Projeto 1 - Métodos para Cálculo de Zeros Reais

Universidade Federal da Bahia

Beatriz Schindler, Krysthian Lessa, Rodrigo da Silva Ramos

25 de maio de 2019

1 Introdução

Conhecer algumas quantidades pode ser de grande interesse em diversas áreas. Na estatística essas quantidades são chamadas de variáveis aleatórias. As variáveis aleatórias podem ser utilizadas para resolver uma gama de problemas práticos. Alguns exemplos de variáveis aleatórias contínuas são: altura ou peso de determinado grupo de pessoas, salário dos indivíduos, notas em uma prova numa turma, tempo de espera até a falha de um determinado aparelho, entre outras.

Em física, por exemplo, como toda medida está sujeita a erros experimentais, houve motivação para o desenvolvimento de uma distribuição de probabilidades para variáveis contínuas, denominada distribuição Normal, ou distribuição Gaussiana. É uma distribuição em forma de "sino"que concentra maior probabilidade para os valores em torno da média. A distribuição Normal e a distribuição Log-Normal são amplamente utilizadas em áreas como engenharia. Um variável tem distribuição Log-Normal quando seu logaritmo possui forma de uma distribuição Normal.

A função que associa cada valor x_i à sua probabilidade de ocorrência é chamada função densidade de probabilidade. A função de densidade de probabilidade da distribuição Log-Normal pode ser escrita como:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{\frac{-(ln(x)-\mu)^2}{2\sigma^2}\}; \qquad -\infty < x < +\infty$$

Na maioria das vezes, conhecer o conjunto de todos os casos (população de um país, total de peças fabricadas num ano, etc) é inviável e pode haver o interesse em tomar conclusões a cerca de uma população através de uma parcela dessa população (amostra). Para descrever as características de uma população através de uma variável aleatória, é necessário conhecer quantidades chamadas parâmetros, neste caso, μ, σ.

Em estatística, existem alguns métodos de estimação de parâmetros e um deles é o método da Máxima Verossimilhança. Para encontrar o Estimador de Máxima Verossimilhança (EMV) a partir de uma amostra de tamanho n, são necessários os passos:

1. Função de verossimilhança do parâmetro:

$$L(\sigma; x) = \prod_{i=1}^{n} f(x) = \prod_{i=1}^{n} (\frac{1}{\sqrt{2\pi}\sigma} \exp\{\frac{-(ln(x)-\mu)^2}{2\sigma^2}\})$$

2. Função de log-verossimilhança do parâmetro:

$$l(\sigma;x) = \prod_{i=1}^{n} log f(x) = \prod_{i=1}^{n} log(\frac{1}{\sqrt{2\pi}\sigma} \exp\{\frac{-(ln(x)-\mu)^2}{2\sigma^2}\})$$

3. Derivada da função de log-verossimilhança para obter o Estimador de Máxima Verossimilhança:

$$l'(\sigma;x) = \frac{-n}{\sigma} + \frac{\sum(ln(x_i)-\mu)^2}{\sigma^3}$$

O estimador será o zero real desta função:

$$\frac{-n}{\sigma} + \frac{\sum(ln(x_i)-\mu)^2}{\sigma^3} = 0$$

Quando o estimador é calculado para uma amostra, obtemos uma estimativa. O objetivo deste trabalho é utilizar os métodos numéricos aprendidos em sala de aula que encontram o zero real de uma função para encontrar a estimativa de máxima verossimilhança de uma amostra gerada aleatoriamente pelo software R, seguindo distribuição Log-Normal.

A função que desejaremos encontrar a raíz real será:

$$f(x) = \frac{-n}{\sigma} + \frac{\sum(ln(x_i)-\mu)^2}{\sigma^3}$$

cuja primeira derivada é:

$$f'(x) = \frac{n}{\sigma^2} - \frac{3\sum(ln(x_i)-\mu)^2}{\sigma^4}$$

e segunda derivada é:

$$f''(x) = \frac{-2n}{\sigma^3} + \frac{12\sum(ln(x_i)-\mu)^2}{\sigma^5}$$

2 Metodologia

2.1 Método de Newton

O método de Newton utiliza a função de Iteração $\phi(x) = x - \frac{f(x)}{f'(x)}$ de modo que a cada nova iteração $x_{k+1} = x_k - \frac{f(x)}{f'(x)}$ Dado um x_0 inicial, pertencente a um intervalo que contenha a raíz, se $f(x) \neq 0$ e $f"(x) \neq 0$ e $f(x_0)f"(x_0) > 0$ a convergência do método é garantida.

2.2 Método das Cordas

O método das cordas é uma alteração do método de Newton e utiliza a função de Iteração: $x_{k+1} = x_k - \frac{f(x_k)(x_k - xk-1)}{f(x_k)-f(x_{k-1}}$

Para este método são necessários dois valores iniciais, x_0 e x_1.

2.3 Simulação dos Dados

Os dados simulados através do software R, foram uma amostra de tamanho 100, seguindo distribuição Log-Normal, com $\mu = 3$ e $\sigma = 2$. Supondo uma situação em que possuímos essa amostra e conhecemos a média μ de sua população, estimaremos o parâmetro σ^2. Os resultados encontrados na amostra necessários para encontrar o estimador foram: $\sum(ln(x_i) - \mu)^2 = 8408621$ Portanto, a função que desejamos encontrar a raíz real será: $\frac{-100}{\sigma} + \frac{8408621}{\sigma^3}$

3 Resultados

Utilizamos as seguintes informações para gerar resultados através do algoritmo:

- Configuração da Máquina: F(10, 5, -15, 15);
- Intervalo: [280,0;300,0];
- Precisão: 0,0001.

Obtemos os seguintes resultados para cada um dos métodos:

Tabela 1: Resultados dos Métodos

k	Método de Newton			Método de Cordas/Secantes		
	Raiz Exata	Raiz Truncada	Raiz Arredondada	Raiz Exata	Raiz Truncada	Raiz Arredondada
0	289,157663	289,15	289,16	289,165681254273	289,16	289,17
1	289,970469	289,97	289,97	289,908972910451	289,90	289,91
2	289,976223	289,97	289,98	289,975754213303	289,97	289,98
3	-	-	289,98	-	-	-

Realizando o cálculo de Erro Relativo para Truncamento, temos os seguintes resultados

- Cota do ER de Truncamento: 0,00001

Tabela 2: Erro Relativo de Truncamento

k	Método de Newton	Método de Cordas/Secantes
	Raiz Exata	Raiz Exata
0	0,00002650181567	0,00001964744181
1	0,000001617408697	0,00003095174354
2	0,00002146084078	0,00001984416768

3.1 Gráficos

- Método de Newton

Figura 1: Raízes Exatas vs Raízes Truncadas

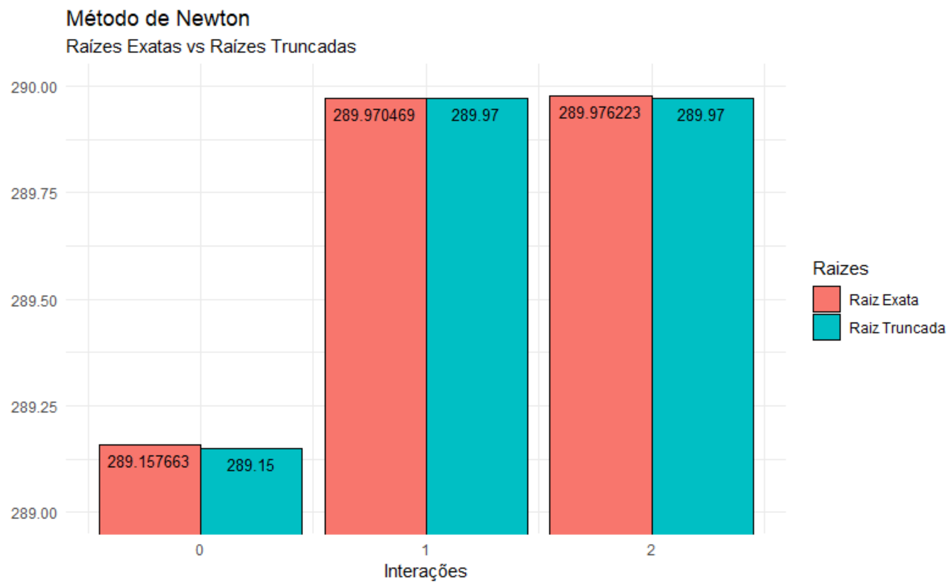


Figura 2: Raízes Exatas vs Raízes Arredondadas

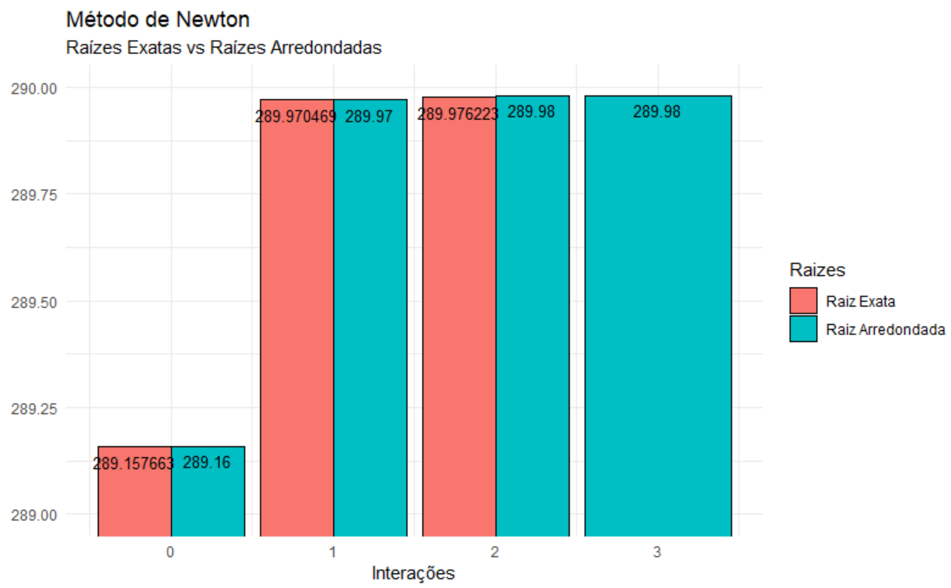
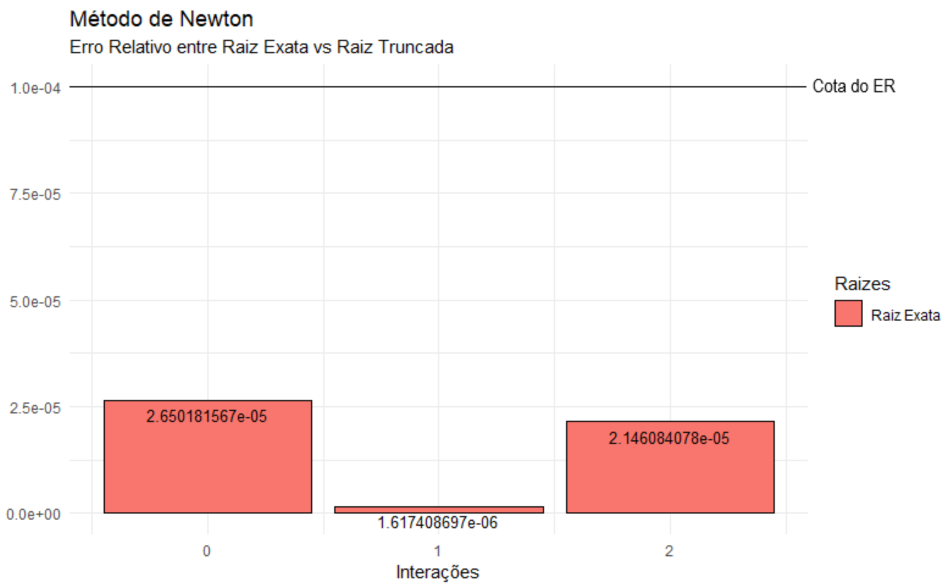


Figura 3: Erro Relativo entre Raízes Exatas vs Raízes Truncadas



- Método de Cortas/Secante

Figura 4: Raízes Exatas vs Raízes Truncadas

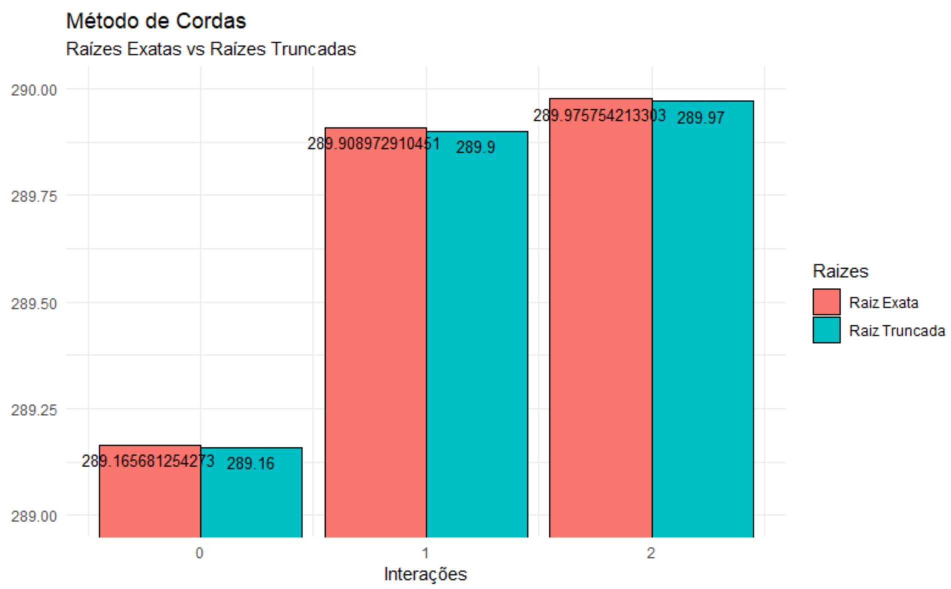


Figura 5: Raízes Exatas vs Raízes Arredondadas

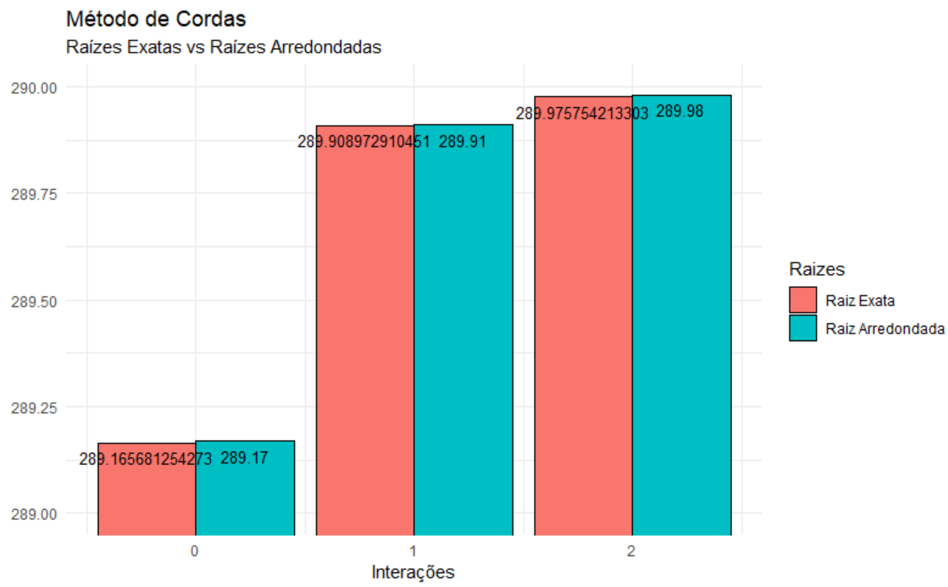
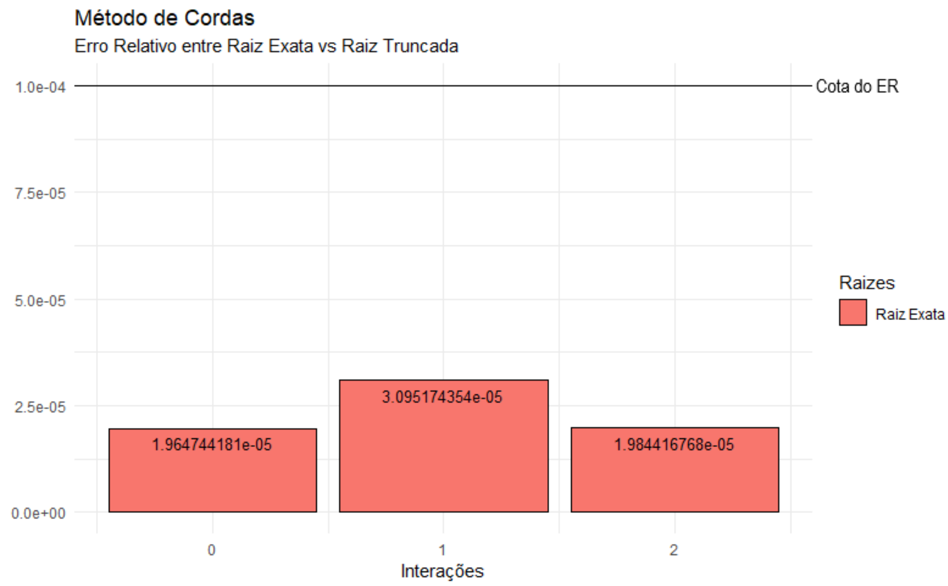


Figura 6: Erro Relativo entre Raízes Exatas vs Raízes Truncadas



4 Conclusão

Os dois métodos convergem ao mesmo tempo, exceto para quando a máquina simulada realiza arredontamentos no método de Newton, demorando uma iteração a mais. Para o método de Newton, nas 3 iterações, a raíz exata e a raíz truncada parecem estar bem próximas, mas quando analisamos o gráfico do erro relativo, podemos perceber que na segunda iteração o erro relativo foi menor que as outras. NO método das Cordas, a raíz truncada exibe uma maior diferença e nesse método, a segunda iteração foi a que apresentou maior erro relativo.