1.68 1.68 1.69  
## [43] 1.69 1.70 1.70 1.70 1.70 1.71 1.71 1.71 1.72 1.72 1.72 1.73 1.74 1.74  
## [57] 1.75 1.76 1.76 1.76 1.77 1.78 1.78 1.78 1.78 1.79 1.80 1.81 1.81 1.82  
## [71] 1.83 1.83 1.83 1.84 1.84 1.84 1.85 1.85 1.86 1.86 1.86 1.87 1.87 1.88  
## [85] 1.88 1.88 1.90 1.90 1.90 1.91 1.91 1.91 1.92 1.93 1.94 1.94 1.94 1.94  
## [99] 1.95 1.95

# MÉDIA:  
mean(altura)

## [1] 1.7289

# MEDIANA:  
median(altura)

## [1] 1.715

# VARIÂNCIA:  
var(altura)

## [1] 0.01647454

# DESVIO PADRÃO:  
sd(altura)

## [1] 0.1283532

# AMPLITUDE (L)  
L2 <- max(altura) - min(altura)  
L2

## [1] 0.45

# QUANTIDADE DE CLASSES:  
K2 <- nclass.Sturges(altura)  
K2

## [1] 8

# AMPLITUDE DA CLASSE:  
H2 <- L2/K2  
H2

## [1] 0.05625

# GRÁFICO DE RAMO-E-FOLHAS  
stem(altura)

##   
## The decimal point is 1 digit(s) to the left of the |  
##   
## 15 | 003344444  
## 15 | 56778888889  
## 16 | 011223344  
## 16 | 55566677778899  
## 17 | 0000111222344  
## 17 | 5666788889  
## 18 | 0112333444  
## 18 | 5566677888  
## 19 | 000111234444  
## 19 | 55

# DISTRIBUIÇÃO DE FREQUÊNCIA  
table(cut(altura, breaks = K2))

##   
## (1.5,1.56] (1.56,1.61] (1.61,1.67] (1.67,1.73] (1.73,1.78] (1.78,1.84]   
## 10 13 12 18 12 8   
## (1.84,1.89] (1.89,1.95]   
## 13 14

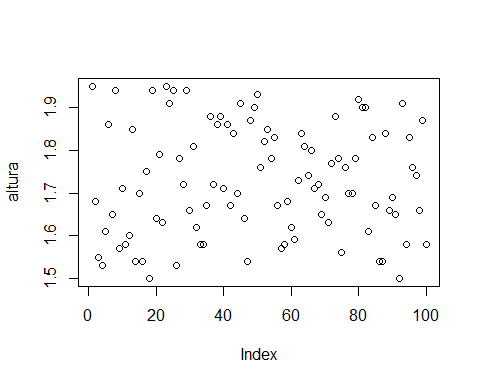
# QUARTIS E INTERQUARTIS  
quantile(altura)

## 0% 25% 50% 75% 100%   
## 1.5000 1.6275 1.7150 1.8400 1.9500

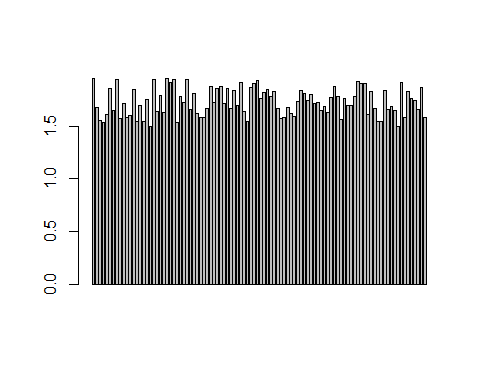
Q1\_questao <- quantile(altura, probs = 0.25)  
Q3\_questao <- quantile(altura, probs = 0.75)  
AIQ\_questao <- Q3\_questao - Q1\_questao  
AIQ\_questao

## 75%   
## 0.2125

# GRÁFICO DE DISPERSÃO  
plot(altura)



# GRÁFICO DE BARRAS  
barplot(altura)



# RESUMO DA QUESTAO  
summary(altura)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.500 1.627 1.715 1.729 1.840 1.950

# Questão Extra (Opcional) ------------------------------------------------  
  
#4) Uma pesquisa identificou que, quando conectados à Internet, os usuários  
# enviam informações somente em 20% de seu tempo de conexão.   
# Sabendo dessa informação, em uma empresa com 35 funcionários, qual a   
# probabilidade de 10 usuários, ou mais, estarem conectados e enviando   
# informações simultaneamente.  
#  
# Dica: probabilidade = C(N, n) \* p^n \* q^(N-n) \* 100  
#  
# C: combinação;   
# N: total de funcionários da empresa;  
# n: número de funcionários enviando informações na Internet;  
# p: probabilidade do funcionário está enviando informações;  
# q: probabilidade do funcionário não está enviando informações.  
N <- 35  
n <- 10  
p <- 20/100  
q <- 80/100  
  
choose(N,n) \* (p^n) \* q^(N-n) \* 100

## [1] 7.101884

#  
# Gere um documento PDF para postar no GoogleClass ------------------------  
  
# Após executar todos os comandos sem erro, compile o script do trabalho para  
# pressionando as teclas 'Ctrl + Shift + k'