

# SOFTWARE DE PROBABILIDADE E ESTATÍSTICA: UMA VISÃO MULTIDISCIPLINAR NA UNIVERSIDADE POSITIVO

**Giancarlo de França Aguiar<sup>1</sup>; Maurício Perreto<sup>2</sup>; Bárbara C. X. C. Aguiar<sup>3</sup>**

Universidade Positivo – UP, Núcleo de Ciências Exatas e Tecnológicas  
Rua Prof. Pedro Viriato Parigot de Souza, 5300, Campo Comprido  
81280-330 – Curitiba – Paraná

<sup>1</sup>giancarl@up.edu.br

<sup>2</sup>mperretto@gmail.com

<sup>3</sup> Universidade Federal do Paraná, Departamento de Matemática  
Centro politécnico, Jardim das Américas  
81531-990 – Curitiba – Paraná  
babimatematica@yahoo.com.br

**Resumo:** *A cada dia, o desenvolvimento de softwares acadêmicos em cursos de graduação se tornará mais comum, sejam eles desenvolvidos por professores ou por estudantes. Neste trabalho será ilustrado o processo de construção de um software desenvolvido nas disciplinas de Probabilidade e Estatística e Fundamentos de Programação, realizado com estudantes do curso de Engenharia da Computação da Universidade Positivo, em Curitiba no estado do Paraná. Os laboratórios virtuais acadêmicos ou softwares podem contribuir na relação ensino-aprendizagem como fator determinante e motivacional, despertando o interesse nos estudantes tanto para a pesquisa como na aplicação dos conhecimentos adquiridos no decorrer do curso. Os estudantes desenvolveram em um primeiro momento um software tutorial com conceitos, definições e capaz de resolver certas Estatísticas. Este instrumento tem como objetivo auxiliar novos estudantes no entendimento da matéria (Estatística), bem como, servir de estrutura para novas implementações em Fundamentos de Programação.*

**Palavras-chave:** *Software Tutorial, Probabilidade e Estatística, Fundamentos de Programação*

## 1. INTRODUÇÃO

A competitividade do mundo moderno fez surgir à necessidade de o homem estar cada vez mais preparado para o mercado de trabalho. As empresas buscam o profissional apto à tomada de decisões, que possua conhecimentos generalistas e que suas decisões sejam baseadas em tecnologia (AGUIAR *et al.*, 2006).

Atualmente, a utilização de recursos computacionais para o tratamento de dados, sejam eles qualitativos ou quantitativos tornou-se imprescindível. O mercado exige agilidade e precisão no estudo de dados. Quando esta tarefa cabe a um ser humano ele tanto pode cometer

erros de precisão, como entrar em fadiga, caso exista o excesso de trabalho. Pelo contrário, os computadores modernos possuem excelente precisão e são muito mais rápidos que os seres humanos (AGUIAR *et al.*, 2006).

Contudo, foi inevitável o desenvolvimento acelerado de recursos computacionais direcionados a resolução dos mais variados problemas, sejam eles de ordem econômica, administrativa, industriais e de engenharia. Tornar-se-á mais comum a cada dia, a construção de laboratórios virtuais acadêmicos em cursos de graduação, sejam eles desenvolvidos por professores ou por estudantes.

Neste trabalho será ilustrado o processo de construção de um software desenvolvido nas disciplinas de Probabilidade e Estatística e Fundamentos de Programação, realizado com estudantes do curso de Engenharia da Computação da Universidade Positivo, em Curitiba no estado do Paraná. Os laboratórios virtuais acadêmicos ou softwares podem contribuir na relação ensino-aprendizagem como fator determinante e motivacional, despertando o interesse nos estudantes tanto para a pesquisa como na aplicação dos conhecimentos adquiridos no decorrer do curso.

O fator ensino-aprendizagem foi de extrema relevância quanto à disciplina de Fundamentos de Programação, pois esta apresenta diversas estruturas de dados que permitem a otimização de diversos processos e resolução de problemas. Contudo, o enfoque dado em materiais bibliográficos da disciplina (DROZDEK, 1998; TANNENBAUM, 1995; LAFORE, 1999), apresentam certas estruturas de forma simples e sem exemplos da aplicação prática, o que desestimula os alunos quanto aos assuntos apresentados.

Como proposta para resolução dos dois problemas, os estudantes desenvolveram softwares tutoriais com conceitos, definições e capaz de resolver certas Estatísticas. Entretanto, para o armazenamento e manipulação dos dados deveriam ser utilizadas as estruturas de dados comumente apresentadas na disciplina, listas, filas, pilhas e matrizes esparsas, assim como métodos de ordenação. Este instrumento tem como objetivo auxiliar novos estudantes no entendimento da matéria (Estatística), bem como, servir de estrutura para novas implementações em Fundamentos de Programação.

## **1.1 Filas e Pilhas**

As filas e pilhas são estruturas básicas da computação, sendo as primeiras a serem apresentadas, elas também são conhecidas como estruturas FIFO (*First-in/First-out*) e LIFO (*Last-in/First-out*), respectivamente. As filas permitem o armazenamento e manipulação de informações na ordem em que esta é repassada a estrutura (DROZDEK, 1998). As filas são utilizadas na prática em sistemas computacionais em roteadores de pacotes em redes, sendo este um meio motivacional.

## **1.2 Algoritmos de Ordenação**

Os algoritmos de ordenação objetivam ordenar uma seqüência de dados armazenados na memória. Ordenar é o processo de criar uma maneira lógica de organizar um grupo de informações que permita a localização rápida de uma informação, sem necessariamente verificar, um por um, os elementos armazenados.

Existe uma série de algoritmos para realizar a ordenação de dados. Entre eles, são trabalhados os seguintes métodos: Bolha, Inserção Direta, Seleção Direta, *MergeSort* e *QuickSort*. Os três primeiros métodos se utilizam do grupo de informações completo, ou seja, realizando a comparação de todas as informações, entretanto, estes algoritmos são essencialmente mais lentos que os demais. Os dois últimos métodos se utilizam da estratégia

“dividir para conquistar”, ou seja, diminuindo o grupo de dados e desta forma realizando um menor número de comparações. Isto ocorre dado ao fato de o número de comparações que devem ser realizadas para ordenar um vetor, crescer exponencialmente em relação ao número de elementos a ser ordenado. (LAFORE, 1999)

## 2. DESENVOLVIMENTO

A seguir são ilustrados 4 (quatro) momentos para o desenvolvimento e construção do laboratório virtual de Probabilidade e Estatística. São eles:

- 1- No início do ano letivo foram dedicadas 2 horas-aula de cada disciplina participante do projeto para a explanação do trabalho proposto. Seguem algumas das orientações:
  - a. Divisão de equipes de trabalho;
  - b. Objetivos de cada equipe;
  - c. Tarefas das equipes durante o ano letivo;
  - d. Apresentação dos trabalhos em seminários para a turma;
  - e. Exposição dos trabalhos em painéis para divulgação no campus universitário.
- 2- Em seguida, utilizando uma ferramenta de programação visual, trabalhada em sala de aula na disciplina de Fundamentos de Programação, os estudantes iniciaram o desenvolvimento dos softwares tutoriais, utilizando para isto, os materiais de sala de aula trabalhados na disciplina de Probabilidade e Estatística.
- 3- A cada bimestre foi realizado um encontro com os professores orientadores para a ponderação de nota para o trabalho desenvolvido até aquele presente momento.
- 4- O prazo de término do trabalho foi o 3º bimestre com a data de apresentação dos seminários já pré-definida. Após as explanações e defesas de trabalhos.

### 2.1 Softwares

Os softwares desenvolvidos são tutoriais que ilustram de forma prática como calcular distribuições de probabilidades, com exemplos didáticos e bem conhecidos de problemas nessa área. Para a implementação dos softwares foram definidas como estruturas de dados para armazenamento as filas, pilhas e listas. Que são as estruturas mais simples para o armazenamento de informação.

Cada equipe ficou responsável pela implementação de uma estrutura de armazenamento diferente. Além disso, as equipes tiveram que definir métodos de ordenação e busca próprios para as informações de probabilidade. Sendo então definida uma competição entre eles para obtenção do menor tempo no cálculo correto da distribuição. Iniciou-se o trabalho implementando os métodos Bolha e *QuickSort* de ordenação em forma crescente e a busca binária como retorno da informação.

A partir deste ponto, cada equipe deveria alterar seu algoritmo em busca de um tempo inferior ao obtido pelo método apresentado na literatura. Na figura 1 abaixo temos a tela inicial de um dos programas desenvolvidos no trabalho.

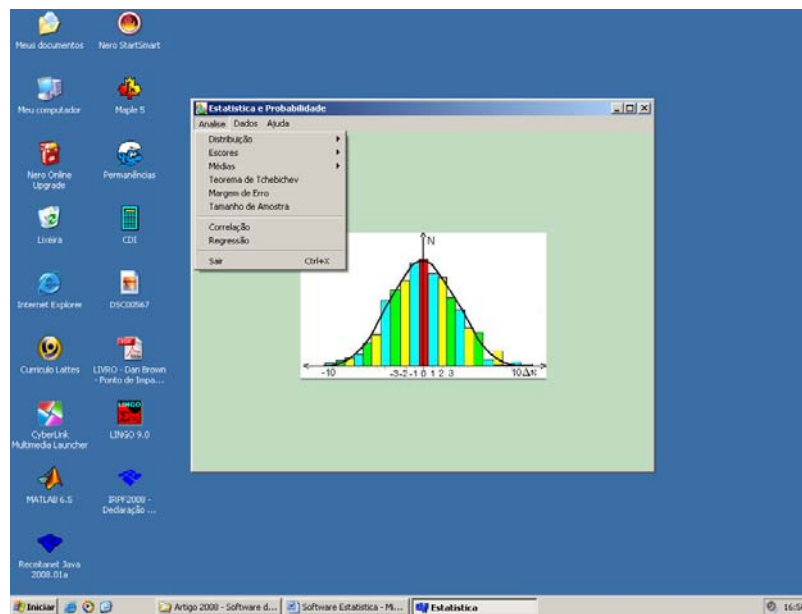


Figura 1 – Tela inicial de um dos programas.

## 2.2 Alguns Resultados

A seguir está ilustrado um conjunto de figuras com algumas das ferramentas de um dos trabalhos selecionados (software desenvolvido). A figura 2 ilustra o cálculo da probabilidade utilizando a Distribuição Binomial. Como exemplo, temos a seguinte situação problema:

**“Dado que 10% das pessoas são canhotas, suponha que queiramos achar a probabilidade de obter exatamente 3 estudantes canhotos em uma turma de 15 estudantes”.**

Como cada prova deve ter todos os resultados classificados em duas categorias, sugere-se a utilização da Distribuição Binomial, com  $x = 3$  (sucesso em  $n$  provas),  $n = 15$  (número de provas),  $p = 0,1$  (probabilidade de sucesso) e  $q = 0,9$  (complemento do sucesso). Após a inserção dos dados no software e clicar em binomial, o resultado é fornecido (0,00129).

Devemos então multiplicar o resultado por dez mil (10.000) devido a uma limitação do software, para obtermos a probabilidade, ou seja, aproximadamente 12,9 % dos estudantes da turma com 15 estudantes, são canhotos.

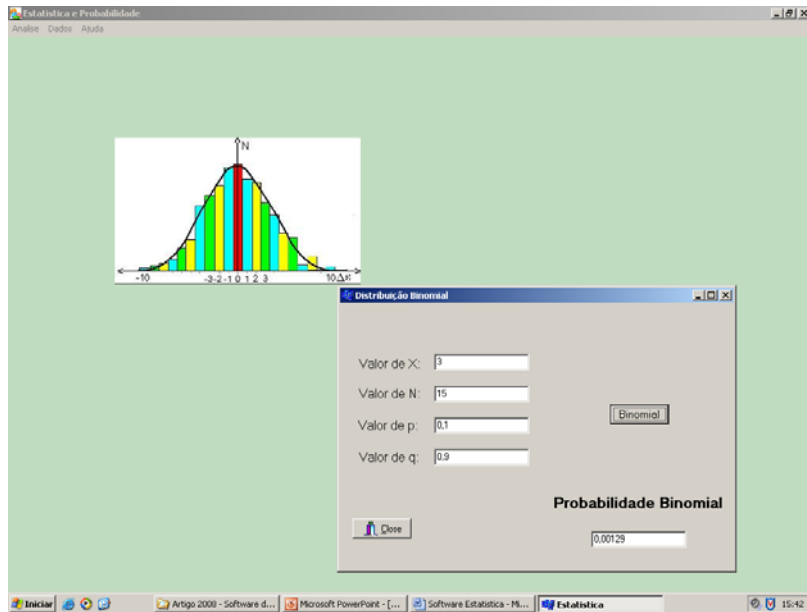


Figura 2 – Distribuição binomial de probabilidades.

Na figura 3 está ilustrado o cálculo de um intervalo de confiança para uma média populacional utilizando o teorema de *Tcheibichev* (a proporção ou fração de qualquer conjunto de dados a menos de  $K$  desvios-padrão a contar da média, é aproximadamente igual à diferença da unidade pelo inverso do quadrado do desvio  $K$ ). A equação 1 abaixo ilustra o teorema.

$$P_p(k) = 1 - \frac{1}{k^2} \quad (1)$$

Para ilustrar a demonstração temos o seguinte problema:

**“O Q.I. (teste ou quociente de inteligência) de um grupo de adultos apresenta distribuição com média 100 e desvio-padrão 15. Iremos encontrar um intervalo de confiança para o QI populacional”.**

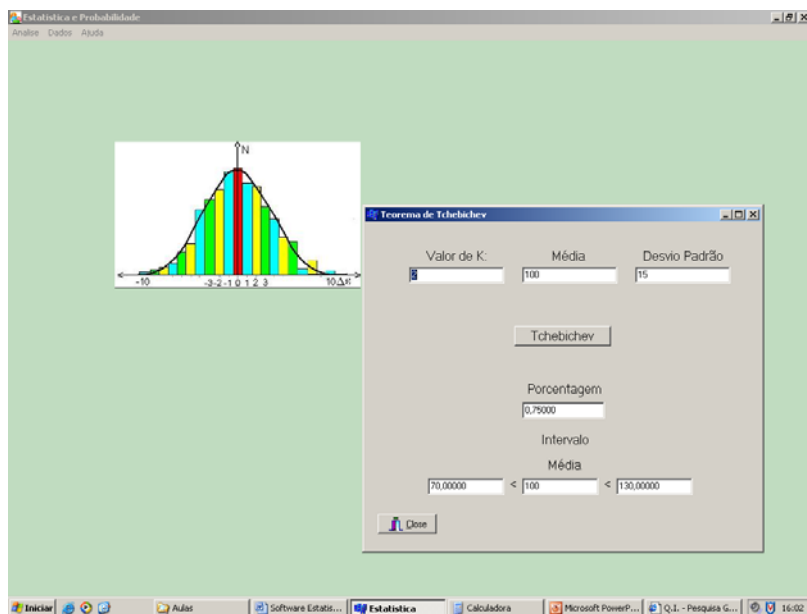


Figura 3 – Intervalo de confiança para a média populacional.

No exemplo anterior, O programa pergunta qual o valor de  $k$  (neste caso 2), a média amostral (100) e desvio-padrão amostral (15). Preenchidos estes campos, basta o operador clicar no botão *Tchebichev*, obtendo o resultado para o intervalo populacional, ou seja, que 75% da população têm um Q.I. no intervalo de 70 e 130 ( $70 \leq \mu \leq 130$ ).

Pode-se observar a seguir (figura 4) uma tela que expõe a média, amplitude, variância e desvio-padrão para um conjunto de dados não agrupados. Como exemplo seja a variável aleatória  $x$  assumindo os seguintes valores  $x = 1,2,3,4,5,6,7,8$ . Então utilizando a ferramenta computacional obtemos as seguintes medidas descritivas aproximadas: média aritmética 4,500, amplitude 7,000, variância 6,000 e desvio-padrão 2,449.

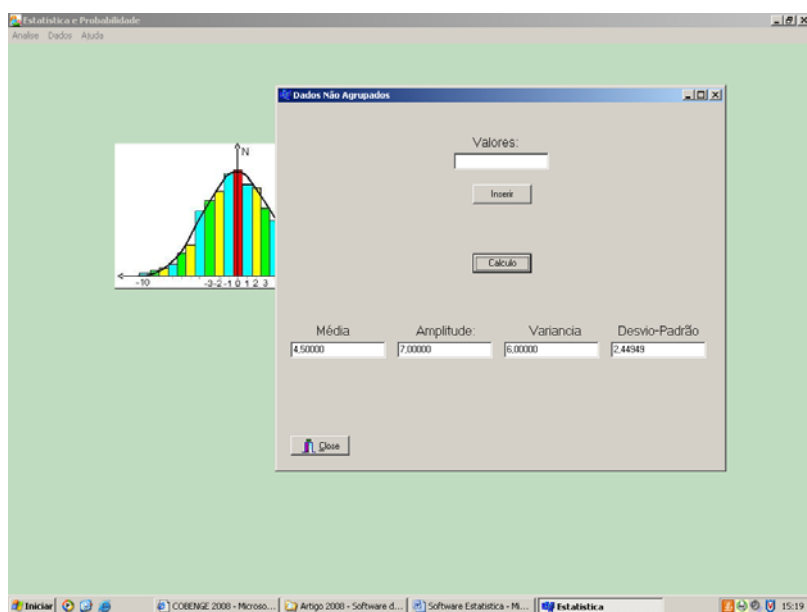


Figura 4 – Medidas descritivas de dados.

O software pode calcular ainda o índice de correlação linear de *Pearson*, as médias harmônica, ponderada, geométrica e quadrática para um conjunto de dados, e também medidas descritivas para dados agrupados. Novos ajustes e validações devem ser realizados para que ele possa ser utilizado como ferramenta de ensino em cursos de graduação.

A seguir estão ilustradas algumas abordagens de outro software, o ProEsa desenvolvido por outro grupo de trabalho. Na figura 5 abaixo foram inseridos os dados de uma distribuição de probabilidades para o cálculo da média, variância e desvio-padrão para a distribuição.

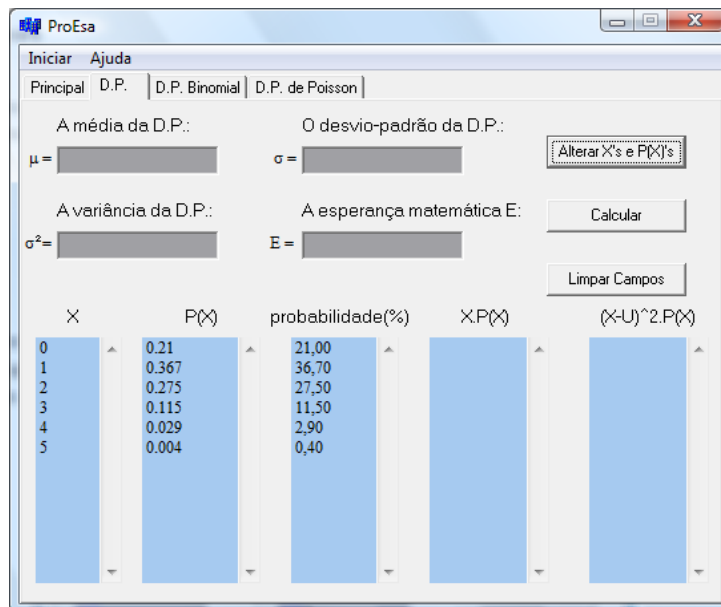


Figura 5 – Dados de uma distribuição de probabilidades inseridos no pacote.

O software solicita o número de variáveis para a distribuição, o valor da variável e sua probabilidade. Ao clicar em calcular, obtemos a média (1,4), o desvio-padrão (1,05), a variância (1,11) e a esperança matemática (1,4) como ilustra a figura 6 a seguir.

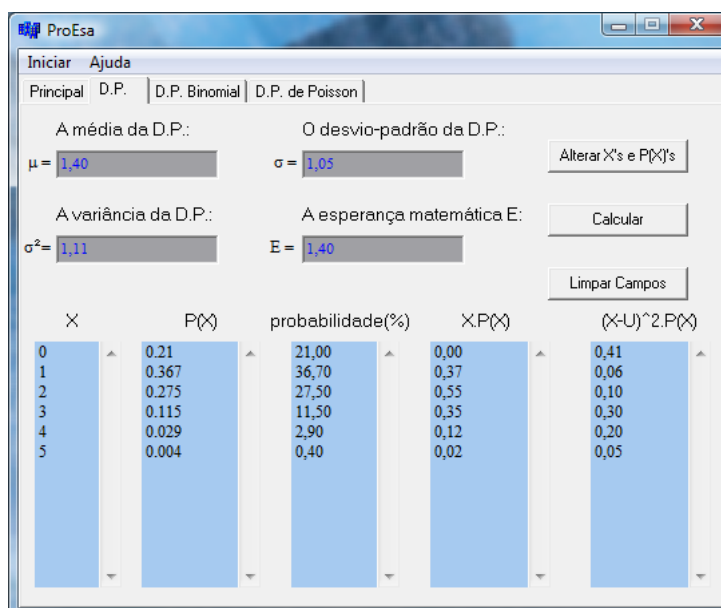


Figura 6 – Cálculo da média, desvio-padrão, variância e esperança matemática.

Ilustramos agora a aplicação do software no cálculo de uma distribuição de probabilidades de variáveis aleatórias discretas utilizando a distribuição binomial (figura 7 a seguir). Como exemplo, temos a seguinte situação problema:

**“Sabe-se que 80% dos bebês que nascem na região sudeste são portadores de icterícia, uma pigmentação amarelada na pele que tende a desaparecer nas primeiras semanas de vida. Contudo 20% dos casos de icterícia devem ser tratados. Em 20 nascimentos escolhidos aleatoriamente, na região sudeste; qual a probabilidade de três crianças selecionadas aleatoriamente precisarem do tratamento?”.**

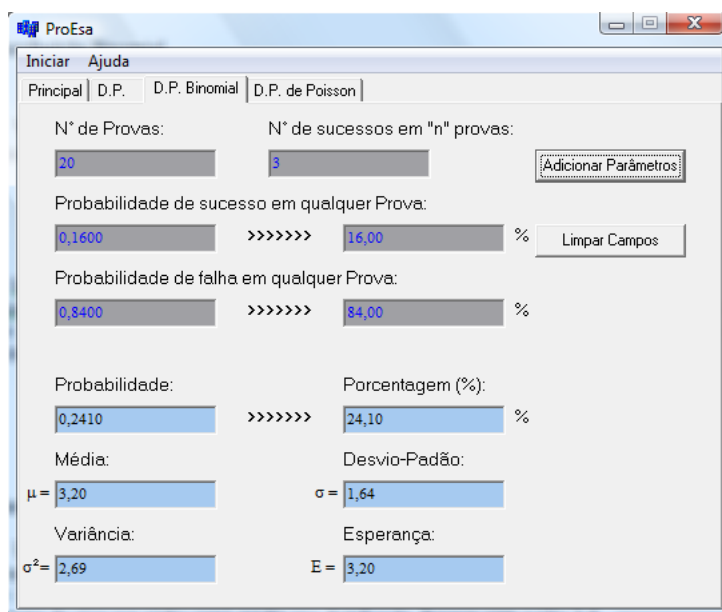


Figura 7 – Cálculo da probabilidade binomial, média, desvio-padrão, variância e esperança.

Adicionamos os parâmetros número de provas (20) e o número de sucessos em “n” provas (3). O Software solicita a probabilidade de sucesso nas “n” provas, que neste contexto é o produto  $p = (0,8) \times (0,2)$  (porcentagem dos bebês que nascem com icterícia versus probabilidade de casos que devem ser tratados). A probabilidade de encontrarmos exatamente três crianças portadoras de icterícia em uma amostra de 20 crianças é de aproximadamente 24,1% de chance de ocorrência como relata o software.

### 3. CONSIDERAÇÕES FINAIS

O trabalho contribuiu para o embasamento teórico (pesquisa dos tópicos selecionados pelo orientador) quando realizado em conjunto com o desenvolvimento físico (criação dos painéis) e com a aplicação tecnológica (desenvolvimento do software).

O trabalho foi de ordem complementar, ou seja, extracurricular sem pontuação na média dos estudantes. Este fato pode ter contribuído para a falta de comprometimento por parte de algumas equipes que não entregaram o projeto. Todavia, a sugestão é ponderar nota ao trabalho e acolher melhores resultados.

Notadamente foi percebido um engajamento dos estudantes quanto à forma de modelar a estrutura de dados para melhor adaptá-la ao problema proposto;



Durante todo o ano letivo, muitos estudantes se mostravam mais motivados no ensino e aprendizagem de novos conteúdos.

O desenvolvimento culminou em uma grande rede de grupos de estudo para a sua incubação, o que aproximou ainda mais os estudantes entre si.

Dadas às duas condições anteriores (motivação e interação), foi construído um processo ensino-aprendizagem que pode gerar resultados mais significativos aos alunos.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

AGUIAR, G. F.; AGUIAR, B. C. X. C.; WILHELM, V. E. Obtenção de Índices de Eficiência para a Metodologia *Data Envelopment Analysis* Utilizando a Planilha Eletrônica Microsoft Excel. **Revista da Vinci**, Curitiba, v.3, n.1, p. 157-169, 2006.

DROZDEK, A. *Estrutura de Dados e Algoritmos em C++*. São Paulo: Thomson Learning Ltda, 1998.

EVARISTO, J. **Aprendendo a programar**: Programação em linguagem C. Rio de Janeiro: Book Express, 2001.

GUIMARÃES, A.; LAGES, N. **Algoritmos e estrutura de dados**. Rio de Janeiro: LTC, 1994.

JAMSA, k.; KLANDER, L. **Programação em C/C++**. São Paulo: Makron Books, 1999.

LAFORE, R. **Aprenda em 24 horas Estrutura de Dados e Algoritmos**. Rio de Janeiro: Campus, 1999.

MORETTIN L.B. **Estatística Básica**. Vol. 1. São Paulo, Saraiva, 2002.

TANNENBAUM, A. M. **Estruturas de Dados usando C**. São Paulo: Makron Books, 1995.

TRIOLA, M. F. **Introdução à Estatística**. 7<sup>a</sup> ed., Rio de Janeiro, LTC, 1999.

## **SOFTWARE OF PROBABILITY AND STATISTICS: A VISION TO MULTIDISCIPLINE IN THE POSITIVO UNIVERSITY**

**Abstract:** *Each day, the development of academics softwares in graduation courses will become more common, being those developed either by professors or by students. In this work the elaboration of one such software, promoted by the disciplines of Probability and Statistics and Programming of Beddings, will be illustrated. It was developed by the students of the Computer Engineering course at Positivo University, in Curitiba state of Paraná. The academic virtual laboratories or softwares can contribute in the teaching/apprenticeship relation as a motivational factor, promoting the student's interest in research as well as in the application of the knowledge acquired during the course. First, the students had to*

*develop a tutorial software with essential concepts and definitions, being capable of solving certain statisticians. This new instrument has as objective assistant students in the agreement of the substance (Statistics), as well as, to serve of structure for new implementations in Beddings of Programming.*

**Key-words:** *Tutorial software, Probability and Statistics, Programming of Beddings*