The Computational Complexity of Problems to Compute Intervals Wrappers for Random Variables Uniform, Exponential and Pareto

Alice Fonseca Finger[1]
Aline Brum Loreto
[1]Programa de Pós-Graduação
em Ciência da Computação
Centro de Desenvolvimento Tecnológico-CDTec
Universidade Federal de Pelotas-UFPel
Pelotas, Brasil
E-mail: {affinger, aline.loreto}@inf.ufpel.edu.br

Marcília Andrade Campos
Filipe Rafael Gomes Varjão[2]
Maria das Graças dos Santos
[2]Pós-Graduação em Ciência da Computação
Centro de Informática-CIn
Universidade Federal de Pernambuco-UFPE
Recife, Brasil
E-mail: {mac, frgv, mgs4}@cin.ufpe.br

***Abstract*—When working with floating point numbers the result is only an approximation of a real value and errors generated by rounding or by instability of the algorithms can lead to incorrect results. We can't affirm the accuracy of the estimated answer without the contribution of an error analysis. Interval techniques compute an interval range, with the assurance the answer belongs to this range. Using intervals for the representation of real numbers, it is possible to control the error propagation of rounding or truncation, between others, in numerical computational procedures. Therefore, intervals results carry with them the security of their quality. The goal is to analyze the computational complexity of the problems of computing enclosures intervals for random variables Uniform, Exponential and Pareto, showing that the intervals algorithms have linear complexity, which together with the security that interval mathematics provides, makes the use of intervals even more justified.**

***Index Terms*—Interval arithmetic, Numerical Algorithms and Problems, Statistical computing.**

I. INTRODUÇÃO

A análise intervalar surgiu com o objetivo inicial de controlar a propagação de erros numéricos em procedimentos computacionais. Mas, aparentemente, a matemática intervalar duplica o problema de representação dos números reais em processadores numéricos, uma vez que ao invés de operar com um número real, operam-se com dois. Entretanto, sua realização é feita por meio de números de ponto flutuante, isto é, os extremos do intervalo $\mathbf{x}$ são números de máquina $\underline{x}_{pf}$ e $\bar{x}_{pf}$ [1][2][3].

Os intervalos foram definidos com o objetivo inicial de automatizar a análise do erro computacional. Através da utilização de intervalos, tem-se um controle automático de erros com limites confiáveis, além de provas de existência e não existência de solução de diversos problemas.

Na matemática intervalar, o valor real x é aproximado por um intervalo $\mathbf{x}$, que possui como limites inferior e superior números de máquina de forma que o intervalo contenha x. O tamanho deste intervalo pode ser usado como medida para avaliar a qualidade de aproximação [4]. Os cálculos reais são substituídos por cálculos que utilizam a aritmética intervalar.

É importante ressaltar que existem questões do tipo: "Por que utilizar técnicas intervalares se existem na maioria dos sistemas de computadores bibliotecas matemáticas avançadas e eficientes, que resolvem a maioria dos problemas?".

A justificativa do uso de técnicas intervalares, segundo Rastchek *et al* [4], inicia pelo fato de que os computadores empregam aritméticas chamadas de ponto flutuante ou ponto fixo. Nestas aritméticas, números reais são aproximados por um subconjunto finito de números reais chamados números de máquina representáveis. Devido a esta representação são gerados erros quando um valor real de entrada é aproximado por um número de máquina; resultados intermediários gerados na execução de cada operação e que vão se acumulando; ou ainda, um outro tipo de erro que está relacionado com a incerteza dos dados de entrada, o que acontece muito em casos de experimentos físicos e químicos onde os dados de entrada são incertos.

No processo de resolução de problemas podem ser constatadas fontes de erros, tais como: propagação dos erros nos dados iniciais, arredondamento e erros de truncamento, causados ao se truncar sequências infinitas de operações aritméticas, após um número finito de etapas. Neste contexto percebe-se a importância de técnicas intervalares. Ressalta-se que uma resposta intervalar carrega com ela a garantia de sua incerteza. Um valor pontual não carrega medidas de sua incerteza. Mesmo quando uma análise de sondagem do erro é executada, o número resultante é somente uma estimativa do erro que pode estar presente.

Segundo Kearfott *et al* [5], são muitas as aplicações de intervalos e nas mais diversas áreas, tais como: programação matemática, manipulação de equações, análise e projeto de circuitos elétricos, psicologia matemática, estatística, equações diferenciais, física e muitos outros.

O método para implementação de operações e algoritmos intervalares em máquinas é realizado por meio do critério de semimorfismo proposto em [6][7]. Considerando que o controle do erro numérico é feito através do uso de intervalos ao invés de números reais, Kulisch [6] e Kulisch e Miranker [7] propuseram que a implementação da aritmética intervalar seja realizada através da chamada aritmética de exatidão máxima, o que significa a busca para que resultados numéricos ou sejam um número de ponto flutuante ou estejam entre dois números de ponto flutuantes consecutivos.

Por conveniência matemática, é importante associar números para cada resultado possível de um experimento aleatório, o que é feito com a definição de variáveis ou vetores aleatórios [8][9][10]. No estudo das variáveis aleatórias contínuas sobre o conjunto dos números reais, $\mathbb{R}$, um dos problemas é o cálculo de probabilidades, visto que é necessário resolver uma integral definida da função densidade que, na maioria das vezes, não possui primitiva explícita ou cuja primitiva não é simples de se obter. Embora integrais de funções densidade de probabilidade como a Uniforme, a Exponencial e a de Pareto, sejam resolvidas analiticamente, seu valor numérico é dado por aproximação, e portanto afetado por erros de arredondamento ou truncamento.

Considerando que métodos númericos devem ser usados para o cálculo de integrais a idéia é que estes sejam suportados pela matemática intervalar e a aritmética de exatidão máxima, o que implica que cálculos numéricos em computadores sejam realizados por meio das linguagens ou bibliotecas que tenham definidos o tipo intervalo e as operações sobre o tipo, usualmente denominadas de linguagens XSC (eXtended Scientific Computation). Dentre estas inclui-se o Python [11].

O termo complexidade, no contexto de algoritmos, refere-se aos recursos necessários para que um algoritmo possa resolver um problema sob o ponto de vista computacional, ou seja, à quantidade de trabalho despendido pelo algoritmo [12]. Quando o recurso é o tempo, são escolhidas uma ou mais operações fundamentais e então são contados os números de execuções desta operação fundamental na execução do algoritmo. Segundo Toscani *et al* [12] a escolha de uma operação como operação fundamental é aceitável se o número de operações executadas pelo algoritmo é proporcional ao número de execuções da operação fundamental.

No contexto da solucionabilidade de problemas, a complexidade pode ser considerada como uma propriedade do problema, o que significa dar uma medida independente do tratamento dado ao problema e independente do caminho percorrido na busca da solução.

O objetivo deste trabalho é realizar a análise da complexidade computacional dos problemas de computar intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto. Para tanto são analisados métodos e algoritmos intervalares [13], baseados na matemática intervalar [2] e na aritmética de exatidão máxima [6][7], que encapsulem probabilidades para as variáveis aleatórias Uniforme, Exponencial e Pareto. Para a Exponencial, o método proposto fundamenta-se em Caprani *et al* [14]. Para as variáveis Uniforme e Pareto o intervalo encapsulador foi definido a partir da função densidade. A escolha das variáveis relaciona-se com questões de avaliação de desempenho.

O trabalho organiza-se da seguinte maneira: na Seção II descreve-se o estado da arte e trabalhos realizados na área de complexidade computacional da matemática intervalar aplicada à estatística; na Seção III é apresentada a forma de representar números reais em intervalos e as definições intervalares para as variáveis aleatórias Uniforme, Exponencial e Pareto; na Seção IV apresenta-se a análise de complexidade dos algoritmos dos intervalos encapsuladores; na Seção V encontram-se os cálculos de erros contidos nos resultados obtidos através do uso de intervalos e na Seção VI apresentam-se as conclusões. Por fim as principais referências.

II. Aritmética Intervalar e Complexidade de Problemas Intervalares

A aritmética intervalar [1][2][3] é baseada no uso de intervalos fechados $[x_1, x_2]$ de números reais como elementos básicos e sua idéia do ponto de vista computacional, é: dada uma função $f(x)$ de variável real x pertencente a um intervalo $\mathbf{x} = [x_1, x_2]$ onde $x_1, x_2 \in \mathbb{R}$, a imagem da f é dada por

$$f(\mathbf{x}) = \{y \mid y = f(x), x_1 \leq \mathbf{x} \leq x_2\},$$

onde este em geral não é representado exatamente, mas é sempre possível determinar um intervalo $\mathbf{y} = [y_1, y_2]$ tal que $f(\mathbf{x}) \subseteq \mathbf{y}$, isto é $y_1 \leq f(x) \leq y_2$. Pode-se então definir uma função intervalar F associada a f pela transformação do intervalo $[x_1, x_2]$ em $[y_1, y_2]$, isto é:

$$f(x) \subseteq F(\mathbf{x}) = \mathbf{y}.$$

Esta função F, chamada extensão intervalar de f, deve ser aquela que possui o mínimo possível de diferença da imagem $f(\mathbf{x})$. O erro obtido no cálculo de $f(x)$ a partir do intervalo $\mathbf{x}$ é obtido através do diâmetro $w(F(\mathbf{x})) = y_2 - y_1$.

O cálculo de expressões na aritmética intervalar consiste em usar extensão das operações aritméticas junto com um conjunto de funções *standards*.

A aritmética intervalar utiliza um arredondamento especial, chamado arredondamento direcionado, o que significa que os resultados são arredondados para o menor e para o maior número de máquina que contém o resultado das operações, obtendo-se com isso um intervalo de máquina, com diâmetro mínimo, no qual a solução se situa.

Uma distinção importante na aritmética intervalar é da imagem intervalar de uma função da avaliação intervalar da função. A Imagem Intervalar de uma função f, contínua no intervalo $\mathbf{x}$, é definida como o intervalo limitado pelo mínimo da imagem de $f(x)$ e pelo máximo da imagem de $f(x)$, sendo x um elemento do intervalo $\mathbf{x}$.

$\mathbf{Im} = I(f, \mathbf{x}) = [min\{f(x) | x \in \mathbf{x}\}, max\{f(x) | x \in \mathbf{x}\}].$

Na computação intervalar, pode-se calcular o intervalo solução $\mathbf{y} = [\underline{y}, \overline{y}] = f(\mathbf{x_1}, \ldots, \mathbf{x_n})$ através de métodos de aproximação, técnicas de otimização, extensão intervalar e, ainda, por métodos considerados mais sofisticados [15] como forma centrada.

Uma das técnicas de otimização é o cálculo dos pontos de máximo e de mínimo da uma função [2], também conhecido como Imagem Intervalar [16].

Ratschek *et al* [4] afirmam que a maioria dos métodos de otimização têm, no mínimo, dois defeitos (falhas): o primeiro é que o método não garante que os pontos de mínimo, ou de máximo, possam ser encontrados para uma dada tolerância; o segundo é que o método, dependendo das condições da função, permite encontrar somente o mínimo local ao invés do global. Estes defeitos dificultam a solução de problemas de otimização global. A otimização global é considerada, por essa razão, um assunto intratável.

Segundo Ferson *et al* [17], historicamente, o primeiro método para computar o intervalo solução é a extensão intervalar [2], ou avaliação intervalar [16]. Este método está baseado no fato que em um computador, todo algoritmo consiste de operações elementares (aritméticas e lógicas). Para cada operação elementar $f(a,b)$, se são conhecidos os intervalos $\mathbf{a}$ e $\mathbf{b}$ para a e b, pode-se computar a imagem exata $f(\mathbf{a},\mathbf{b})$ através da aritmética intervalar definida por Moore em [1]. Na extensão intervalar, repete-se a computação formando o programa f passo-a-passo, substituindo cada operação elementar de números reais pela correspondente operação da aritmética intervalar.

Um dos primeiros problemas a ser analisado quanto a complexidade computacional foi o PBCI, ou seja, o Problema Básico de Computar o Intervalo imagem de uma função contínua e computável f.

(Problema Básico da Computação Intervalar: PBCI) O problema básico da computação intervalar é definido por:

DADOS:

- n intervalos racionais $\mathbf{x_i}$ (intervalos com extremos racionais), e

- uma função contínua e computável f que transforma n números reais $x_1,\dots,x_n$ em um número real $y = f(x_1,\dots,x_n)$.

CALCULAR: o intervalo dos possíveis valores de y:

$\mathbf{y} = [\underline{y},\overline{y}] = f(\mathbf{x_1},\dots,\mathbf{x_n}) = \{y \mid y = f(x_1,\dots,x_n), x_1 \in \mathbf{x_1},\dots,x_n \in \mathbf{x_n}\}$.

Moore [18] foi o primeiro a considerar a computação do intervalo imagem de uma função polinomial com entradas intervalares. Mais tarde, A.A. Gaganov [19] provou que o problema de calcular a imagem de um polinômio com dados intervalares é de fato computacionalmente intratável. Gaganov baseou-se no problema formalizado por Matiyasevich [20], de que nenhum algoritmo podia encontrar a imagem de um polinômio com entradas intervalares.

Com o propósito de responder a questão: "É possível ter um algoritmo que sempre calcula a imagem exata (isto é, os extremos inferior e superior do intervalo) em tempo razoável?", Kreinovich *et al* [15] realizaram a análise da complexidade do PBCI, verificando que o mesmo pertence à classe de problemas NP-Difícil.

É importante observar que a NP-dificuldade do PBCI está relacionada com o processamento dos dados de entrada do problema. No PBCI, os dados de entrada de uma função contínua f são valores intervalares, e a forma de calcular o intervalo imagem do PBCI é através da imagem intervalar [16]. Neste caso calculam-se os valores dos extremos inferior e superior do intervalo $\mathbf{Im}$, isto é, no pior caso quando tem-se m números arbitrários de valores, deve-se considerar todos os valores compreendidos entre $\underline{x}$ e $\overline{x}$ do intervalo $\mathbf{x}$ para encontrar estes extremos. Como no problema de computar a imagem de uma função de várias variáveis $f(x_1,...,x_n)$, tem-se $\mathbf{x_i}$ intervalos, $i = 1,\dots,n$, calcula-se o extremo inferior $\underline{y}$ avaliando todos os valores de todos os n intervalos (da mesma forma para $\overline{y}$), totalizando n^m avaliações numéricas (operações fundamentais), o que caracteriza uma complexidade exponencial. Conforme o tamanho da entrada do problema, o tempo de processamento da computação cresce exponencialmente, e pela dificuldade do processamento de todos os dados, não se tem algoritmo de tempo de processamento polinomial conhecido.

Outro fato importante, e que completa a caracterização do problema como NP-Difícil, é que o problema em questão é um problema de decisão, pois deseja-se saber se existe o intervalo $\mathbf{y} = [\underline{y},\overline{y}]$ que contenha a solução aproximada do problema. Observa-se, ainda, que o problema PBCI é computável; que não foi provado que o problema PBCI é NP; e que a questão em relação a complexidade deste problema está em aberto.

Mesmo o PBCI sendo NP-Difícil, Ferson *et al* [21] consideraram valores intervalares $[\underline{x}_i,\overline{x}_i]$, $i = 1,\dots,n$, para cada um dos n valores de amostra $\{x_1,\dots,x_n\}$ como dados de entrada de problemas da estatística, mais precisamente para os indicadores estatísticos variância, covariância e coeficiente de correlação. Da análise da complexidade computacional do problema de computar o valor da variância va contido em um intervalo $\mathbf{VA} = [\underline{va},\overline{va}]$, foi verificado que o mesmo é NP-Difícil [22][17], e considerando algumas restrições sobre os dados de entrada, Ferson *et al* [22] e Wu [23], apresentaram algoritmos razoáveis [12] que computavam o extremo inferior $\underline{va}$ do intervalo $\mathbf{VA}$ em tempo polinomial. Os problemas de computar o intervalo da covariância intervalar $\mathbf{CO}$ e o intervalo da correlação intervalar $\mathbf{CC}$ também pertencem a classe de problemas NP-Difícil [21] [24], porém para estes problemas não foram consideradas restrições sobre os dados de entrada para melhorar o resultado da complexidade.

A análise da complexidade do problema de computar medidas de dispersão com valores intervalares justifica-se devido a afirmação de Traylor [25] de que nenhum algoritmo razoável [12] é possível para computar a estimativa de um intervalo ótimo, isto é, a menos que o tempo de processamento seja exponencial. Esta afirmação foi constatada nos resultados encontrados na bibliografia pesquisada sobre a complexidade de problemas com entradas intervalares, porém não com a utilização da computação intervalar (operações aritméticas intervalares e aritmética de exatidão máxima) devido a superestimação no intervalo obtido, e sim com processamento de dados utilizando as operações de aritmética real (computação da imagem intervalar).

Loreto, em [26], verificou por meio da análise de complexidade computacional que os problemas de medidas de tendência central, dispersão e separatrizes, com entradas intervalares e algoritmos intervalares (aplicando a extensão intervalar na solução destes problemas), pertencem à classe de problemas P, ou seja, propôs algoritmos intervalares de tempo polinomial e que retornam como solução intervalos com qualidade de aproximação.

Em decorrência dos resultados obtidos em Loreto [26], o presente artigo propõe a aplicação da extensão intervalar na solução dos problemas de computar intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto, considerando que no estudo das variáveis aleatórias contínuas sobre o conjunto dos números reais, $\mathbb{R}$, um dos problemas é o cálculo de probabilidades, onde deve-se resolver uma integral definida da função densidade que não possui primitiva explícita ou cuja primitiva não é simples de se obter.

III. Intervalos encapsuladores para Uniforme, Exponencial e Pareto

A distribuição de probabilidade de uma variável aleatória contínua, [8] [9] [10], é caracterizada por sua função densidade de probabilidade, a qual satisfaz as propriedades:

(i) $f(x) \geq 0$,
(ii) $\int_a^b f(x)dx = P(a \leq x \leq b), a < b$,
(iii) $\int_{-\infty}^{\infty} f(x)dx = 1$.

O item (ii) indica que a probabilidade da variável aleatória assumir valor em um intervalo é dada pela integral da função nesse intervalo. Entretanto, o cálculo dessa probabilidade implica em resolver dois tipos de problemas: (1) encontrar primitivas na forma analítica, o que é possível no caso da Uniforme, Exponencial e Pareto e (2) o valor da probabilidade, em geral, é um número real, necessariamente não representável em computadores, isto é, não é um número de ponto flutuante [27][28][29][30].

O método de Simpson Intervalar [14] é fundamentado na propriedade aditiva da integral definida e no teorema do valor médio para integrais [31]. Supondo que uma função f é quatro vezes continuamente derivável em um intervalo A = [a,b], o método retorna um intervalo que encapsula (contém) a integral definida.

1) Uniforme: A distribuição Uniforme possui densidade com primitiva na forma analítica, entretanto valores da probabilidade podem não ser representáveis em computadores.

A densidade de uma Uniforme X no intervalo $A = [a, b]$ é

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b], \\ 0, & x \notin [a,b]. \end{cases}$$

Seja $B = [c, d] \subset \mathbb{R}$. Para calcular $P(c < X \leq d)$ tem-se as seguintes situações possíveis:

(i) Se $d \leq a$,

$$\text{P(c} < \text{X} \leq d) = 0$$

(ii) Se $c < a \leq d \leq b$,

$$\text{P(c} < \text{X} \leq d) = P(a \leq X \leq d) = \int_a^d \tfrac{1}{b-a}dx = \tfrac{d-a}{b-a}.$$

(iii) Se $a < c \leq b \leq d$,

$$\text{P(c} < \text{X} \leq d) = P(c \leq X \leq b) = \int_c^b \tfrac{1}{b-a}dx = \tfrac{b-c}{b-a}.$$

(iv) Se $a \leq c < d \leq b$,

$$\text{P(c} < \text{X} \leq d) = P(c \leq X \leq d) = \int_c^d \tfrac{1}{b-a}dx = \tfrac{d-c}{b-a}.$$

(v) Se $c \leq a < b \leq d$,

$$\text{P(c} < \text{X} \leq d) = P(a \leq X \leq b) = \int_a^b \tfrac{1}{b-a}dx = \tfrac{b-a}{b-a} = 1.$$

(vi) Se $c \geq b$,

$$\text{P(c} < \text{X} \leq d) = 0.$$

Analisando as probabilidades acima, observa-se que os cálculos anteriores poderiam ter sido realizados através da seguinte definição:

Definição 1.

$$P(c < X \leq d) = \begin{cases} \dfrac{w([c,d] \cap [a,b])}{b-a}, & [c,d] \cap [a,b] \neq 0, \\ 0, & [c,d] \cap [a,b] = 0. \end{cases}$$

onde w é o diâmetro do intervalo.

A densidade da Uniforme tem primitiva na forma analítica, portanto, pelo Teorema Fundamental do Cálculo Integral, pode-se calcular qualquer integral definida, por exemplo

$$P(c < X \leq d) = \int_c^d f(x)dx,$$

diretamente desta primitiva.

Entretanto, a existência da primitiva na forma analítica não impediu a ocorrência de problemas numéricos relacionados com o cálculo das integrais, como visto no exemplo acima, os quais justificam a busca para um intervalo encapsulador para a probabilidade real.

A definição a seguir, a função $UNIF$, é uma proposta de como calcular um intervalo encapsulador para probabilidades da Uniforme no intervalo $[a, b]$. Enfatiza-se que $UNIF$ é uma extensão intervalar inclusão monotônica [1][2][3].

Definição 2. *Encapsulando Probabilidades para a Uniforme.*

$$UNIF(c,d) = \left[\underline{\frac{w([c,d] \cap [a,b])}{b-a}}, \overline{\frac{w([c,d] \cap [a,b])}{b-a}}\right],$$

$[c, d] \cap [a, b] \neq \emptyset$.

Se $[c, d] \cap [a, b] = \emptyset$,então $UNIF(c, d) = 0_\varepsilon$, onde 0_ε é o menor intervalo de máquina que contém o zero [32].

$UNIF$, definida com suporte na matemática intervalar [1][2][3] e na aritmética de exatidão máxima [6][7], fornece o intervalo de menor amplitude para encapsular probabilidades para a distribuição Uniforme.

2) Exponencial: Uma variável aleatória contínua com função densidade dada por

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & 0 \leq x \leq \infty, \\ 0, & x < 0. \end{cases}$$

com $\alpha \in \mathbb{R}_+^*$, é uma variável aleatória Exponencial.

Seja X uma Exponencial com parâmetro α. Para computar a probabilidade intervalar para esta variável aleatória usa-se o método de Simpson Intervalar, in Caprani *et al* [14], para resolver integrais numéricas. Como a probabilidade da variável X assumir valores negativos é 0, pode-se restringir, na aplicação do método de Simpson Intervalar, o domínio de f e de sua derivada de ordem 4, $f^{(4)}$, ao conjunto $\mathbb{R}_+$. Desta forma, tem-se, respectivamente,

$$F_E(X) = \alpha[e^{-\alpha\bar{x}}, e^{-\alpha\underline{x}}], \quad \underline{x} > 0$$

e

$$G_E(X) = \alpha^5[e^{-\alpha\bar{x}}, e^{-\alpha\underline{x}}],$$

como extensão intervalar para a função densidade $f_X|\mathbb{R}_+$ e a derivada $f_X^{(4)}|\mathbb{R}_+$.

Verifica-se que

$$F_E(X) = f|\mathbb{R}_+(X) = \tilde{f}|\mathbb{R}_+(X) = \bar{f}|\mathbb{R}_+(X)$$

e

$$G_E(X) = f^{(4)}|\mathbb{R}_+(X) = \tilde{f}^{(4)}|\mathbb{R}_+(X) = \bar{f}^{(4)}|\mathbb{R}_+(X), \quad (1)$$

onde $\tilde{f}$ é a cobertura intervalar da função f [1][2][3] e $\bar{f}$ é extensão unida da função f [1][2][3].

Teorema 1. *G_E é uma extensão intervalar monotônica linear para $f^{(4)}|\mathbb{R}_+$.*

Prova 1. $d(G_E(Y), \tilde{f}^{(4)}|\mathbb{R}_+(Y)) = 0 \leq Kw(Y), \forall K > 0$, $\forall Y \subseteq X$, *por* (1). □

3) Pareto: A distribuição de Pareto possui densidade com primitiva na forma analítica, porém os valores da probabilidade, como na Uniforme, podem não ser representáveis em computadores.

Uma variável aleatória contínua tem distribuição de Pareto com parâmetros α e c (onde α e c são constantes positivas) se a sua densidade é:

$$f(x) = \alpha\frac{c^\alpha}{x^{\alpha+1}} \quad \text{para} \quad x \geq c.$$

Seja X uma Pareto com parâmetros α e c. Para computar o intervalo encapsulador da probabilidade intervalar para esta variável aleatória, define-se a função densidade como extensão intervalar:

$$F_P(X) = \alpha.\frac{c^\alpha}{X^{\alpha+1}} = \alpha\left(\left[\frac{c^\alpha}{\overline{x}^{\alpha+1}}, \frac{c^\alpha}{\underline{x}^{\alpha+1}}\right]\right),$$

para $\quad 0 \notin X = [\underline{x}, \overline{x}]$.

IV. Complexidade: Análise Sobre o Probelma de Computar Cada Variável Aleatória

Entre as medidas de complexidade, o critério de avaliação mais utilizado é a complexidade no pior caso ou simplesmente Complexidade [12], como será chamada neste trabalho. Neste critério a complexidade é tomada como a máxima para qualquer entrada de um dado "tamanho".

A complexidade também pode ser vista como uma propriedade do problema, o que significa dar uma medida independente do tratamento dado ao problema, independente do caminho percorrido na busca da solução, portanto independente de algoritmos. Alguns problemas são bem comportados, isto é, permitem chegar a limites de complexidades bem definidos, outros estão em classes com contornos não bem claros [12].

Os problemas definidos na literatura [17][22][33][23] consideram intervalos $\mathbf{x_i}$ contidos no domínio da função f, isto é, $Dom(f) = \mathbb{R}^n$, onde n é o número de argumentos da função.

Neste trabalho, considera-se como domínio dos problemas de computar os intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto, um conjunto de valores intervalares, ou seja, $Dom(f) = \mathbb{IR}^n$ (n é o número de argumentos da função). Ressalta-se que, quando se altera a definição do domínio (dados de entrada) de um problema, define-se um novo problema.

Importante salientar que os problemas em análise no presente trabalho não foram abordados na literatura pesquisada, e sim problemas relacionados à indicadores estatísticos e cálculo de probabilidades intervalares.

Os problemas de computar os intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto é um problema de localização associado ao problema de decisão, pois deseja-se saber se os intervalos encapsuladores contêm as soluções reais para estes problemas.

Os algoritmos desenvolvidos para o cáculo de intervalos encapsuladres para as variáveis aleatórias Uniforme, Exponencial e Pareto são encontrados em [13]. Estes algoritmos utilizam aritmética intervalar definida por Moore [1] e a extensão intervalar [2][3] como método da computação intervalar.

A seguir, apresenta-se a análise de complexidade de cada algoritmo elaborado para cada variável aleatória Uniforme, Exponencial e Pareto, descrevendo as ordens de complexidade.

Análise do algoritmo de intervalos encapsuladores para a variável aleatória Uniforme: o algoritmo utiliza a estrutura de lista para armazenar os valores dos intervalos. Dentro dessa estrutura de dados utilizada, a operação executada é o cálculo de elementos mínimo e máximo dessa lista, cálculo que é executado em tempo linear. Após, faz-se também uma remoção de elementos de uma lista, o que também é executado em tempo linear.

Com isso, pode-se afirmar que a complexidade do algoritmo de intervalos encapsuladores para a variável aleatória Uniforme é de ordem linear, $O(n)$.

Análise do algoritmo de intervalos encapsuladores para a variável aleatória Exponencial: no algoritmo são encontradas duas estruturas "for". O primeiro faz atribuições, comparações e operações aritméticas para todos os subintervalos p dos intervalos para a variável. Logo, possui uma complexidade de ordem linear. Já o segundo "for"percorre todos os elementos de uma lista fazendo atribuições e operações aritméticas, ou seja, tem um custo computacional linear. Como essas estruturas ocorrem uma após a outra, considera-se apenas uma soma dessas duas complexidades. Assim, afirma-se que a complexidade do algoritmo de intervalos encapsuladores para a variável aleatória Exponencial também é de ordem linear, $O(n)$.

Análise do algoritmo de intervalos encapsuladores para a variável aleatória Pareto: o algoritmo contém estruturas de testes, realiza atribuições, comparações e operações matemáticas, onde cada um possui complexidade constante. Logo, o algoritmo de intervalos encapsuladores para a variável aleatória Pareto possui uma complexidade constante, $O(1)$.

A partir da investigação da complexidade computacional dos problemas baseados em computar os intervalos encapsuladores das variáveis aleatórias, realiza-se a classificação destes quanto a classe de complexidade. Conhecendo as classes de complexidade a que os problemas pertencem, os projetistas de algoritmos, e até mesmo os profissionais da área de estatística, podem ter uma medida real quanto às soluções disponíveis e a expectativa de melhorar esses resultados.

A Tabela I apresenta as classes de complexidade que os problemas de computar os intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto pertencem.

Tabela I
VARIÁVEIS ALEATÓRIAS UNIFORME, EXPONENCIAL E PARETO E COMPLEXIDADE DOS PROBLEMAS DOS INTERVALOS ENCAPSULADORES

Variáveis Aleatórias	**Complexidade dos Problemas dos Intervalos Encapsuladores**
Uniforme	P
Exponencial	P
Pareto	P

Verifica-se que os problemas de computar os intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto, pertencem à classe de problemas P, ou seja, as soluções (através da extensão intervalar) para estes problemas possuem algoritmos intervalares de tempo linear. Os resultados obtidos na análise da complexidade ratificam a viabilidade da computação dos intervalos encapsuladores para as respectivas variáveis.

V. QUALIDADE DOS INTERVALOS ENCAPSULADORES

Segundo Ratschek [4], os computadores utilizam aritmética de ponto flutuante. Nesta aritmética números reais são aproximados por um subconjunto de números reais chamados representação numérica da máquina. Devido a esta representação são gerados dois tipos de erros: o primeiro ocorre quando uma entrada de valor real é aproximada por um número de máquina e o segundo é causado por resultados intermediários aproximados pelos números de máquina. A aritmética intervalar fornece uma ferramenta para estimar e controlar esses erros automaticamente. No lugar de aproximar um valor real x por um número de máquina, o valor real x, usualmente desconhecido, é aproximado por um intervalo $\mathbf{x}$ tendo número de máquina nos extremos inferior e superior. O intervalo $\mathbf{x}$ contém o valor x. O comprimento (ou diâmetro) deste intervalo pode ser usado como medida para qualidade da aproximação. Os cálculos são executados usando intervalos ao invés de números reais e, consequentemente, a aritmética real é substituída pela aritmética intervalar. A computação com utilização de intervalos fornece as seguintes estimativas para o erro:

- *Erro Absoluto*: $|x - m(\mathbf{x})| < \frac{w(\mathbf{x})}{2}$,
  onde $m(\mathbf{x}) = (\frac{\underline{x}+\overline{x}}{2})$ é o ponto médio do intervalo $\mathbf{x}$ e $w(\mathbf{x}) = \overline{x} - \underline{x}$ é o diâmetro do intervalo $\mathbf{x}$;
- *Erro Relativo*: $\left|\frac{x-m(\mathbf{x})}{x}\right| \leq \frac{w(\mathbf{x})}{2min|\mathbf{x}|}$ se $0 \notin \mathbf{x}$.

Observa-se que nas medidas de erros, utiliza-se o ponto médio $m(\mathbf{x})$ do intervalo $\mathbf{x}$ para medir a distância do valor real em relação ao valor pontual (ponto médio) do intervalo. Segundo Sunaga [34] a interpretação usualmente aceita para um intervalo no contexto da aritmética intervalar é a de envoltória intervalar de um número real. Esta semântica, de envoltória intervalar, sugere a representação dos intervalos na forma $m(\mathbf{x}) \pm \frac{w(\mathbf{x})}{2}$, aludindo à idéia que o ponto médio $m(\mathbf{x})$ seria o número real "medido" e o raio $\frac{w(\mathbf{x})}{2}$ indicaria a incerteza gerada pelas restrições de precisão e ambientais existentes. Dessa forma o valor exato estaria limitado pelo intervalo apresentado.

Aplicam-se estas medidas de erros nos intervalos encapsuladores obtidos para as variáveis aleatórias contínuas Uniforme, Exponencial e Pareto com o objetivo de verificar a qualidade do intervalo solução, obtido após o processamento de operações aritméticas intervalares. Através das medidas de erros verifica-se se que o método da extensão intervalar fornece como resposta intervalos que englobam a resposta real exata.

Para todos os algoritmos intervalares, construídos por Varjão [13] para serem executados no ambiente intervalar IntPy, realiza-se a análise da complexidade para determinar a ordem de complexidade dos mesmos.

Os algoritmos desenvolvidos para calcular os intervalos encapsuladores das variáveis aleatórias Uniforme e Exponencial possuem complexidade de ordem linear, O(n), e para a variável Pareto a complexidade é constante. Salienta-se que a complexidade obtida destes algoritmos deve-se à implementação na linguagem de programação Python e no pacote IntPy, o qual reconhece o tipo intervalo e as respectivas operações para este tipo.

Para verificação da qualidade de aproximação nos intervalos encapsuladores, apresentam-se exemplos de cálculos numéricos de cada variável aleatória Uniforme, Exponencial e Pareto usando sistema de ponto flutuante F(10, 14, -10, 10) e arredondamento direcionado [7]. As operações intervalares envolvidas podem ser encontradas em Moore [1].

Na validação das soluções intervalares encontradas no trabalho, realiza-se um comparativo entre as soluções dos cálculos reais das variáveis aleatórias, realizados no software NetBook [35], e as soluções dos cálculos intervalares das mesmas variáveis, realizados no IntPy [13].

O NetBook é uma ferramenta gratuita desenvolvida no CIn/UFPE, com o objetivo de suportar análise de desempenho de sistemas de comunicação, em particular, redes de computadores. É composto por quatro módulos: Estatística, Geração, Transformação e Gráfico. Exceto pelo módulo de geração de tráfego auto-similar, que foi implementado em C++, todos os demais módulos do NetBook foram implementados em Java.

IntPy é um pacote desenvolvido em Python que implementa o tipo intervalo e as operações sobre o tipo. A versão atual, a 0.1.3, implementa intervalos com arredondamentos direcionados, reconhece entrada de intervalos como strings e realiza as operações entre intervalos.

A. Uniforme

Exemplo 1. *Um ponto é escolhido ao acaso no segmento de reta [0,3]. Qual é a probabilidade de que o ponto escolhido esteja entre 1 e 2?*

Solução. *Fazendo X representar a coordenada do ponto escolhido, tem-se a função densidade de probabilidade f(x):*

$$f(x) = \begin{cases} \dfrac{1}{3}, & x \in [0,3], \\ 0, & x \notin [0,3]. \end{cases}$$

Portanto,

$$P(1 \leq X \leq 2) = \int_1^2 \frac{1}{3} dx = \frac{1}{3}.$$

O exemplo a seguir resolve este problema sobre a Uniforme, encapsulando as probabilidades.

Exemplo 2. *Se $X \sim U(0,3)$, qual a probabilidade de que o ponto escolhido esteja entre 1 e 2? Ainda neste exemplo, qual é a probabilidade de que o ponto escolhido esteja entre $1/2$ e $4/7$?*

A Tabela II mostra as probabilidades reais p e os intervalos encapsuladores Ip para estas probabilidades.

Tabela II
VALORES DE PROBABILIDADE REAL p E PROBABILIDADE INTERVALAR Ip DA UNIFORME

Exemplo 2	p	Ip
$P(1 \leq X \leq 2)$	0.33333333333333	[0.33333333333331, 0.33333333333337]
$P(1/2 \leq X \leq 4/7)$	0.02380952380952	[0.02380952380952, 0.02380952380954]

Verifica-se, pela Tabela II, que os intervalos encapsuladores para a Uniforme apresentam qualidade, ou seja, contêm a probabilidade real com erro no último dígito.

A Tabela III apresenta os Erros Absoluto e Relativo do Intervalo Encapsulador obtidos para a variável aleatória contínua Uniforme.

Tabela III
ERROS ABSOLUTO E RELATIVO DO INTERVALO ENCAPSULADOR PARA UNIFORME

Exemplo	**Erro Absoluto**	**Erro Relativo**
$P(1 \leq X \leq 2)$	$10^{-14} < 3 \times 10^{-14}$	$3 \times 10^{-14} \leq 9 \times 10^{-14}$
$P(1/2 \leq X \leq 4/7)$	$10^{-14} < 10^{-14}$	$2 \times 10^{-14} \leq 4 \times 10^{-14}$

Os erros obtidos de acordo com a Tabela III são muito próximos de zero, o que ratifica a qualidade dos intervalos encapsuladores para a Uniforme.

B. Exponencial

Exemplo 3. *Seja T a duração de vida, em horas, de um dispositivo eletrônico com densidade $f_T(t) = 0.01e^{-0.01t}$, $t > 0$. Qual a probabilidade de que o ponto escolhido:*
- **(i)** $T < 50$*;*
- **(ii)** *T esteja entre 20 e 50.*

Solução.
- **(i)** *Probabilidade real:* $P(T < 50)$
  Precisão dupla: 0.39346934028737.
- **(ii)** *Probabilidade real:* $P(20 \leq T \leq 50)$
  Precisão dupla: 0.21220009336535.

Exemplo 4. *O número de defeitos de um tecido segue uma lei de Poisson com média de um defeito a cada 500m. Qual a probabilidade que o intervalo entre dois defeitos consecutivos*
- **(a)** *seja no mínimo 1250m;*
- **(b)** *esteja entre 1000 e 1250m;*
- **(c)** *seja menor do que 1000m.*

Solução. *Se, X, o número de defeitos é Poisson, então a distância entre dois defeitos consecutivos é Exponencial com parâmetro $\alpha = \frac{1}{500} = 0.002$. Portanto,*
- **(a)** *Probabilidade real:* $P(X \geq 1250)$
  Precisão dupla: 0.08208499862390.
- **(b)** *Probabilidade real:* $P(1000 < X < 1250)$
  Precisão dupla: $0,05325143167565$.
- **(c)** *Probabilidade real:* $P(X < 1000)$
  Precisão dupla: 0.86466471676339.

A Tabela IV mostra as probabilidades reais p e os intervalos encapsuladores Ip para estas probabilidades dos Exemplos 3 e 4.

Tabela IV
VALORES DE PROBABILIDADE REAL p E PROBABILIDADE INTERVALAR Ip DA EXPONENCIAL

Exemplos 3 e 4	p	Ip
3) $P(T < 50)$	0.39346934028737	[0.39346934028736, 0.39346934028737]
$P(20 \leq T \leq 50)$	0.21220009336535	[0.21220009336534, 0.21220009336535]
4) $P(X \geq 1250)$	0.08208499862390	[0.08208499862389, 0.08208499862390]
$P(1000 < X < 1250)$	0,05325143167565	[0.05325143167564, 0.05325143167566]
$P(X < 1000)$	0.86466471676339	[0.86466471676338, 0.86466471676339]

A Tabela V apresenta os Erros Absoluto e Relativo do intervalo encapsulador obtidos para a variável aleatória contínua Exponencial.

Tabela V
ERROS ABSOLUTO E RELATIVO DO INTERVALO ENCAPSULADOR PARA EXPONENCIAL

Exemplos	**Erro Absoluto**	**Erro Relativo**
3) $P(T < 50)$	0	$10^{-14} \leq 10^{-14}$
$P(20 \leq T \leq 50)$	$0 < 10^{-14}$	$2 \times 10^{-14} \leq 2 \times 10^{-14}$
4) $P(X \geq 1250)$	0	$6 \times 10^{-14} \leq 6 \times 10^{-14}$
$P(1000 < X < 1250)$	$0 < 10^{-14}$	$0 \leq 1,9 \times 10^{-13}$
$P(X < 1000)$	0	$10^{-14} \leq 10^{-14}$

Os resultados dos erros contidos na Tabela V mostram a ocorrência de diferença apenas no último dígito, onde para todas as variáveis obteve-se intervalos solução com qualidade e ainda contendo a solução exata do problema com entradas reais.

C. Pareto

Exemplo 5. *Seja X uma variável aleatória com distribuição de Pareto com parâmetros α e c. Calculando-se a probabilidade real $P(1 < X < 2)$, em precisão dupla, nos casos abaixo tem-se:*

Solução.

(i) *Se* $c = 1$ *e* $\alpha = 0.25$ *então* $P(1 < X < 2) = 0.15910358474629$.

(ii) *Se* $c = 1$ *e* $\alpha = 0.5$ *então* $P(1 < X < 2) = 0.29289321881345$.

(iii) *Se* $c = 1$ *e* $\alpha = 0.75$ *então* $P(1 < X < 2) = 0.40539644249864$.

(iv) *Se* $c = 1$ *e* $\alpha = 1$ *então* $P(1 < X < 2) = 0.50000000000000$.

(v) *Se* $c = 1$ *e* $\alpha = 1.25$ *então* $P(1 < X < 2) = 0.57955179237314$.

(vi) *Se* $c = 1$ *e* $\alpha = 1.3$ *então* $P(1 < X < 2) = 0.59387380182188$.

A Tabela VI mostra as probabilidades reais p e os intervalos encapsuladores Ip para estas probabilidades.

Tabela VI
VALORES DE PROBABILIDADE REAL p E PROBABILIDADE INTERVALAR Ip DA PARETO

Exemplo 5	p	Ip
i)	0.15910358474629	[0.15910358474628, 0.15910358474629]
ii)	0.29289321881345	[0.29289321881345, 0.29289321881346]
iii)	0.40539644249864	[0.40539644249863, 0.40539644249864]
iv)	0.50000000000000	[0.5, 0.5]
v)	0.57955179237314	[0.5795517923731426, 0.5795517923731428]
vi)	0.59387380182188	[0.5938738018218823, 0.5938738018218824]

A Tabela VII apresenta os Erros Absoluto e Relativo do intervalo encapsulador obtidos para a variável aleatória contínua Pareto.

Tabela VII
ERROS ABSOLUTO E RELATIVO DO INTERVALO ENCAPSULADOR PARA PARETO

Exemplo 5-$P(1 < X < 2)$	**Erro Absoluto**	**Erro Relativo**
i)	0	$3 \times 10^{-14} \leq 3 \times 10^{-14}$
ii)	0	$2 \times 10^{-14} \leq 2 \times 10^{-14}$
iii)	0	$10^{-14} \leq 10^{-14}$
iv)	0	$0 \leq 0$
v)	0	$10^{-16} \leq 10^{-16}$
vi)	0	$0 \leq 0$

Os exemplos ilustram a computação científica intervalar confirmando que os resultados obtidos contemplam a solução real, com intervalos encapsuladores de qualidade com erros (absoluto e relativo) próximos a zero.

Os algoritmos desenvolvidos para calcular os erros absoluto e relativo dos intervalos encapsuladores das variávies aleatórias Uniforme, Exponencial e Pareto possuem complexidade constante, pois possuem operações aritméticas intervalares básicas e, ainda, consideram os extremos dos intervalos encapsuladores para o ponto médio e diâmetro destes intervalos, o que diminui o número de operações fundamentais a serem executadas.

VI. CONCLUSÃO

No estudo das variáveis aleatórias sobre o conjunto dos números reais, $\mathbb{R}$, um dos problemas é o cálculo de probabilidades, visto que é necessário resolver uma integral definida da função densidade que, na maioria das vezes, não possui primitiva explícita ou cuja primitiva não é simples de se obter. Embora integrais de funções densidade de probabilidade como a Uniforme, a Exponencial e a de Pareto, sejam resolvidas analiticamente, seu valor numérico no computador é dado por aproximação, e portanto afetado por erros de arredondamento ou truncamento. Outras funções densidade como a Normal ou Gama, por exemplo, não possuem primitivas na forma analítica, sendo necessário o uso de integração numérica onde erros de arredondamentos e truncamentos são propagados devido às operações aritméticas realizadas no computador.

Para as variáveis Uniforme e Pareto o intervalo encapsulador foi definido a partir da função densidade e para a variável Exponencial o método proposto fundamenta-se na aplicação do método de Simpson Intervalar.

O objetivo deste trabalho foi mostrar a importância e justificativa de se utilizar a matemática intervalar no cálculo de intervalos encapsuladores para as variáveis aleatórias Uniforme, Exponencial e Pareto. A partir do desenvolvimento dos métodos e algoritmos para encontrar intervalos, baseados na matemática intervalar e na aritmética de exatidão máxima que encapsulassem probabilidades reais para as variáveis aleatórias realizou-se uma análise de complexidade de cada algoritmo para justificar seu uso.

Com o sistema de ponto flutuante F(10, 14, -10, 10) (ou com quatorze casas decimais) verifica-se através das medidas de erros (absoluto e relativo) que todas as probabilidade reais estão contidas nos intervalos encapsuladores.

Ao analisar a complexidade de um problema pode-se informar qual a classe de complexidade que o mesmo pertence. Conhecendo a classe em que o problema pertence, os projetistas de algoritmos, e até mesmo profissionais da área de estatística, poderiam se concentrar naqueles problemas para os quais existem algoritmos razoáveis, ou seja, que possam ser resolvidos através do computador em tempo polinomial de processamento.

Segundo Toscani [12], identificar a tratabilidade e a intratabilidade dos problemas, mesmo aqueles que possuem algoritmos imediatos (algoritmos de fácil construção), é de extrema importância para os projetistas de algoritmos, pois conhecendo as classes de complexidade em que os problemas pertencem, os projetistas poderiam ter uma medida real quanto às soluções disponíveis e a expectativa de melhorar esses resultados.

Nesterov [36] salienta que, do ponto de vista técnico, muitos problemas práticos e importantes da computação intervalar são conhecidos como computacionalmente intratáveis, ou conhecidos como computacinalmente razoáveis. Existem poucos problemas dos quais não se sabe se são razoáveis ou não. Assim, pode-ser ter a impressão de que esta área de pesquisa está quase finalizada, o que é uma falta impressão. Acredita-se que existem muitos problemas da computação intervalar em aberto e que as pesquisas nesta área mostram-se recentes e atuais.

Como trabalhos futuros, pretende-se realizar uma análise mais abrangente em relação a complexidade computacional. Deseja-se analisar o custo em termos de tempo de máquina para os algoritmos implementados para cada variável aleatória e em diferentes ambientes bem como o custo envolvido para o controle dos erros de arredondamento e truncamento. Nesta análise deseja-se verificar a qualidade dos intervalos encapsuladores em quatro ambientes intervalares: XSC (eXtended for Scientific Computation), Maple Intervalar, IntLab e IntPy (utilizado no presente trabalho). Tais ambientes possibilitam a programação utilizando operações definidas na matemática intervalar [1] e, através da aplicação dos cálculos dos erros Absoluto e Relativos nos intervalos encapsuladores, verifica-se a qualidade do intervalo e define-se qual dos ambientes retorna um melhor resultado.

Agradecimentos

Os autores agradecem à CAPES pelo suporte financeiro na realização do presente trabalho.

Referências

[1] R. E. Moore, *Interval Analysis*. Englewood Cliffs: Prentice Hall, 1966.

[2] ——, *Methods and Applications of Interval Analysis*, 2nd ed. SIAM, 1979.

[3] R. Moore, M. Kearfott, and J. Cloud, *Introduction to Interval Analysis*. Philadelphia: SIAM, 2009.

[4] H. Ratschek and R. Rokne, *New Computer Methods for Global Optimization*. Ellis Horwood, 1988.

[5] R. B. Kearfott, "Interval computations: Introduction, uses, and resources," *Euromath Bulletin*, vol. 2, pp. 95–112, 1996.

[6] U. W. Kulisch. (2008, apr) Complete interval arithmetic and its implementation on the computer. [Online]. Available: http://www.math.kit.edu/iwrmm/seite/preprints.

[7] U. Kulisch and L. Miranker, *Computer Arithmetic in Theory and Practice*, 1st ed. Academic Press, 1981.

[8] W. Feller, *An Introduction to Probability and Its Applications*, 3rd ed. John Wiley & Sons, 1968.

[9] B. R. James, *Probabilidade: um curso em nível intermediário*, 3rd ed. IMPA, 2006.

[10] P. L. Meyer, *Probabilidade Aplicações à Estatística*, 2nd ed. LTC, 1983.

[11] Python Software Foundation. (2012, feb). [Online]. Available: http://www.python.org/.

[12] L. Toscani and P. Veloso, *Complexidade de Algoritmos: análise, projetos e métodos*. Sagra-Luzzato, 2001.

[13] F. R. G. Varjão, "IntPy: Computação científica auto validável em Python," Master's thesis, Universidade Federal de Pernambuco, Recife, 2011.

[14] H. N. O. Caprani, K. Madsen, "Introduction to interval analysis," *IMM - Informatics and Mathematical Modelling*, 2002.

[15] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*. Dordrecht: Kluwer, 1998.

[16] P. Oliveira, T. Diverio, and D. Claudio, *Fundamentos de Matemática Intervalar*. Sagra-Luzzato, 1997.

[17] S. Ferson, L. Ginzburg, and V. Kreinovich, "Absolute bounds on the mean of sum, product, etc.: A probabilistic extension of interval arithmetic," in *SIAM WORKSHOP ON VALIDATED COMPUTING*, Toronto, 2002.

[18] R. E. Moore, *On computing the range of a rational function of variables over a bounded region*. Computing, 1976, vol. 16.

[19] A. Gaganov, "Computational complexity of the range of the polynomial in several variables," *Cybernetics*, vol. 11, pp. 418–421, 1985.

[20] Y. Matiyasevich, "Enumerable sets are diophantine," *Soviet Math. Doklady*, vol. 11, pp. 354–357, 1970.

[21] S. F. et al., "Exact bounds on sample variance of interval data," *IAM WORKSHOP ON VALIDATED COMPUTING*, 2002.

[22] ——, *Computing variance for interval data is NP-Hard*, 2nd ed. ACM SIGACT News, 2002.

[23] B. Wu, H. Nguyen, and V. Kreinovich, "Real-time algorithms for statistical analysis of interval data," in *INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY, INTECH*, Thailand, 2003, pp. 483–490.

[24] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpre, and M. Aviles, *Exact bounds on finite populations of interval data*. Reliable Computing, 2004, vol. 3.

[25] B. Traylor and V. Kreinovich, *A bright side of NP-harness of interval computations: interval heuristics applied to NP-problems*. Reliable Computing, 1995, vol. 3.

[26] A. B. Loreto, "Análise da complexidade computacional de problemas de estatística descritiva com entradas intervalares," Ph.D. dissertation, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

[27] G. E. Forsythe, "Pitfalls in computation, or a why a math book isn't enough," *Amer. Math. Monthly*, pp. 931–955, 1977.

[28] D. Goldberg, "What every computer scientist should know about floating-point arithmetic," *ACM Computing Surveys*, vol. 23, no. 1, pp. 5–48, 1991.

[29] ANSI/IEEE STD 754, "Ieee standard for binary floating - point arithmetic," *ACM GIGPLAN*, vol. 22, pp. 5–48, 1987.

[30] M. A. G. Ruggiero and V. L. da Rocha Lopes, *Cálculo numérico: aspectos teóricos e computacionais*, 2nd ed. São Paulo: Pearson Makron Books, 1996.

[31] W. Rudin, *Princípios de Análise Matemática*. Livro Técnico S.A., 1971.

[32] M. G. Santos, "Probabilidades autovalidáveis para as variáveis aleatórias exponencial, normal e uniforme." Ph.D. dissertation, Universidade Federal de Pernambuco, Recife, 2010.

[33] V. Kreinovich, "Probabilities. intervals, what next? optimization problems related to extension of interval computations to situation with partial information about probabilities," *Global Optimization*, vol. 29, no. 3, pp. 265–280, 2003.

[34] T. Sunaga, *Theory of an Interval Algebra and its Applications to Numerical Analysis*. RAAG Memoirs, 1958.

[35] M. A. Campos, E. L. Silva, D. C. Pedrosa, J. A. Loureiro, J. L. C. SIlva, and C. A. Ferraz, "Netbook: uma ferramenta para avaliação de desempenho em redes de comunicação," in *Proc. Salão de Ferramentas, Simpósio Brasileiro de Redes de Computadores*, 2004, pp. 967–974.

[36] S. Nesterov and V. Kreinovich, "The worse, the better: a survey of paradoxical computational complexity of interval computations," in *Proc. Abstract of the II Workshop on Computer Arithmetic, Interval and Symbolic Computations - WAI'96*, Recife, Pernambuco, Brasil, 1996, pp. 61A–63A.