



ESTIMATIVA DAS MEDIDAS DESCRITIVAS DE VARIÁVEIS QUANTITATIVAS COM PROGRAMAÇÃO WEB PHP

Estimate of Descriptive Measures of Quantitative Variables with PHP Web Programming

DA SILVA, Higor de Barros¹; BARNASKI, Maria²; DORNELES, Daniel³;
DUTRA, Vinicius⁴; SATHES, Lidiandro⁵; FIGUEIRÓ, Michele F.⁶

Resumo: Este artigo descreve a elaboração de uma calculadora estatística escrita na linguagem de programação web PHP. Nela, é possível estimar as medidas de posição e de dispersão de variáveis quantitativas estudadas na Estatística Descritiva da disciplina de Probabilidade e Estatística, quando se faz necessário o resumo e a organização de um banco de dados. A tabulação dos dados se dará diretamente na calculadora ou via arquivo, e os resultados das análises descritivas poderão ser salvos em um arquivo para facilitar suas interpretações.

Palavras-chave: Estatística Descritiva. PHP. Medidas Descritivas. Variáveis Quantitativas.

Abstract: This article describes the construction of a statistical calculator written in the PHP web programming language. In it, it is possible to estimate the position and dispersion measures of quantitative variables were studied in Probability and Statistics subject, when it is necessary the summary and the organization of a database. The tabulation of data will occur directly on the calculator or via file, and the results of descriptive analysis will be able to save in a file in order to facilitate its interpretation.

Keywords: Descriptive Statistics. PHP. Descriptive Measures. Quantitative Variables.

Introdução

A estatística descritiva é a etapa inicial de qualquer trabalho de análise estatística de uma base de dados, cuja a finalidade é obter algum conhecimento ou informação sobre estes dados coletados. Através da análise descritiva, é possível apresentar, descrever, organizar, simplificar e resumir os dados (Bussab & Morettin, 2014; Barbetta, 2014). Antes da aplicação das técnicas da análise descritiva, é necessário classificar as variáveis (ou características de interesse), que estão sendo investigadas na pesquisa, como qualitativas ou quantitativas. Para

¹ Acadêmico do curso de graduação em Ciência da Computação, UNICRUZ .E-mail: higor.jpc2012@gmail.com.

² Acadêmica do curso de graduação em Ciência da Computação, UNICRUZ.E-mail: mariabarnaski@gmail.com.

³ Acadêmico do curso de graduação em Ciência da Computação, UNICRUZ. E-mail: danielm.dorneles@gmail.com.

⁴ Acadêmico do curso de graduação em Ciência da Computação, UNICRUZ.E-mail: viniciusdutra10@hotmail.com.

⁵ Acadêmico do curso de graduação em Ciência da Computação, UNICRUZ.E-mail: lidiandro15@hotmail.com.

⁶ Doutora em Física, professora do curso de graduação em Ciência da Computação e Coordenadora do Núcleo de Estatística Aplicada (NEA), UNICRUZ e ENGETEC. E-mail: mfigueiro@unicruz.edu.br.



as variáveis quantitativas, podem ser determinadas as medidas (ou estatísticas) descritivas que podem ser classificadas em dois tipos: de posição e de dispersão (Goos & Meintrup, 2015; Devore & Berk, 2012). De acordo com Estatística Descritiva, as medidas de posição são dadas pelo mínimo, máximo, média e percentis, enquanto que as medidas de dispersão são amplitude, intervalo interquartil, variância, desvio padrão e coeficiente de variação. O cálculo destas estatísticas será válido somente para amostras de populações finita ou infinita.

O PHP é uma linguagem interpretada por um servidor, utilizada quase que principalmente em páginas *web*, que se diferencia de outras linguagens, tais como as linguagens C ou Lua, que são linguagens compiladas. O código implementado nessa linguagem será executado no próprio servidor *web*, gerando então um HTML, que será enviado para o servidor.

A linguagem foi criada em 1995, por Rasmus Lerdorf, permitindo, inicialmente, a criação de aplicativos simples para a *web*. Seu criador, desde o início, liberou o código fonte do PHP, o que permitiu que usuários adicionassem ou mudassem diversas funções e corrigissem possíveis problemas. Hoje, o PHP funciona na maior parte dos sistemas operacionais (Windows, Linux, Mac, entre outros) e funciona bem em diversos bancos de dados como o MSSQL, Oracle, DB2 e o MySQL.

Alguns dos principais exemplos do uso do PHP estão em *sites* como a Wikipédia e o Facebook.

Procedimentos Metodológicos

Com a finalidade de desenvolver um sistema computacional que oferecesse a possibilidade de calcular as medidas descritivas de variáveis quantitativas de uma base de dados, utilizou-se a linguagem de programação *Hypertext Preprocessor* (PHP), a linguagem de marcação *Hypertext Markup Language* (HTML), a linguagem *Cascading Style Sheets* (CSS), o *framework Bootstrap* e linguagem *JavaScript*. Juntando todas estas tecnologias, foi criada uma calculadora estatística que se adapta conforme os dados inseridos e as opções operacionais selecionadas. O HTML foi utilizado para criar todo o layout da *webpage*, enquanto o CSS, juntamente com Bootstrap, foram responsáveis pela estilização de todos os componentes do layout. O PHP foi utilizado para realizar todas as operações matemáticas das medidas descritivas do sistema. Utilizou-se o JavaScript realizar as buscas e respostas do sistema computacional, isto é, interligar todas as funções e operações.



Resultados e discussões

Tendo como base teórica as aulas de Probabilidade e Estatística, ministradas na Universidade de Cruz Alta aos acadêmicos de Ciências da Computação, foi desenvolvido o projeto da calculadora estatística usando a Linguagem PHP. Em aula, conceitos estatísticos foram apresentados e explicados, sendo posteriormente aplicados no programa criado.

Utilizando de um layout simples e intuitivo, a calculadora permitirá a efetuação de diversas operações matemáticas na área estatística, podendo ser uma auxiliadora no processo de compreensão e aprendizagem de diversas pessoas, ligadas ou não a computação.

Seu uso não é complexo; consta digitar os números desejados para os cálculos em um espaço indicado, podendo estes números serem separados por vírgula, espaço ou quebra de linha. Após isso, o usuário irá indicar quais as operações que deseja fazer e então deverá clicar no botão “Calcular” logo abaixo das opções.

CALCULADORA ESTATÍSTICA

Abaixo é apresentada uma calculadora estatística, que faz a análise descritiva de variáveis quantitativas, onde esta permite calcular um número de propriedades estatísticas de uma amostra, como por exemplo: média, mediana, mínimo, máximo, entre outras mais!

INSTRUÇÕES: Digite no mínimo 3 valores, sendo estes separados por espaço em branco, vírgula ou nova linha!

Valores para o calculo

Operações

- | | | |
|---------------------------------|---|--|
| <input type="checkbox"/> Mínimo | <input type="checkbox"/> Primeiro Quartil | <input type="checkbox"/> Variância Amostral |
| <input type="checkbox"/> Máximo | <input type="checkbox"/> Segundo Quartil | <input type="checkbox"/> Desvio padrão amostral |
| <input type="checkbox"/> Soma | <input type="checkbox"/> Terceiro Quartil | <input type="checkbox"/> Coeficiente de variação |
| <input type="checkbox"/> Média | <input type="checkbox"/> Amplitude | |
| <input type="checkbox"/> Moda | <input type="checkbox"/> Intervalo-interquartil | |

Calcular

Imagem do layout da calculadora. Fonte: do Autor.

• Sobre as operações

São oferecidas treze possíveis operações na calculadora, podendo elas serem executadas simultaneamente, dependendo da vontade do usuário. Essas operações são:

Mínimo – o programa vai checar e localizar o menor número digitado na caixa de texto.



Máximo – assim como no mínimo, a calculadora procura um número digitado pelo usuário, porém agora, o maior entre eles.

Soma – soma-se todos os números expostos no programa.

Média – o usuário entra com os valores, que é somado e dividido pelo número de valores em questão, resultando assim, a média.

Moda – nada mais é que o número que aparece mais vezes na caixa de texto.

Primeiro Quartil (Q1) - valor que deixa 25% das observações à sua esquerda.

Segundo Quartil (Q2 = mediana) - valor que deixa 50% das observações à sua esquerda.

Terceiro Quartil (Q3) - valor que deixa 75% das observações à sua esquerda.

- Os quartis inferior e superior, Q1 e Q3, são definidos como os valores abaixo dos quais estão um quarto e três quartos, respectivamente, dos dados. Estes três valores normalmente são usados para resumir os dados juntamente com o mínimo e o máximo. Eles são obtidos ordenando-se os dados do menor para o maior e, então, contando-se o número apropriado de observações (n). Inicialmente encontramos o termo e depois localizamos nos dados ordenados os seus respectivos valores.

Amplitude - representa a diferença entre o maior e o menor valor de um conjunto de dados. Ela mostra a dispersão dos valores de uma série. Se a amplitude for um número elevado, então os valores da série estão distribuídos afastados; se a amplitude for um número baixo, então, os valores na série estão próximos uns dos outros.

Intervalo-Interquartil - utilizada em análises estatísticas para ajudar nas conclusões a cerca de um conjunto de dados. Geralmente é mais preferível que a variância para se medir a dispersão de um conjunto de dados pois exclui a maioria dos elementos discrepantes do conjunto. Enquanto o desvio padrão e a variância calculam a medida de dispersão sem levar em conta a ordem dos dados, o intervalo interquartil avalia a dispersão de dados somente depois de ordená-los em ordem crescente. O intervalo interquartil é calculado com base no cálculo de quartis, sendo o primeiro quartil (inferior), o quartil intermediário (mediana), o terceiro quartil (superior), que estão ligados ao conceito de quartil. A diferença entre o quartil superior e o quartil inferior determina o intervalo interquartil.



Variância amostral - a variância de uma amostra $\{x_1, \dots, x_n\}$ de n elementos é definida como a soma ao quadrado dos desvios dos elementos em relação à sua média \bar{x} dividido por $(n-1)$. Ou seja, a variância amostral é dada por:

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n - 1}$$

- Ao utilizarmos a média amostral como estimador de m para calcularmos a variância amostral, perdemos 1 grau de liberdade em relação à variância populacional.

Desvio padrão amostral - O desvio padrão amostral de um conjunto de dados é igual à raiz quadrada da variância amostral. Desta forma, o desvio padrão amostral é dado por:

$$s = \sqrt{s^2} = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n - 1}}$$

Coefficiente de variação - O coeficiente de variação é usado para expressar a variabilidade dos dados estatísticos excluindo a influência da ordem de grandeza da variável.

- O coeficiente de variação é usado para analisar a dispersão em termos relativos a seu valor médio quando duas ou mais séries de valores apresentam unidades de medida diferentes. Dessa forma, podemos dizer que o coeficiente de variação é uma forma de expressar a variabilidade dos dados excluindo a influência da ordem de grandeza da variável.

O cálculo do coeficiente de variação é feito através da fórmula:

$$CV = \frac{s}{\bar{x}} \cdot 100$$

Onde:

$s \rightarrow$ é o desvio padrão

$\bar{x} \rightarrow$ é a média dos dados

$CV \rightarrow$ é o coeficiente de variação

- Como o coeficiente de variação analisa a dispersão em termos relativos, ele será dado em %. Quanto menor for o valor do coeficiente de variação, mais homogêneos serão os dados, ou seja, menor será a dispersão em torno da média. De uma forma geral, se o CV:

For menor ou igual a 15% \rightarrow baixa dispersão: dados homogêneos.



For entre 15 e 30% → média dispersão.

For maior que 30% → alta dispersão: dados heterogêneos.

Considerações finais

O desenvolvimento deste sistema computacional, que calcula as medidas descritivas de variáveis quantitativas, permitiu que conhecimentos abordados de forma apenas teórica em sala de aula pudessem ser postos em prática usando os conhecimentos das áreas da Probabilidade e Estatística e da Ciência da Computação. Ao construir esta calculadora estatística, é possível concluir o quanto as linguagens de programação podem auxiliar na elaboração de sistemas computacionais aplicados em qualquer área do conhecimento científico, assim como, em metodologias de ensino diferenciadas na resolução de exercícios em sala de aula.

REFERÊNCIAS

Barbetta, P. A. (2014). Estatística Aplicada às Ciências Sociais. 9 ed. Florianópolis: Ed. da UFSC.

Bussab, W. de O., Morettin, P. A. (2014). Estatística Básica. 8 ed. São Paulo: Saraiva.

Devore, J. L., Berk, K. N. (2012). Modern Mathematical Statistics with Applications. 2 ed. New York: Springer.

Goos, P., J. F., Meintrup, D. (2015). Statistics with JMP: Graphs, Descriptive Statistics and Probability. West Sussex: Wiley.

Manual do PHP. **O que é PHP**. Disponível em: <https://secure.php.net/manual/pt_BR/intro-what-is.php> Acessado em 05 de setembro de 2018.

Herbert, F., G. PHP Guia de Consulta Rápida. São Paulo: Novatec.

Dalloglio, P. (2007). PHP Programando com Orientação a Objetos. 1 ed. São Paulo: Novatec.

Triola, M., F.(2008). Introdução á estatística. 10 ed. Rio de Janeiro: LTC.