



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
ESTATÍSTICA PARA ENGENHARIA
SEMESTRE 2022.2

TRABALHO COMPUTACIONAL DE ESTATÍSTICA

ALUNO: RYAN ERIK LIMA DA SILVA

MATRÍCULA: 542103

CURSO: ENGENHARIA DA COMPUTAÇÃO

PROFESSOR: CHARLES CASIMIRO CAVALCANTI

Todo o trabalho foi realizado na linguagem de programação R.

1 - Gerar uma sequência de variáveis aleatória (10.000 amostras) com distribuição uniforme no intervalo [4, 12].

- Descrever quais os parâmetros utilizados no algoritmo de geração.

Primeiramente, escolhi o seguinte método congruencial para que gerasse números aleatórios

$$x(n) = (a * x(n - 1) + c) \text{ módulo } m$$

Para isso escolhi números para adquirirem as seguintes variáveis.

```
a <- 4997
```

```
c <- 3333
```

```
m <- 10000
```

```
x <- c(7)
```

Para a escolha de a, escolhi um número que quando dividisse por 4 o módulo dessa divisão fosse igual a 1, já para a escolha de c, escolhi um número que fosse ímpar, para a variável m escolhi o total de amostras que a questão pedia, para ter uma maior variedades de números e para o primeiro elemento do vetor x foi um número escolhido ao acaso.

A geração de números aleatórios na linguagem utilizada foi a seguinte:

```
for (i in 2:10000){  
  x[i] = (a * x[i - 1] + c) %% m  
}
```

Utilizei um loop para que a operação dentro se repetisse 9999 vezes, para que os outros 9999 números fossem gerados. Por fim, dividi todos os elementos do veto por m - 1, para que os elementos do vetor ficassem entre o intervalo de 0 a 1.

$$x = x / (m - 1)$$

- Mostrar o histograma (construído por você!)

Para a montagem do histograma, precisei calcular a CDF inversa da distribuição uniforme que é descrita da seguinte forma:

$$x = y * (b - a) + a$$

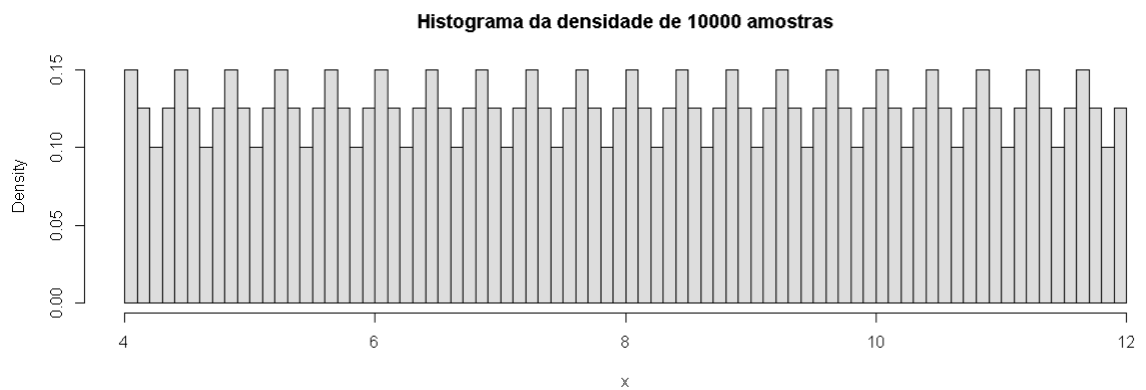
Em que, a e b representam o intervalo que a questão fornece, substituindo os valores a fórmula ficou dessa maneira.

$$x = y * 8 + 4$$

Com essa fórmula em mãos, coloquei para o programa aplicar em todos os números do

vetor e em seguida plotei o histograma usando o seguinte código.

```
x = x * 8 + 4  
hist(x,  
     breaks = 100,  
     freq = F,  
     main = "Histograma da densidade de 10000 amostras")
```

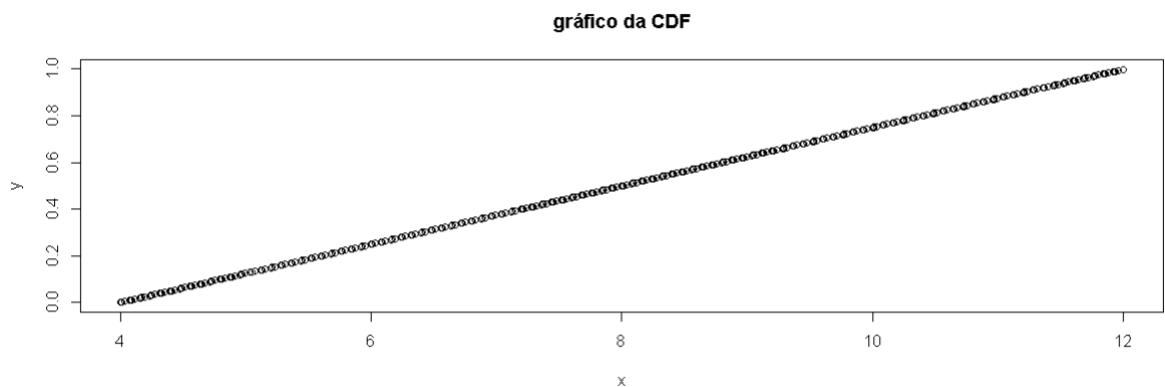


Fonte: elaborada pelo autor.

- Mostrar o gráfico da CDF (construído por você!)

Tendo os vetores x e y, plotei os dois para que gerassem o gráfico da CDF, utilizando o seguinte código.

```
plot(x, y, main = "gráfico da CDF")
```



Fonte: elaborada pelo autor.

- Calcular a média e variância da v.a. gerada. Explicar porque os valores teóricos e medidos são diferentes.

Para o cálculo da média e da variância fiz um código que calculasse de forma conhecida, os dois dados

```
soma = 0
for(i in 1:10000){
  soma = soma + x[i]
}
media = soma/10000

som = 0
for(i in 1:10000){
  som = som + (x[i] - media)^2
}
variancia = som / 10000
variância
```

O resultado da média pelo programa foi 7,99 e para a variância foi 5,33

O resultado teórico diz que a média é para ser 8,00 e a variância 5,33

$$\mu = \frac{a + b}{2} = \frac{4 + 12}{2} = \frac{16}{2} = 8,00$$
$$\sigma^2 = \frac{(b - a)^2}{12} = \frac{8^2}{12} = \frac{64}{12} = 5,33$$

Apesar da média não ter sido exatamente igual, deu um valor bem próximo.

2 - Gerar uma sequência de v.a. com distribuição exponencial (10.000 amostras) com parâmetro $\lambda = 3$.

- Descrever quais os parâmetros utilizados no algoritmo de geração.

Utilizei o mesmo método utilizado na questão 1.

- Mostrar o histograma (construído por você!)

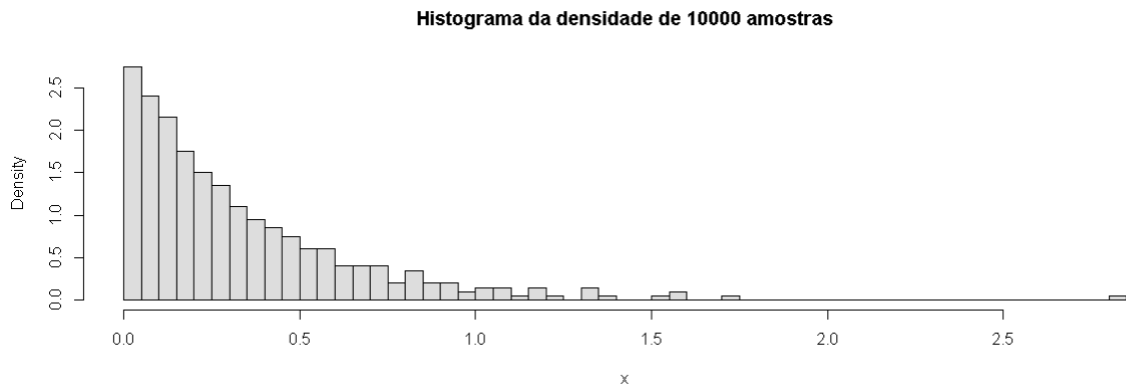
Para a montagem do histograma, precisei calcular a CDF inversa da distribuição exponencial que é descrita da seguinte forma:

$$x = -\frac{\log(1 - y)}{3}$$

Com essa fórmula em mãos, coloquei para o programa aplicar em todos os números do vetor e em seguida plotei o histograma usando o seguinte código.

$x = -\log(1-y)/3$

```
hist(x,
     breaks = 100,
     freq = F,
     main = "Histograma da densidade de 10000 amostras")
```

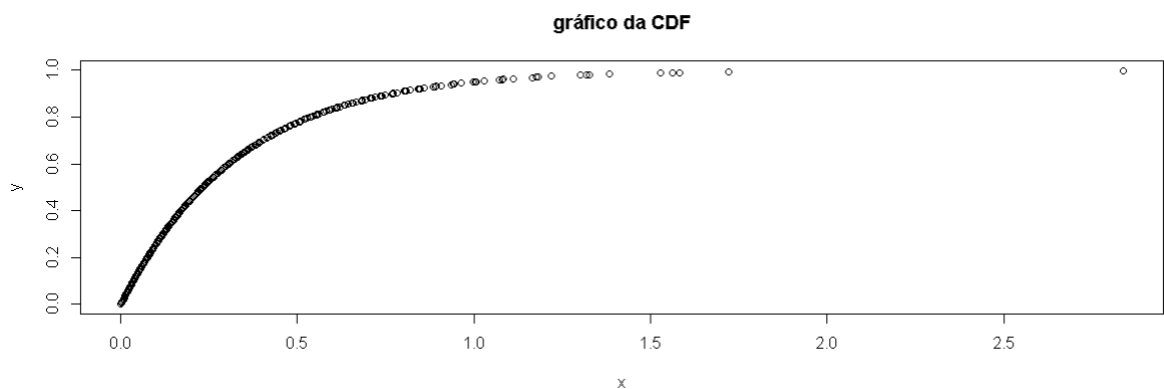


Fonte: elaborada pelo autor.

- Mostrar o gráfico da CDF (construído por você!)

Tendo os vetores x e y, plotei os dois para que gerassem o gráfico da CDF, utilizando o seguinte código.

```
plot(x, y, main = "gráfico da CDF")
```



Fonte: elaborada pelo autor.

- Calcular a média e variância da v.a. gerada. Explicar porque os valores teóricos e medidos são diferentes.

Para o cálculo da média e da variância fiz um código que calculasse de forma conhecida, os dois dados

```
soma = 0
```

```
for(i in 1:10000){
  soma = soma + x[i]
}
media = soma/10000
```

```
som = 0
for(i in 1:10000){
  som = som + (x[i] - media)^2
}
variância = som / 10000
variância
```

O resultado da média pelo programa foi 0,33 e para a variância foi 0,11

O resultado teórico diz que a média é para ser 0,33 e a variância 0,11

$$\mu = \frac{1}{\lambda} = 0,33$$

$$\sigma^2 = \frac{1}{\lambda^2} = 0,11$$

Tanto o valor para a média e variância foram bem preciso.

3 - Gerar uma sequência de v.a. com distribuição de Weibull (10.000 amostras) considerando $\lambda = 1$ e $\beta = 0,5$

- Descrever quais os parâmetros utilizados no algoritmo de geração.

Utilizei o mesmo método utilizado na questão 1.

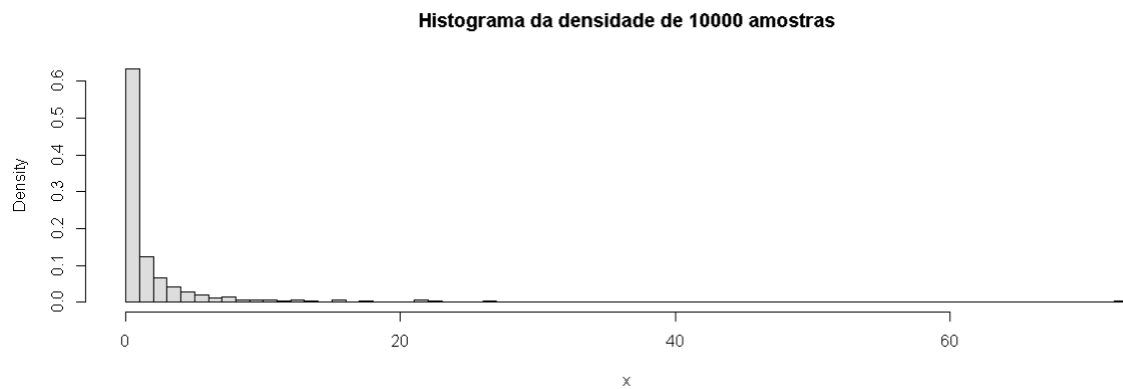
- Mostrar o histograma (construído por você!)

Para a montagem do histograma, precisei pesquisar a CDF inversa da distribuição de Weibull que é descrita da seguinte forma:

$$x = \lambda * (-\ln(1 - y))^{\frac{1}{\beta}}$$

Com essa fórmula em mãos, coloquei para o programa aplicar em todos os números do vetor e em seguida plotei o histograma usando o seguinte código.

```
x = (-log((1 - y), base = exp(1)))^(2)
hist(x,
  breaks = 100,
  freq = F,
  main = "Histograma da densidade de 10000 amostras")
```

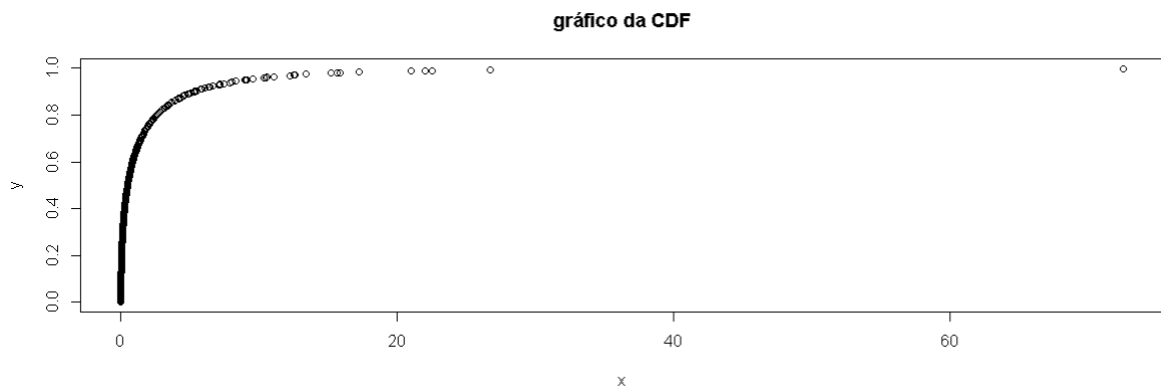


Fonte: elaborada pelo autor.

- Mostrar o gráfico da CDF (construído por você!)

Tendo os vetores x e y , plotei os dois para que gerassem o gráfico da CDF, utilizando o seguinte código.

```
plot(x, y, main = "gráfico da CDF")
```



Fonte: elaborada pelo autor.

- Calcular a média e variância da v.a. gerada. Explicar porque os valores teóricos e medidos são diferentes.

Para o cálculo da média e da variância fiz um código que calculasse de forma conhecida, os dois dados

```
soma = 0
for(i in 1:10000){
  soma = soma + x[i]
}
media = soma/10000
```

```
som = 0
for(i in 1:10000){
  som = som + (x[i] - media)^2
}
variância = som / 10000
variância
```

O resultado da média pelo programa foi 2,02 e para a variância foi 24,7

O resultado teórico diz que a média é para ser 2,00 e a variância 20,0

$$\mu = \lambda * \Gamma\left(1 + \frac{1}{\beta}\right) = 2,00$$

$$\sigma^2 = \lambda * \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2 \right] = 20,0$$

$$\text{ERRO PERCENTUAL} = \frac{24,7 - 20}{20} = 0,235 \rightarrow 23,5\%$$

O valor da variância deu uma diferença teve uma diferença significativa, com um erro percentual de 23,5%, talvez o método não tenha sido muito eficiente.

4 - Gerar uma sequência de v.a. com distribuição gaussiana (10.000 amostras) com média $\mu = 6$ e variância $\sigma = 4$.

- Descrever quais os parâmetros utilizados no algoritmo de geração.

Utilizei o mesmo método utilizado na questão 1.

- Mostrar o histograma (construído por você!)

Para a montagem do histograma, precisei pesquisar a CDF inversa da distribuição gaussiana que é descrita da seguinte forma:

$$x = \mu + \sigma\sqrt{2}\text{erf}^{-1}(2p - 1)$$

Em que, erf quer dizer função erro, que para utilizar a inversa dessa função tive que colocar no seguinte parâmetro

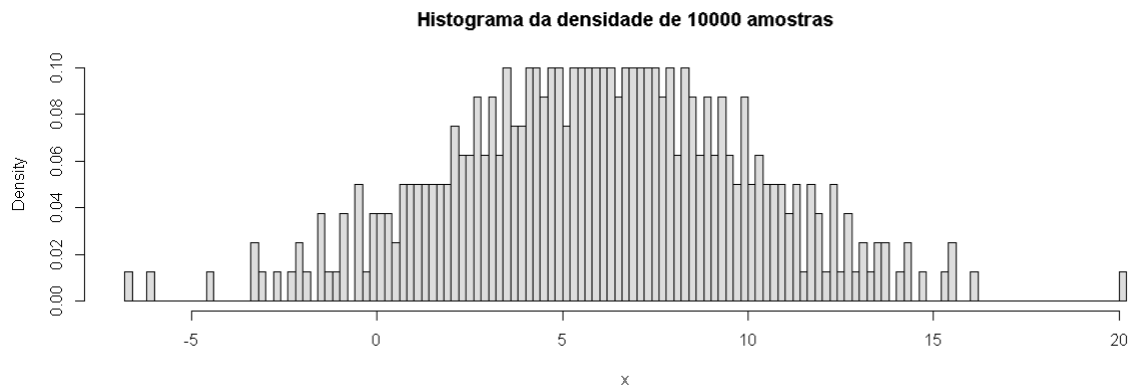
```
erfinv <- function (x) qnorm((1 + x)/2)/sqrt(2)
```

Com essa fórmula em mãos, coloquei para o programa aplicar em todos os números do vetor e em seguida plotei o histograma usando o seguinte código.

```
x = 6 + 4 * sqrt(2) * qnorm((2 * y)/2)/sqrt(2)
hist(x,
  breaks = 100,
  freq = F,
```



```
main = "Histograma da densidade de 10000 amostras")
```

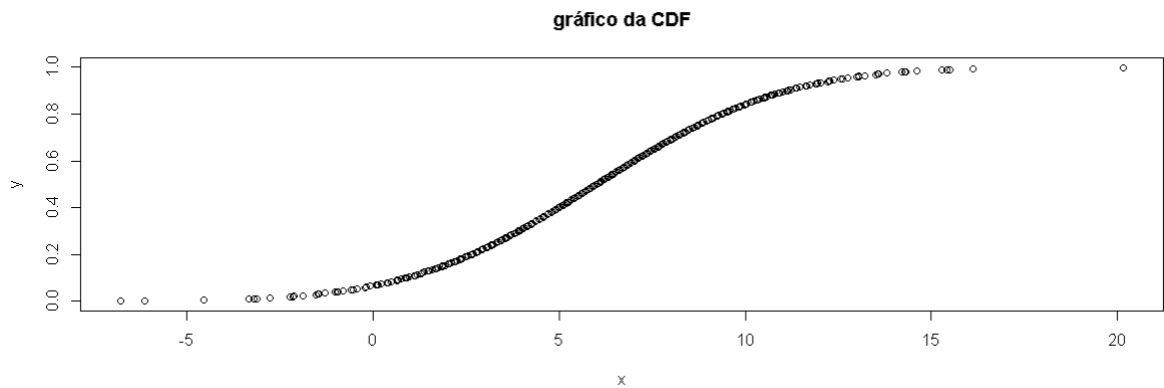


Fonte: elaborada pelo autor.

- Mostrar o gráfico da CDF (construído por você!)

Tendo os vetores x e y, plotei os dois para que gerassem o gráfico da CDF, utilizando o seguinte código.

```
plot(x, y, main = "gráfico da CDF")
```



Fonte: elaborada pelo autor.

- Calcular a média e variância da v.a. gerada. Explicar porque os valores teóricos e medidos são diferentes.

Para o cálculo da média e da variância fiz um código que calculasse de forma conhecida, os dois dados

```
soma = 0
```

```
for(i in 1:10000){
```

```
  soma = soma + x[i]
```

```
}  
media = soma/10000  
  
som = 0  
for(i in 1:10000){  
  som = som + (x[i] - media)^2  
}  
variância = som / 10000  
variância
```

O resultado da média pelo programa foi 5,98 e para a variância foi 16,2

O resultado teórico diz que a média é para ser 6,00 e a variância 16,0

A própria questão dá a média e o desvio padrão que quando eleva ao quadrado obtemos a variância. Os valores para a média e variância foram bem próximos.