EXTENSÃO INTERVALAR DA DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA GAMA

ALICE FONSECA FINGER[1]; VINICIUS SIGNORI FURLAN[2]; MAURICIO DORNELES CALDEIRA BALBONI[2]; ALINE BRUM LORETO[3]

[1]*Universidade Federal de Pelotas – affinger@inf.ufpel.edu.br*
[2]*Universidade Federal de Pelotas – vsfurlan@inf.ufpel.edu.br, baalbis@gmail.com*
[3]*Universidade Federal de Pelotas – aline.loreto@inf.ufpel.edu.br*

1. INTRODUÇÃO

A aritmética intervalar utiliza intervalos reais para representar valores infinitos, valores desconhecidos ou para representar valores contínuos que podem ser conhecidos ou não. Os intervalos servem para representar dados inexatos, aproximações e erros de truncamento de procedimentos.

Na matemática intervalar, o valor real *x* é aproximado por um intervalo **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 (RATSCHEK; ROKNE, 1988). Os cálculos reais são substituídos por cálculos que utilizam a aritmética intervalar (MOORE, 1966).

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 (FELLER, 1968) (JAMES, 2006) (MEYER, 1983). 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. Na resolução de funções densidade de probabilidade, seu valor numérico é obtido por aproximação e, portanto, afetado por erros de arredondamento ou truncamento, além disso, é necessário utilizar integração numérica.

Considerando que métodos numéricos devem ser usados para o cálculo de integrais, a ideia 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*) (KLATTE, 1993).

Para utilizar a aritmética intervalar nas distribuições de probabilidade das variáveis aleatórias contínuas, primeiramente foram definidas as distribuições com intervalos, onde em SANTOS (SANTOS, 2010) definiu-se a extensão intervalar das distribuições Uniforme, Normal e Exponencial e em FINGER et al. (2012) definiu-se extensão intervalar para Pareto.

A fim de completar a definição intervalar para as distribuições de probabilidade das variáveis aleatórias contínuas, o presente trabalho tem como objetivo apresentar a extensão intervalar para a distribuição Gama, bem como sua validação através da comparação dos resultados obtidos da forma real com a intervalar.

2. METODOLOGIA

A aritmética intervalar é baseada no uso de intervalos fechados [x_1, x_2] de números reais como elementos básicos, e sua ideia, do ponto de vista computacional, é: dada uma função *f(x)* de variável real x pertencente a um intervalo X = [x_1, x_2], onde $x_1, x_2 \in \mathbb{R}$, o domínio de *f* é dado por:

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

onde este em geral não é representado exatamente, mas é sempre possível determinar um intervalo **y** = [y_1, y_2] tal que *f(x)* ⊆ y, isto é $y_1 \le f(x) \le y_2$. 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:

$$f(x) \subseteq F(x) = y$$

Esta função F, chamada extensão intervalar de *f*, deve ser aquela que se afasta o mínimo possível de *f(x)* (KREINOVICH et al., 1998).

A função $\Gamma = \Gamma(x)$ dá origem a distribuição, a qual é utilizada nos fenômenos limitados, como, por exemplo, os intervalos de tempo de espera numa fila de banco ou para analisar o tempo de permanência de pacientes num hospital (VILCHES, 2013).

A função de densidade de probabilidade Gama, de parâmetros λ > 0 e $\nu \in \mathbb{R}$, é definida por:

$$f(x) = \begin{cases} \dfrac{\lambda^{\nu}}{\Gamma(\nu)} e^{-\lambda x} \cdot x^{\nu - 1} & \text{se } x > 0 \\ 0 & \text{outro caso.} \end{cases}$$

Utilizando a definição da função Gama, obtém-se que:

$$\int_0^{+\infty} e^{-\lambda x} \cdot x^{\nu-1} dx = \frac{\Gamma(\nu)}{\lambda^{\nu}}.$$

Definição Gama Intervalar: Seja X uma Gama com parâmetros λ e ν**.** Para computar o intervalo encapsulador da probabilidade intervalar para esta variável aleatória, define-se a função densidade com a extensão intervalar:

$$F_G(X) = \frac{\Gamma(\nu)}{\lambda^2} \cdot \left(e^{-\lambda X} \cdot X^{\nu-1}\right),$$

onde $X = [x_1, x_2]$ e $x_1 \le x \le x_2$.

3. RESULTADOS E DISCUSSÃO

Exemplos numéricos foram computados a fim de comparar o resultado da aplicação da distribuição Gama na forma real e intervalar.

Para obter os valores na forma real, utilizou-se o software matemático Maple (MAPLE). Já os intervalos solução foram computados através da implementação da distribuição Gama no ambiente de programação intervalar IntPy (INTPY).

Abaixo são apresentados dois exemplos de aplicação da distribuição de probabilidade da variável Gama.

Exemplo 1: Se o tempo de sobrevivência no mercado, em anos, de um certo tipo de microempresa segue a distribuição Gama para λ = 0.81 e ν = 7.81, determine qual a probabilidade de que a sobrevivência seja menor que 10 anos.

Exemplo 2: Se o tempo utilizando um computador com *time-sharing* segue uma distribuição Gama com média de 20 minutos e variância de 80 minutos, pede-se: Qual é a probabilidade de um usuário utilizar o computador entre 20 e 40 minutos?

Tabela 1: Soluções dos exemplos, real e intervalar, da distribuição de probabilidade Gama.

Exemplos	Maple	IntPy
1	0.587755193314091	[0.587755193314091, 0.587755193314092]
2	$3.220751040236955 \times 10^{-46}$	[$3.220751040236954 \times 10^{-46}$, $3.220751040236955 \times 10^{-46}$]

A partir dos resultados apresentados na Tabela 1, pode-se verificar que a extensão intervalar desenvolvida para a distribuição Gama retorna um intervalo que contém a resposta real. Para analisar a qualidade do intervalo encapsulador, utiliza-se o comprimento (ou diâmetro) (RATSCHEK; ROKNE, 1998) deste intervalo como medida para verificar a qualidade da aproximação, o qual é calculado através da fórmula: $w(x) = \overline{x} - \underline{x}$.

Na Tabela 2 são apresentados os resultados obtidos para o cálculo do diâmetro do intervalo encapsulador de cada exemplo.

Tabela 2: Cálculo do comprimento (ou diâmetro) do intervalo solução resultante dos exemplos.

Exemplos	Diâmetro do intervalo
1	0.000000000000001
2	$0.000000000000001 \times 10^{-46}$

De acordo com os resultados obtidos na análise da qualidade do intervalo, é possível afirmar que o intervalo encapsulador retornado com a definição intervalar para Gama é de qualidade, uma vez que ele contém o valor real e os limites, inferior e superior, diferem apenas no 15º dígito, considerando um sistema de Ponto Flutuante (10, 14, -10, 10).

4. CONCLUSÕES

Quando se trabalha com computação numérica, um dos fatores de maior importância é a exatidão da resposta desses cálculos. O que sempre se procura são resultados cada vez mais exatos e com um menor erro possível contido neles.

No estudo das variáveis aleatórias um dos problemas é o cálculo de probabilidades devido à necessidade de resolver uma integral definida da função densidade. A matemática intervalar tem como objetivo realizar um controle automático de erro dos cálculos, retornando respostas com maior exatidão possível.

Neste contexto, o trabalho apresenta a definição intervalar para a distribuição de probabilidade da variável Gama, bem como sua validação através de exemplos numéricos. Os resultados comprovam que a definição intervalar apresentada retorna como solução um intervalo que encapsula o valor real, resultado importante, o qual justifica o uso da matemática intervalar na resolução das distribuições de probabilidade.

5. REFERÊNCIAS BIBLIOGRÁFICAS

FELLER, W. **An Introduction to Probability and Its Applications**, 3rd ed., John Wiley & Sons, 1968.

FINGER, A. F.; LORETO, A. B.; CAMPOS, M. A.; VARJÃO, F. R. G.; SANTOS, M. G. A Complexidade Computacional dos Problemas de Computar Intervalos Encapsuladores para as Variáveis Aleatórias Uniforme, Exponencial e Pareto. In: **XXXVIII CONFERENCIA LATINOAMERICANA EN INFORMÁTICA.** Medellín, 2012, **Anais...**

INTPY. Interval Arithmetic package. Online. Disponível em: https://pypi.python.org/pypi/IntPy/0.1.3.

JAMES, B. R. **Probabilidade: um curso em nível intermediário**, 3rd ed., IMPA, 2006.

KLATTE, R.; KULISCH, U.; WIETHOFF, A.; LAWO, C.; RAUCH, M. **C-XSC - A C++ Class Library for Extended Scientific Computing**. Springer-Verlag, Heidelberg, 1993.

KREINOVICH, V.; LAKEYEV, A.; ROHN, J.; KAHL, P. **Computational Complexity and Feasibility of Data Processing and Interval Computations**. Springer, 1998.

MAPLE. Power Interval Arithmetic. Online. Disponível em: www.math.uniwuppertal.de/wrswt/software/intpakX.

MEYER, P. L. **Probabilidade: Aplicações à Estatística**, 2nd ed., LTC, 1983.

MOORE, R. E. **Interval Analysis**. Englewood Cliffs: Prentice Hall, 1966.

RATSCHEK, H.; ROKNE, R. **New Computer Methods for Global Optimization**. Ellis Horwood, 1988.

SANTOS, M. G. **Probabilidades autovalidáveis para as variáveis aleatórias exponencial, normal e uniforme.** 2010. Tese (Doutorado em Matemática Computacional) – Programa de Pós-Graduação em Matemática Computacional, Universidade Federal de Pernambuco.

VILCHES, M. A. Cálculo para economia e administração. Acessado em 15 mai. 2013. Online. Disponível em: http://magnum.ime.uerj.br/~calculo/ecomat.html.