

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Biblioteca Estatística Descritiva Intervalar utilizando Python

Lucas Mendes Tortelli¹

Programa de Pós-Graduação em Computação, CDTEc, UFPel, Pelotas, RS

Dirceu A. Maraschin Jr.²

Programa de Pós-Graduação em Computação, CDTEc, UFPel, Pelotas, RS

Alice Fonseca Finger³

Programa de Pós-Graduação em Computação, CDTEc, UFPel, Pelotas, RS

Aline Brum Loreto⁴

Campus de Cachoeira do Sul, UFSM, Cachoeira do Sul, RS

1 Introdução

A computação científica é largamente utilizada quando há necessidade de exatidão e confiabilidade. Ao representar dados numéricos no sistema de ponto flutuante, estes são aproximados para um subconjunto finito dos números reais. Através desta aproximação ocorrem erros pela impossibilidade da máquina representar estes valores reais.

A matemática intervalar, proposta por Moore [3], surge para automatizar o cálculo do erro computacional científico com limites confiáveis. Nesta forma de representação numérica os valores pontuais são representados por um intervalo X que os contenha. Todo erro originado no processamento dos valores é tratado pela aritmética intervalar.

Este trabalho visa a criação de uma Biblioteca Estatística Descritiva Intervalar na linguagem Python, fornecendo para a literatura formas de garantir resultados exatos e controlar o erro gerado em dados estatísticos.

2 Metodologia e Resultados

A Estatística Descritiva, que tem como objetivo a descrição e organização dos dados de uma população ou amostra, necessita de grande precisão durante todas as etapas que consistem seu processamento.

Devido a exigência de exatidão, como também a garantia da confiança do resultado em cada etapa deste processo, se faz necessário o uso de técnicas que controlem o erro numérico gerado. A fim de alcançar tais objetivos utiliza-se a Aritmética Intervalar [3].

¹lmtortelli@inf.ufpel.edu.br

²dirceumaraschin@gmail.com

³alicefinger@unipampa.edu.br

⁴aline.loreto@gmail

Para a realização da implementação de dados intervalares, o ambiente de desenvolvimento deve conter o tipo de dado intervalar com as definições das operações e funções básicas desta aritmética. O IntPy consiste em uma biblioteca que importa o tipo intervalar dentro da linguagem de programação Python [2]. Sua forma de operação garante redução dos erros absolutos e relativos em comparação com outros ambientes intervalares.

Os indicadores estatísticos descritivos intervalares foram implementados, conforme definido por Loreto [1]. A Figura 1 apresenta a execução de cada indicador descritivo definido na biblioteca, como também os resultados obtidos em cada etapa do processamento. No passo (a) ocorre a importação da biblioteca, na etapa (b) são declarados dois conjuntos de dados (data1 e data2) e a respectiva representação intervalar com precisão = 0,01. No passo (c) ocorre a chamada da rotina de cada indicador estatístico sobre os dados intervalares no passo (b). A etapa (d) apresenta os resultados obtidos no passo (c).

```
(a)>>> from intStatistics import *
(b)>>> data1 = [5.0,12.0,7.0,47.0,7.0,79.0,10.0]
    >>> data2 = [16.0,68.0,21.0,27.0,25.0,97.0,27.0]
    >>> data1 = initInterval(data1,0.01)
    >>> data2 = initInterval(data2,0.01)
(c)>>> md = iAverage(data1)
    >>> med = iMedian(data1)
    >>> at = iRange(data1)
    >>> va = iVariance(data1)
    >>> dp = iSDeviation(data1)
    >>> co = icoVariance(data1,data2)
    >>> cc = icoefCorrelation(data1,data2)
    >>> cv = icoefVariance(data1)
(d)>>> md
    [23.84714285714285, 23.867142857142866]
    >>> med
    [9.989999999999998, 10.010000000000002]
    >>> at
    [74.0, 74.0]
    >>> va
    [689.5138693877544, 691.3032571428579]
    >>> dp
    [26.258596104661695, 26.29264644616167]
    >>> co
    [530.0893795918361, 531.6665224489802]
    >>> cc
    [0.7166765008646055, 0.720623030645976]
    >>> cv
    [1.1001985558905356, 1.1025491231242543]
```

Figura 1: Exemplo de execução da biblioteca

Salienta-se que todos os resultados reais foram contemplados nos resultados intervalares obtidos pela biblioteca.

3 Conclusões

A biblioteca desenvolvida neste trabalho visa fornecer uma ferramenta de fácil acesso e manipulação para processamento de dados estatísticos, além de garantir resultados exatos e controlar o erro gerado. Essa encontra-se disponível no repositório do Python.

Referências

- [1] A.B. Loreto, Análise da Complexidade Computacional de Problemas de Estatística Descritiva com Entradas Intervalares, Tese de Doutorado, UFRGS, 2006.
- [2] F.R.G Varjão, Computação científica auto validável em Python. Dissertação de Mestrado, UFPE, 2011.
- [3] R.E. Moore, Interval Analysis, Englewood Cliffs, 1966.