Prática 4

Grupo MV:

Vitor Martins Basso - 11611BCC034

Marcos Gabriel Leão Muñoz - 11611BCC026

Exercise 4.1.7

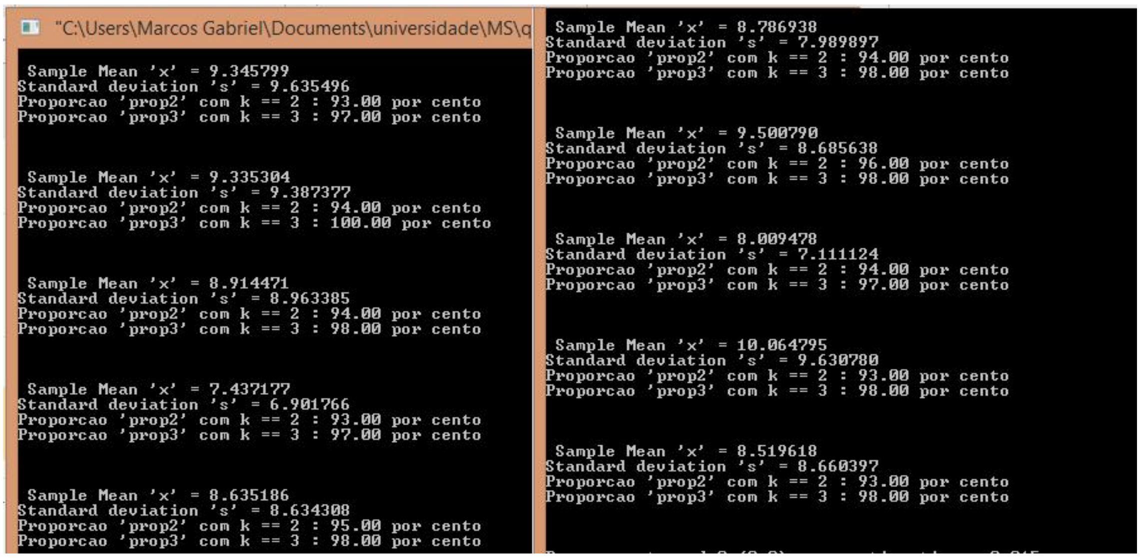
(a) Generate an Exponential (9) random variate sample of size n = 100 and compute the proportion of points in the sample that fall within the intervals xʋ ± 2s and xʋ ± 3s. Do this for 10 different rngs streams.

(b) In each case, compare the results with Chebyshev's inequality.

(c) Comment.

Resposta:

A) Código em
anexo



B) A desigualdade de Chebyshev dá percentuais de 75%, quando k = 2, e aproximadamente 89%, quando k = 3. Os resultados de todas as simulações variam no máximo até 3%, ou seja, as proporções de dados dentro da faixa 2k se mantém por volta de 94% quando k = 2 e 98% quando k = 3.

C) Como explicado no material, a desigualdade de Chebyshev não têm o propósito de acertar precisamente a proporção dos dados dentro de 2k, mas sim de dar uma idéia de onde a grande maioria dos dados irá se concentrar independentemente dos fatores em que se baseiam. Portanto, o fato de existir uma diferença notável entre os 75% estimado pela desigualdade e os 94% encontrados no experimento não é de grande relevância nem mesmo inesperado.

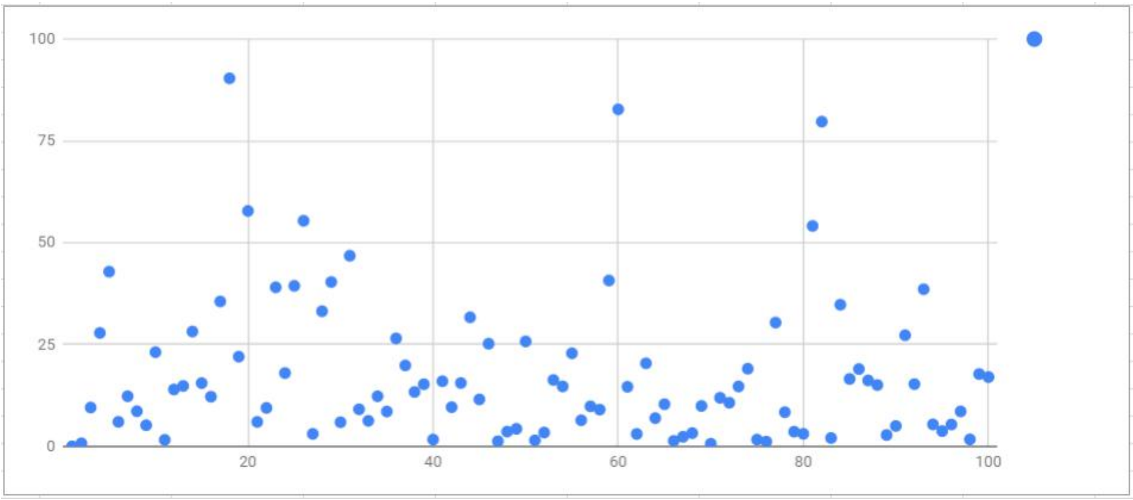
Exercise 4.1.8

Generate a plot similar to that in Figure 4.1.2 with calls to Exponential (17), rather than Random to generate the variates. Indicate the values to which the sample mean and sample standard deviation will converge.

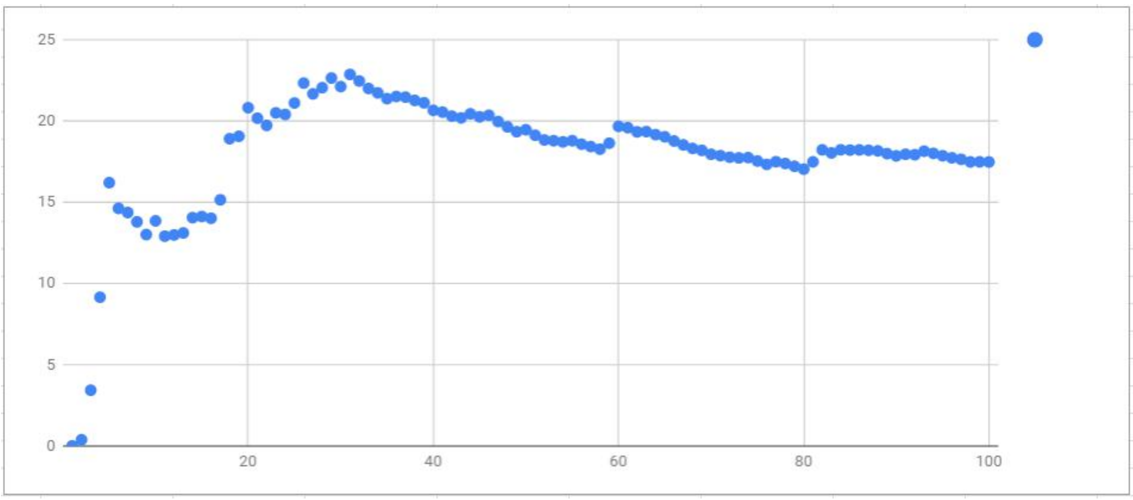
Resposta:

Código em anexo

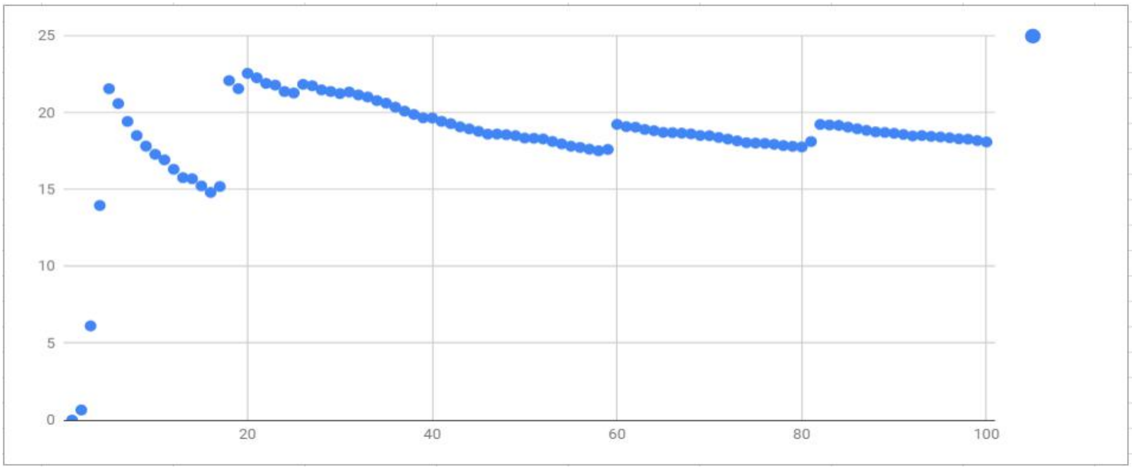
Data – cada ponto é um número gerado por Exponential(17).



Sample mean – Cada ponto é a média para cada número de amostras. Tende a 17, pois é o parâmetro da função Exponential(x) usada.



Standard deviation – Cada ponto é o desvio padrão para cada número de amostras. Tende a 17 pelo mesmo motivo da média.



Exercise 4.1.11 Calculate xȳ and s by hand using the 2-pass algorithm, the 1-pass algorithm, and Welford's algorithm in the following two cases.

(a) **The data based on n = 3 observations: x 1 = 1, x 2 = 6, and x 3 = 2.**

(b) **(b) The sample path x(t) = 3 for 0 < t ≤ 2, and x(t) = 8 for 2 < t ≤ 5, over the time interval 0 < t ≤ 5.**

Resposta:

A) Pelo 2-pass algorithm :

xȳ = (1 + 6 + 2)/3 = 3

s^2 = [(1-3)^2 + (6-3)^2 + (2-3)^2]/3 = 14/3

s = sqrt(s^2) = sqrt(14/3)

Pelo 1-pass algorithm:

xȳ = [(1 + 6 + 2)/3] = 3

s^2 = {[(1^2) + (6^2) + (2^2)]/3 } – xȳ^2 = 14/3

s = sqrt(s^2) = sqrt(14/3)

Pelo método de Welford:

n = 0 | xȳ = 0 | v = 0

x = 1 | n = 0 + 1 = 1 | d = 1 – 0 = 1|v = 0 + 1 * 1 *(1-1) / 1 = 0 | xȳ = 0 + 1 / 1 = 1

x = 6 | n = 1 + 1 = 2 | d = 6 – 1 = 5|v = 0 + 5 * 5 * (2–1) / 2 = 25/2 | xȳ = 1 + 5/2 = 7/2

x = 2 | n = 2 + 1 = 3 | d = 2 – 7/2 = -3/2| v = 25/2 + (-1) * (-1) * (3-1)/ 3 = 79/6
xȳ = 7/2 + (-3/2 * 1/3) = 3

s = sqrt(v/n) = sqrt(79/18)

B)

Pelo 2-pass algorithm:

xʋ = [(2–0) * 3 + (5-2) * 8]/ 5 = 6

s^2 = {(2-0) * [(3-6)^2] + (5-2)*[(8-6)^2]}/ 5 = 6

s = sqrt(6)

Pelo 1-pass algorithm:

xʋ = [(2–0) * 3 + (5-2) * 8]/ 5 = 6

s^2 = [(2-0) * (3^2) + (5-2)*(8^2)]/ 5 – (6^2) = 6

s = sqrt(6)

Pelo método de Welford

xʋ0 = 0 | xʋ1 = 0 + (2/2)*(3-0) = 3 | xʋ2 = x3 + (3/5)*(8 – 3) = 6 ---> xʋ = 6

v0 = 0 | v1 = 0 + [(2*0)/2] * (3 – 0)^2 = 0| v2 = 0 + [(3*2)/ 5) * (8 – 3)^2 = 30

s = sqrt(v2/ 5) = sqrt(6)

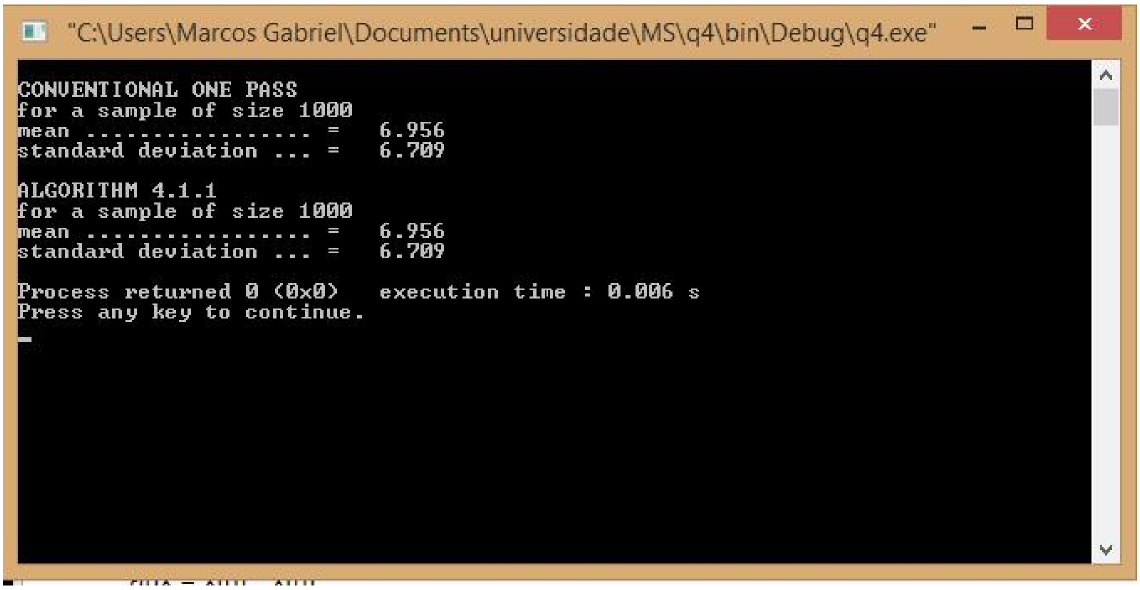
Exercise 4.1.13

Generate an Exponential(7) random variate sample of size n = 1000 and compute the mean and standard deviation using the Conventional One pass algorithm and the Algorithm 4.1.1. Comment on the results.

Resposta:

A) Código em anexo

Os dois algoritmos têm resultado iguais, mas geralmente o algoritmo de Welford é mais confiável. Isso se dá pelo fato do algoritmo de passada única convencional ser mais suscetível a overflows pela sua manipulação de floats envolver às vezes valores muito pequenos, não podendo às vezes utilizar o número com um alto grau de precisão em suas operações.



Exercise 4.2.2

(a) Generate the 2000-ball histogram in Example 4.2.2.

(b) Verify that the resulting relative frequencies f(x) satisfy the equation

f(x) ∼ (2^x * exp(−2))/x!   x = 0, 1, 2, . . .

(c) Then generate the corresponding histogram if 10 000 balls are placed, at random, in 1000 boxes.

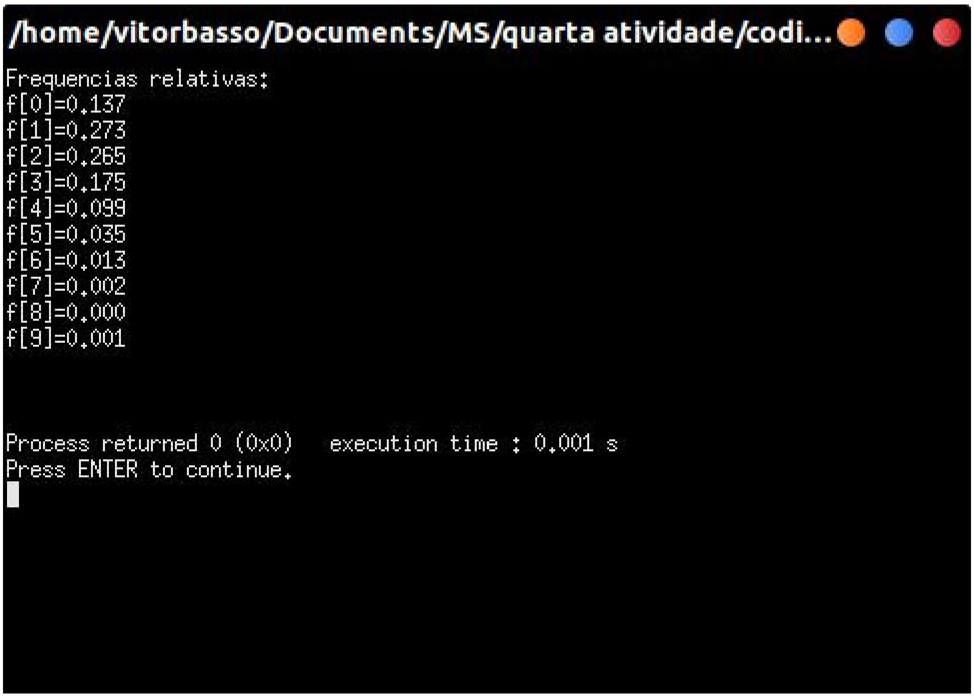
(d) Find an equation that seems to fit the resulting relative frequencies well and illustrate the quality of the fit.

Resposta:

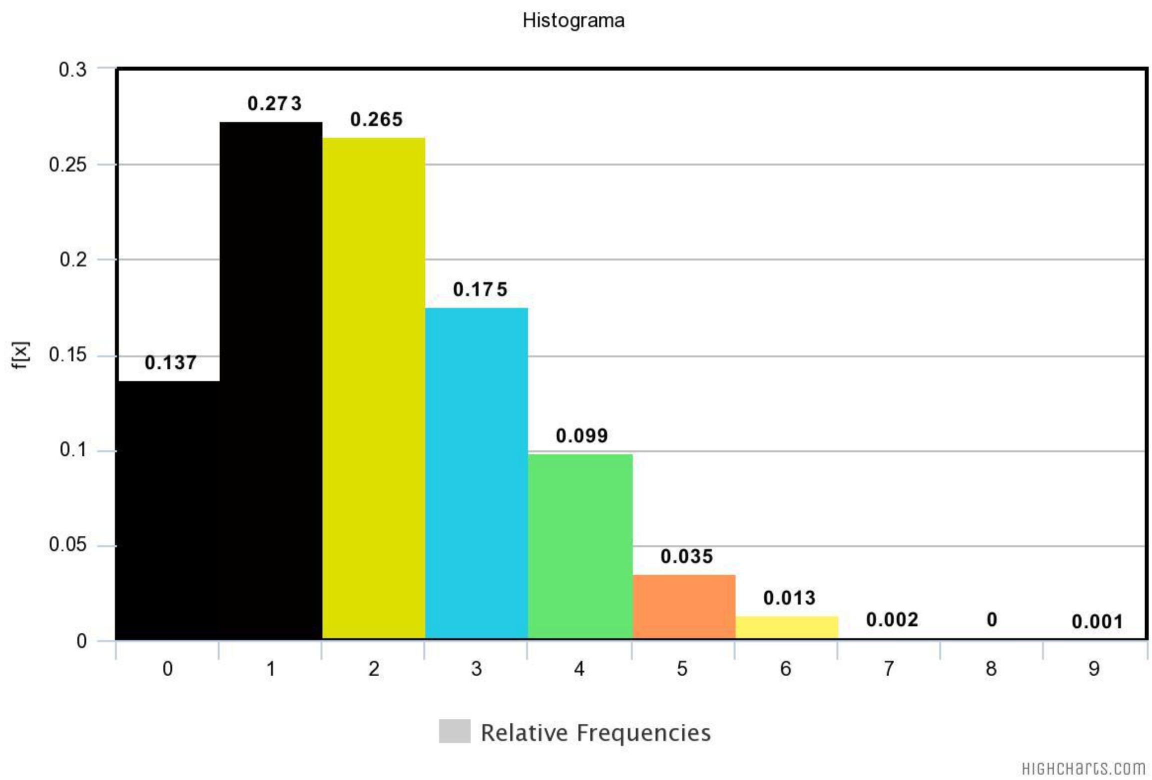
A)

Código na pasta 422

Resultado:



Com o seguinte histograma como resultado:



B)

Conferindo:

f[0] = (2^0 exp^(-2))/0! ~ 0.135

f[1] = (2^1 exp^(-2))/1! ~ 0.273

f[2] = (2^2 exp^(-2))/2! ~ 0.265

f[3] = (2^3 exp^(-2))/3! ~ 0.175

f[4] = (2^4 exp^(-2))/4! ~ 0.090

f[5] = (2^5 exp^(-2))/5! ~ 0.036

f[6] = (2^6 exp^(-2))/6! ~ 0.012

f[7] = (2^7 exp^(-2))/7! ~ 0.003

f[8] = (2^8 exp^(-2))/8! ~ 0.000

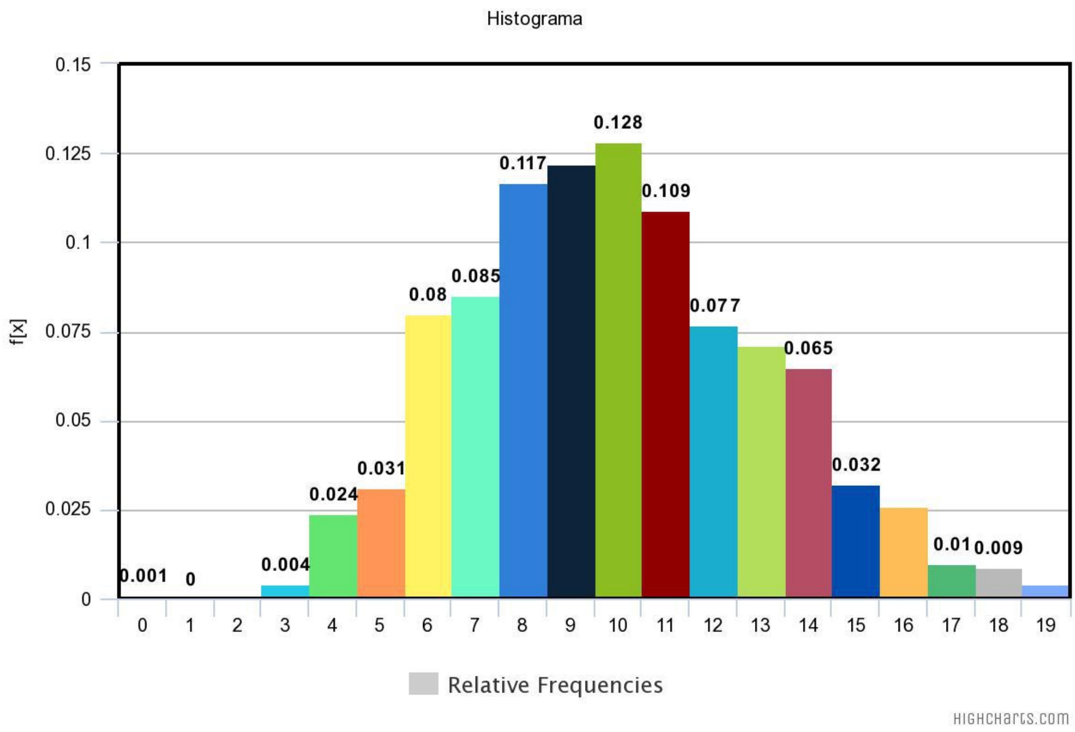
f[9] = (2^9 exp^(-2))/9! ~ 0.001

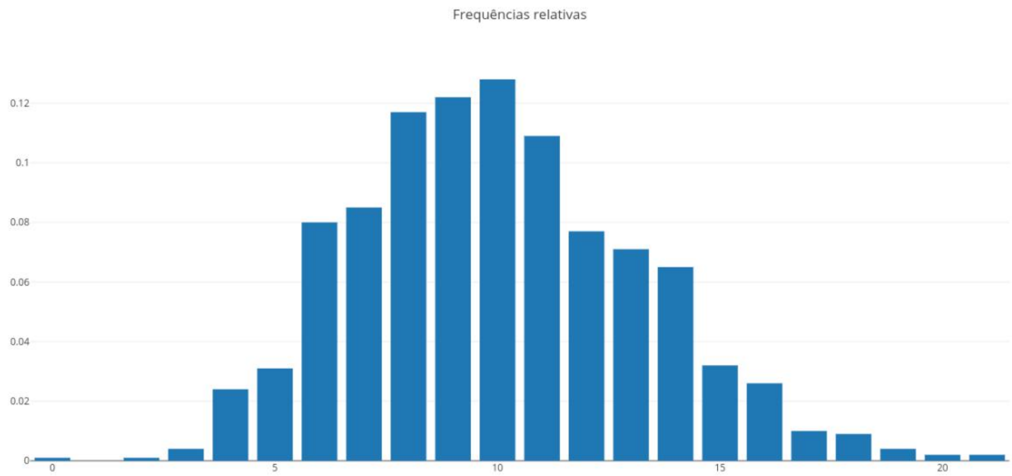
C)

Mesmo código da pasta 422 - Alterando BOLAS pra 10000 e MAXIMO para 22

```
/home/vitorbasso/Documents/MS/quarta atividade/codi...
Frequencias relativas:
f[0]=0.001
f[1]=0.000
f[2]=0.001
f[3]=0.004
f[4]=0.024
f[5]=0.031
f[6]=0.080
f[7]=0.085
f[8]=0.117
f[9]=0.122
f[10]=0.128
f[11]=0.109
f[12]=0.077
f[13]=0.071
f[14]=0.065
f[15]=0.032
f[16]=0.026
f[17]=0.010
f[18]=0.009
f[19]=0.004
f[20]=0.002
f[21]=0.002
```

Com o seguinte histograma como resultado:


Histograma
0.15
0.125
0.1
0.075
0.05
0.025
0
f[x]
0.001
0
0.004
0.024
0.031
0.08
0.085
0.117
0.128
0.109
0.077
0.065
0.032
0.01
0.009
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
Relative Frequencies
Highcharts.com


Frequências relativas
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20

D)

Exercise 4.3.1

(a) **Use program cdh to construct a continuous-data histogram like the one on the left in Example 4.3.1, but corresponding to a needle of length**

r = 2.

(b) Based on this histogram what is the probability that the needle will cross at least one line.

(c) What is the corresponding axiomatic probability that a needle of length r = 2 will cross at least one line?

Resposta:

A)

Código na pasta 431.

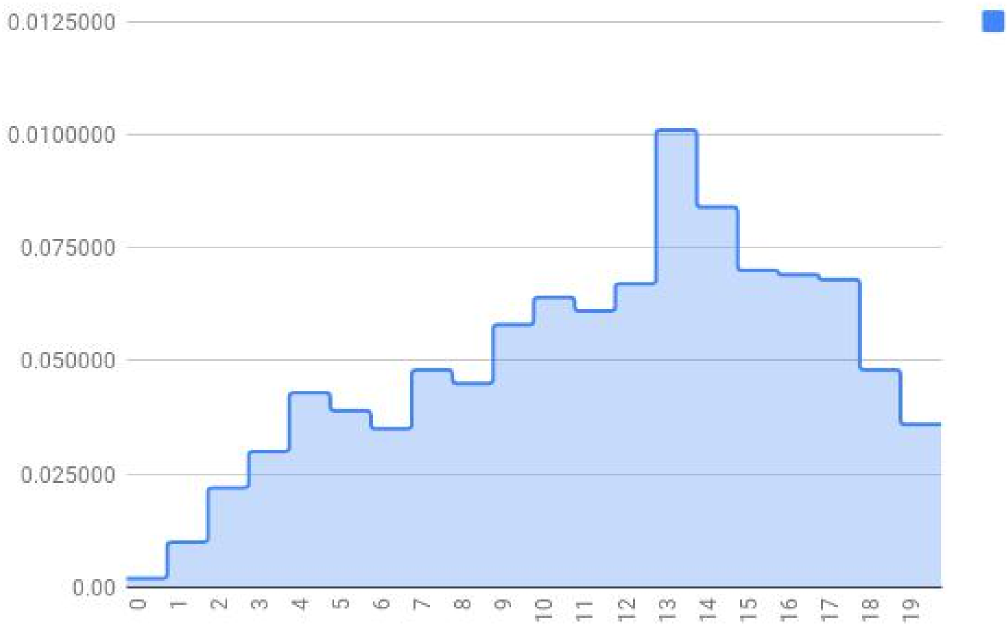
Resultado do programa:

```
vitorbasso@vitorbasso-asus: ~/Documents/MS/quarta atividade/codigos/431
File  Edit  View  Search  Terminal  Help
 bin      midpoint      count    proportion     density

   1       0.075          2        0.002        0.013
   2       0.225         10        0.010        0.067
   3       0.375         22        0.022        0.147
   4       0.525         30        0.030        0.200
   5       0.675         43        0.043        0.287
   6       0.825         39        0.039        0.260
   7       0.975         35        0.035        0.233
   8       1.125         48        0.048        0.320
   9       1.275         45        0.045        0.300
  10       1.425         58        0.058        0.387
  11       1.575         64        0.064        0.427
  12       1.725         61        0.061        0.407
  13       1.875         67        0.067        0.447
  14       2.025        101        0.101        0.673
  15       2.175         84        0.084        0.560
  16       2.325         70        0.070        0.467
  17       2.475         69        0.069        0.460
  18       2.625         68        0.068        0.453
  19       2.775         48        0.048        0.320
  20       2.925         36        0.036        0.240

sample size .... =    1000
mean ........... =   1.785
stdev .......... =   0.701
```

Histograma resultante:



B)

Realizando a soma da área sobre o histograma, obtemos a probabilidade que a variável randomica tem um valor entre a e b, portanto a soma da área desse histograma é de 0.939. Dessa forma, a probabilidade de que a agulha cruse pelo menos uma linha é de 93.9%.