

CRIAÇÃO DE FRAMEWORK PARA A REALIZAÇÃO DE CÁLCULOS DE MEDIDAS DE DISPERSÃO E AGRUPAMENTO DE DADOS EM INTERVALOS DE CLASSES

Rodrigo Haering¹, Rodrigo Curvello²

¹Instituto Federal Catarinense

rodrigohaering333@gmail.com.br, rodrigo.curvello@ifc.edu.br

Abstract. *This meta-article describes the use of Object-Oriented Programming in the construction of a Framework that has the goal to perform dispersion measures calculation, as well, the grouping of data in a class range and graphics exhibition. Also, the concept of Framework will be presented, as well as the meaning of dispersion measures and its relevance to the statistics area.*

Key-words: *OOP; Dispersion measures; Statistics.*

Resumo. *Este meta-artigo descreve a utilização de programação orientada a objetos na construção de um framework que tem como objetivo a realização de cálculos de medidas de dispersão, bem como, o agrupamento de dados em intervalos de classes e exibição de gráficos. Além disso, será apresentado o conceito de Framework, bem como, o significado de medidas de dispersão e sua relevância para a área da estatística.*

Palavras-chave: *Framework; Medidas de dispersão; Intervalos de Classes.*

1. Introdução

O presente artigo foi desenvolvido com o objetivo de explicar e exemplificar em forma de código o desenvolvimento de um framework responsável pela realização de cálculos de medidas de dispersão, bem como, o agrupamento de dados em uma tabela de intervalos de classe e a criação de três tipos de gráficos com base nesses dados. O Framework desenvolvido permite ao usuário que irá utilizá-lo realizar o cálculo das principais medidas de dispersão da área da estatística com base nos dados quantitativos informados, bem como, gerar uma tabela com os intervalos de classes e três tipos de gráficos com os respectivos dados da tabela. As medidas de dispersão que serão trabalhadas podem ser descritas em: amplitude, média, moda, mediana, variância, desvio padrão e coeficiente de variação. Essas medidas fornecem valores que mostram a dispersão dos dados em relação à média.

2. Framework

Um framework normalmente é um conjunto de bibliotecas para conseguir executar uma operação maior e, costuma dar consistência a esse conjunto. Um Frameworks pode ser entendido como uma plataforma de desenvolvimento, no qual

possui lacunas que devem ser preenchidas pelo programador para funcionar conforme sua necessidade. Pode ser visto como um esqueleto de uma aplicação. Frequentemente são confundidos, de forma certa ou errada, como um conjunto de classes.

3. Medidas de Dispersão

No estudo da estatística, existem algumas estratégias que ajudam a verificar se os valores apresentados em um conjunto de dados estão dispersos ou não, e o quão distantes um do outro eles podem estar. É aí que entram as ferramentas classificadas como medidas de dispersão. As principais são denominadas em: amplitude total, variância, desvio padrão e coeficiente de variação.

A amplitude total tem o intuito de mostrar a dispersão entre os dados, quando maior for a amplitude, mais heterogêneo é o grupo de dados. Ela é obtida através da diferença entre o maior e o menor valor dos dados observados.

A variância mostra em média, o quão distante cada valor de um conjunto de dados está em relação à média. Quanto menor é a variância, mais próximos os valores estão da média; mas quanto maior é, mais os valores estão distantes da média (RIBEIRO, 2018).

O desvio padrão expressa o grau de dispersão de um conjunto de dados. Ou seja, o desvio padrão indica o quanto um conjunto de dados é uniforme. Quanto mais próximo de 0 for o desvio padrão, os dados são mais homogêneos e tendem a estar mais próximos da média. Se o desvio padrão for alto, isso indica que os dados estão espalhados por uma ampla gama de valores (GOUVEIA, 2018).

O coeficiente de variação, fornecido em porcentagem, fornece a variação dos dados obtidos em relação à média. Geralmente essa medida é utilizada quando se quer comparar a variabilidade de dois conjuntos de dados. Quanto menor for o seu valor, mais homogêneos serão os dados. O coeficiente de variação é considerado baixo quando for menor que 15%, médio quando estiver entre 15 e 30% e alto quando for maior que isso (RIGONATTO, 2018). Embora o desvio padrão também fornecer o grau de variabilidade dos dados, para calcular a variabilidade entre dois conjuntos de dados deve-se comparar o coeficiente de variação de cada conjunto de dados em vez do desvio padrão, pois, conforme lido no artigo publicado pelo site Portal Action (2014), um desvio padrão pode ser considerado grande ou pequeno dependendo da ordem de grandeza da variável. Sendo assim, uma maneira de se expressar a variabilidade dos dados tirando a influência da ordem de grandeza da variável é através do coeficiente de variação.

4. Intervalos de classe e Gráficos.

Distribuição de frequência sem intervalos de classe é a simples condensação dos dados conforme as repetições de seus valores. Para um ROL de tamanho razoável esta distribuição de frequência é inconveniente, já que exige muito espaço.

É aí que entram os intervalos de classe, pois quando o tamanho da amostra é elevado, é mais racional efetuar o agrupamento dos valores em vários intervalos em vez de trabalhar com os dados em ROL.

A figura a seguir mostra um exemplo de uma tabela com intervalos de classe.

Peso de pessoas	Qtd de pessoas
41 ----- 45	7
45 ----- 49	3
49 ----- 53	4
53 ----- 57	1
57 ----- 61	5
Total	20

Figura 1: Exemplo de tabela com intervalos de classes.

4.1. Gráficos

A figura a seguir mostra o gráfico gerado ao pegar os dados da coluna “Qtd de pessoas” da tabela mostrada na Figura 1.

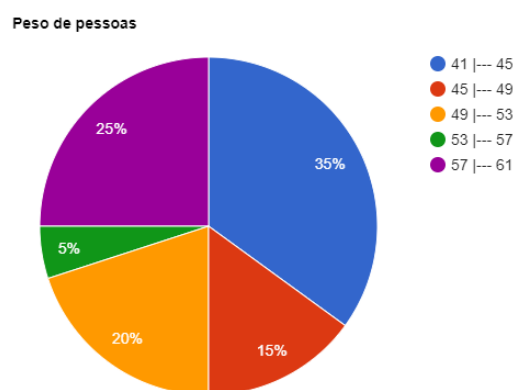


Figura 2: Gráfico de pizza.

5. Desenvolvimento

Neste tópico, será apresentada e exemplificada em forma de código a classe desenvolvida para ser o framework.

```

60.  /**
61.   * @param <T> - subclasse da classe calculo
62.   * @param <E> - tipo do número. (inteiro, float, double...)
63.   */
64.  public class Framework<T extends Calculo, E extends Number> {
65.
66.      protected CalculoCustom calculo;
67.      protected TabelaFrequencia tabela;
68.      protected HTMLViewer html;
69.      protected CalculoBO calcBO;
70.      private List<Double> dados;
71.
72.      public Framework(CalculoCustom
73.  clazz) throws InstantiationException, IllegalAccessException, ClassNotFoundException {
74.          this.calculo = clazz;
75.          this.initAttributes();
76.      }

```

```
17. public void realizarCalculo() {
18.     calculo.realizarCalculo();
19. }
20.
21. public void setDados(List<E> lista) {
22.     for (E l : lista) {
23.         this.dados.add(Double.parseDouble(String.valueOf(l)));
24.     }
25.     calculo.setDados(dados);
26. }
27.
28. public void importarCalculo() {
29.     JFileChooser fc = new JFileChooser();
30.     fc.setCurrentDirectory(new File("ArquivosExportados"));
31.     int returnVal = fc.showOpenDialog(fc);
32.     File file = null;
33.     if (returnVal == JFileChooser.APPROVE_OPTION) {
34.         file = fc.getSelectedFile();
35.     }
36.     if (fc.getSelectedFile() != null) {
37.         if (file.getAbsolutePath().endsWith(".xml")) {
38.             calcBO.setPersistencia("xml");
39.         } else if (file.getAbsolutePath().endsWith(".json")) {
40.             calcBO.setPersistencia("json");
41.         }
42.         calculo = calcBO.importar(file.getAbsolutePath());
43.         calculo.realizarCalculo();
44.     }
45. }
46.
47. public DefaultTableModel gerarTabelaDeFreq() {
48.     this.tabela = new TabelaFrequencia(dados);
49.     return this.tabela.dtm;
50. }
51.
52. public JPanel gerarGrafico(String htmlFileName) {
53.     try {
54.         return html.generateHTML(this.tabela.grafico.drawChart(tabela.dtm)
55. , "teste");
56.     } catch (IOException ex) {
57.         return null;
58.     }
59. }
```

Figura 3: Classe do framework

Na figura 3 apresentada acima pode-se ver o código fonte da classe framework. A classe contém 4 atributos: calculo, tabela, html, calcBO e dados. O atributo calculo é a classe cálculo, que contém os métodos e atributos necessários para realizar o calculo de medidas de dispersão. O atributo tabela é o objeto responsável por gerar a tabela de intervalo de classes e os gráficos. O atributo html será usado para mostrar o gráfico gerado pela classe TabelaFrequencia. E por fim, o atributo dados que é uma lista de double, essa lista serve como uma lista auxiliar para converter a lista passada pelo usuário em uma lista de double pois ele possui a liberdade de trabalhar com outros tipos de números (inteiro, double ou float).

6. Considerações finais

//Falta conclusão.

6. Referências

GOUVEIA, Rosimar. **Desvio Padrão**. 2018. Disponível em:
<<https://www.todamateria.com.br/desvio-padrao/>>. Acesso em: 09 dez. 2018.

PORTAL ACTION. **MEDIDAS DE DISPERSÃO**. 2014. Disponível em:
<<http://www.portalaction.com.br/estatistica-basica/22-medidas-de-dispersao>>. Acesso em: 09 dez. 2018.

RIBEIRO, Amanda Gonçalves. **Medidas de dispersão: variância e desvio padrão**. Disponível em: <<https://brasilescola.uol.com.br/matematica/medidas-dispersao-variancia-desvio-padrao.htm>>. Acesso em: 09 dez. 2018.

RIGONATTO, Marcelo. **Coeficiente de variação**. Disponível em:
<<https://mundoeducacao.bol.uol.com.br/matematica/coeficiente-variacao.htm>>. Acesso em: 09 dez. 2018.