Capítulo 9

Distribuições estatísticas

Objetivo

- Apresentar as principais distribuições estatísticas discretas e contínuas.

1. Distribuições discretas

1.1 Distribuição binomial

A densidade binomial é uma das mais importantes em teoria da probabilidade. Ela surge como uma idealização matemática de diversas situações comuns na "vida real" e está intimamente ligada à amostragem com reposição. A situação clássica em que usamos uma densidade binomial é a seguinte.

Uma experiência tem apenas 2 resultados possíveis : "sucesso" e "falha", em que a probabilidade de "sucesso" é p e a probabilidade de "falha" é $q = 1 - p$.

A experiência é repetida um número fixo (n) de vezes, sempre nas mesmas condições, de tal forma que as probabilidades de "sucesso" (p) e "falha" ($q = 1 - p$) se mantêm inalteradas a cada repetição. As diversas repetições da experiência são feitas de maneira independente, ou seja, o resultado de uma repetição não afeta o resultado das outras.

A variável aleatória X que mede o número de "sucessos" nas n repetições da experiência é uma variável discreta, com valores possíveis 0, 1, 2, ... , n. Dizemos que esta variável tem densidade binomial com parâmetros n e p, e escrevemos que X ~ Bin(n, p).

Nota : a escolha de um tipo de resultado como "sucesso" ou "falha" não implica em qualquer julgamento sobre o resultado ser "bom" ou "ruim", é apenas uma questão de nomenclatura. Na verdade, a escolha do que é um "sucesso" ou "falha" depende da questão de interesse ao analisar o problema, de forma que o que é "sucesso" numa situação pode ser a "falha" num problema semelhante.

A equação da distribuição binomial é a seguinte.

$$P(X = x) = f(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

1.2 Distribuição hipergeométrica

Suponha que, na população, existem r objetos do tipo A ("sucessos") e $N-r$ objetos do tipo B ("falhas").

Seja X o número de objetos do tipo A na amostra. Então as probabilidades dos diversos valores de X são dadas pela seguinte fórmula.

$$P(X = x) = f(x) = \frac{\binom{r}{x} \cdot \binom{N-r}{n-x}}{\binom{N}{n}},$$

onde x = 0, 1, 2, ..., mín(n,r).

Esta é a densidade hipergeómétrica, usada para calcular probabilidades no caso de amostragem sem reposição.

Pode-se provar que f(x) acima definida integra a 1 (não é muito fácil) .

1.3 Densidade Poisson

Esta densidade é usada principalmente para modelar o número de ocorrências de um evento "raro" (de probabilidade baixa) durante um intervalo de tempo especificado. Por exemplo, o número de acidentes numa estrada durante um fim de semana; o número de bactérias presentes numa solução após um certo período são, entre outros, eventos modelados pela distribuição de Poisson. Além disso, a distribuição de Poisson surge como um caso limite da distribuição Bin(n,p) quando n é grande, e p é pequeno (próximo de zero) e, neste contexto, é muito útil em aproximações numéricas.

A derivação da densidade Poisson pode ser feita de duas formas: a primeira está relacionada com o "processo de Poisson", e a segunda surge como uma aproximação da densidade binomial.

1.4 Processo de Poisson

Considere uma sequência de eventos que ocorrem ao longo do tempo, como o número de carros vermelhos que param num sinal, o número de chamadas telefônicas que chegam a uma estação durante um certo intervalo de tempo.

Seja Xt o número de ocorrência no intervalo de tempo [0, t]. Claramente, Xt é uma variável aleatória discreta com valores possíveis 0, 1, 2, ... Para derivar a densidade de Xt, partimos das seguintes premissas.

Seja Δt um intervalo de tempo pequeno. Então.

- A probabilidade de exatamente uma ocorrência em um intervalo de tempo **D**t é aproximadamente k Δ t.
- A probabilidade de exatamente zero ocorrências em um intervalo de tempo **D**t é aproximadamente $(1 – k)\Delta t$.
- A probabilidade de duas ou mais ocorrências em um intervalo de tempo Δt é igual a um certo $o(\Delta t)$, onde $o(\Delta t)/\Delta t$ tende a zero à medida que Δt tende a zero. Em outras palavras, a probabilidade de duas ou mais ocorrências em um intervalo de tempo Δt é um valor muito pequeno, e esse valor decresce a zero mais rapidamente que o comprimento do intervalo Δt.

Essas três premissas definem o tipo de processo que pode ser chamado de um processo de Poisson.

O parâmetro *k* acima é um número real > 0, chamado de **taxa média de ocorrência**.

Para cada instante *t* > 0, seja

$P(X_t = x) = p_x(t)$, onde x = 0,1,2,...

Fixando um instante qualquer *t* e aplicando a segunda premissa, nos dá

$$p_1(t + \Delta t) \cong [1 - k\Delta t]p_0(t)$$

Subtraindo p0(t) de ambos os lados e dividindo o resultado por Δt, leva a

$$\frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} \cong kp_0(t)$$

Tomandose o limite desta última expressão quando Dt tende a zero, encontramos, do lado esquerdo, a derivada de p0(t). Isso nos dá a equação diferencial

$$p_0(t) = - kp_0(t)$$

Para x > 0, pode-se provar que as premissas resultam no seguinte sistema de equações diferenciais

$$p_x(t) = -kp_x(t) + kp_{x-1}(t),\ \text{onde } x = 1,2,3,...$$

A solução do sistema dado por (4) e (5) é

$$p_x(t) = \frac{(kt)^x e^{-kt}}{x!},\ \text{onde } x = 1,2,3,...$$

Para qualquer intervalo [0, t], se fixarmos t e fizermos l = kt, a equação acima reduzse a

$$P(X = x) = f(x) = \frac{\lambda^x e^{-\lambda t}}{x!},\ \text{onde } x = 1,2,3,...$$

A densidade anterior é a densidade Poisson com parâmetro l > 0 e escrevemos

X ˜ Poisson(λ).

1.5 Como usar a densidade Poisson na prática

Selecione um intervalo de tempo fixo. Conte o número de ocorrência de um certo evento de interesse neste intervalo. Esse número de ocorrências é uma variável discreta com valores possíveis 0, 1, 2, ... Se o evento é tal que a probabilidade do número de ocorrências no intervalo ser 0 ou 1 é "grande", então o evento pode ser na prática modelado pela distribuição de Poisson. Uma densidade Poisson modela bem eventos "raros", isto é, que não acontecem com grande frequência para qualquer intervalo de tempo fixo.

Por exemplo, o número de automóveis Corsa que entram num estacionamento no Rio de Janeiro num intervalo de 1 hora certamente não é uma variável Poisson, mas o número de Ferraris que entram no estacionamento no mesmo período de tempo deve ser Poisson.

2. Distribuições contínuas

2.1 Densidade Uniforme

A densidade uniforme serve para modelar o seguinte fenômeno: "escolhe-se um número aleatoriamente num intervalo dado", por exemplo, o intervalo (0,1). A função "random", presente na maioria das linguagens de computador, nada mais é do que um mecanismo para gerar números com distribuição uniforme no intervalo (0,1).

Uma variável aleatória *X* tem densidade uniforme no intervalo (*a*,*b*), e escrevemos *X* ˜ Unif(*a*,*b*) se a sua densidade é

$$f(x) = \begin{cases} \dfrac{1}{b-a}, x \notin (a,b) \\ 0, x \notin (a,b) \end{cases}$$

Uma variável aleatória X com densidade uniforme no intervalo (a,b) tem a seguinte propriedade: qualquer subintervalo de comprimento d localizado dentro do intervalo (a,b) tem a mesma probabilidade.

A função de distribuição de uma variável aleatória Unif(*a*,*b*) é:

$$F(x) = \begin{cases} 0, x < a \\ \dfrac{x-a}{b-a}, x \notin [a,b] \\ 1, x > b \end{cases}$$

2.2 Densidade Exponencial

Uma variável aleatória com densidade exponencial é usada para modelar tempos de duração de equipamentos. Na verdade, existem densidades mais apropriadas para modelar esse fenômeno, pois, como veremos mais tarde, a densidade exponencial não leva em conta o desgaste do equipamento ao longo do tempo. A densidade exponencial é definida para variáveis contínuas e maiores que zero e depende de um parâmetro positivo, l.

Notação: X ˜ Expo(λ)

A densidade exponencial é dada pela fórmula:

$$f(x) = \lambda e^{-\lambda x}, x \geq 0$$

E f(x) = 0 se x < 0. Note que λ é > 0 sempre.

A função de distribuição é dada por

$$F(x) = \begin{cases} 0, x < 0 \\ \int_0^x \lambda e^{-\lambda u} du = -e^{\lambda u} \Big|_0^x = 1 - e^{-\lambda x}, x \geq 0 \end{cases}$$

Note que o limite da função de distribuição quando x tende a +∞ é um.

O próximo gráfico apresenta as densidades exponenciais com parâmetros $\lambda = 2, 4 \text{ e } 8$. Note que a densidade decai mais rápido quando λ é grande.

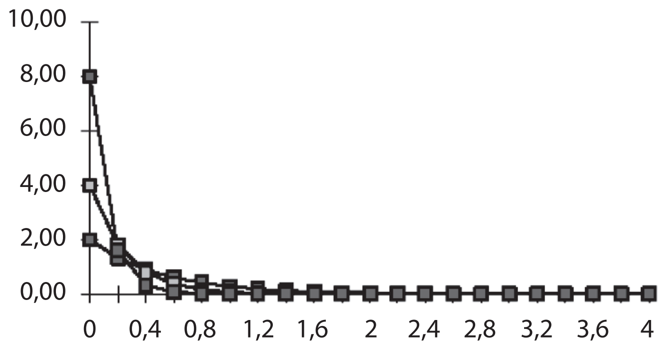

Figura 5 - Densidades exponenciais

O próximo gráfico exibe a função de distribuição de uma variável aleatória com parâmetros λ = 1 e λ = 2.

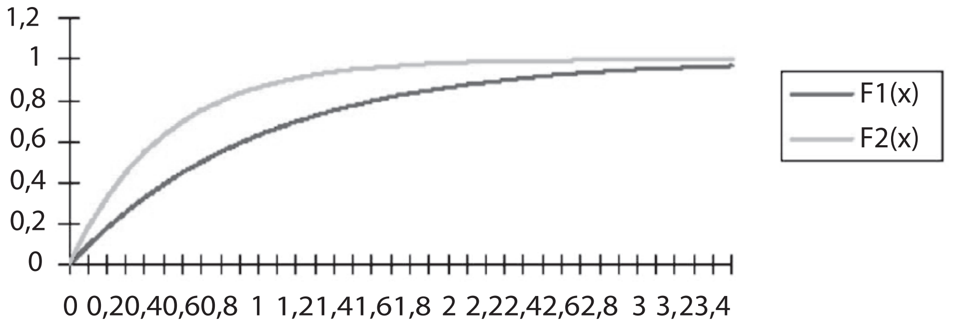

Figura 6 - Funções de distribuição densidades Expo(1) e Expo(2)

2.3 Função Gama

Seja **a** um número real maior que zero, não necessariamente inteiro. A função Gama com argumento a é definida por

$$\Gamma(a) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

a) Propriedades da Função Gama

1) $\Gamma(n) = (n-1)\Gamma(n-1)$ para $n > 1$

A demonstração deste fato usa integração por partes.

2) $\Gamma(n) = (n-1)!$ se n é inteiro > 1

3) $\Gamma(1) = 0! = 1$

4) $\Gamma\left(\frac{1}{1}\right) = \sqrt{\pi}$

b) Densidade Gama

Seja X uma variável aleatória contínua definida no intervalo(0, ∞). Dizemos que X tem densidade Gama com parâmetros α e β, e escrevemos X ~ Gama(α, β) se a densidade de X é

$$f(x) = \begin{cases} \dfrac{\beta^{\alpha} x^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x}, \text{onde } x \geq 0 \\ 0 \text{ se } x < 0 \end{cases}$$

Os parâmetros a e β são números reais positivos **a** é conhecido como parâmetro de forma, e β é o parâmetro de escala.

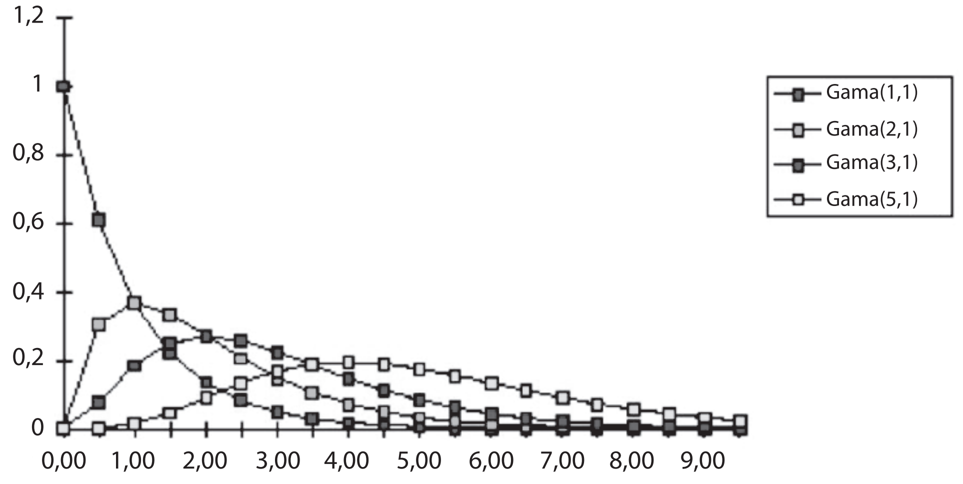

Figura 7 - Densidades Gama: gama(1,1), Gama(1,2), Gama(1,3) e Gama(1,5)

2.4 Densidade quiquadrado com k graus de liberdade)

Seja *X* uma variável aleatória contínua e positiva com densidade dada por

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{n-2}{2}} e^{\frac{-x}{2}} \text{ onde } x > 0$$

Tem densidade quiquadrado com (n - 1) graus de liberdade, e escrevemos:

$$X \sim \chi_k^2$$

A densidade quiquadrado com k graus de liberdade é apenas um caso particular da densidade gama. Na verdade,

$$\chi_k^2 = \text{Gama}\left(\alpha = \frac{k}{2}, \beta = \frac{1}{1}\right)$$

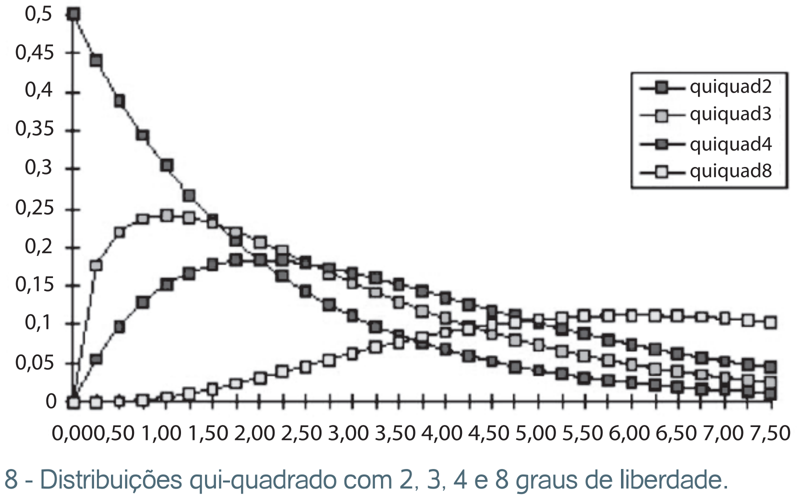

Figura 8 - Distribuições qui-quadrado com 2, 3, 4 e 8 graus de liberdade.

2.5 Distribuição Normal (Gaussiana)

A distribuição normal é, talvez, a mais importante das distribuições de probabilidade. Erros de mensuração de fenômenos físicos ou econômicos são frequentemente modelados pela distribuição normal, mas esta não é a única aplicação desta densidade. Por exemplo, a distribuição dos pesos, alturas e QI's das pessoas numa população também já foram modelados com sucesso por esta distribuição. A distribuição normal tem a forma de um sino e possui dois parâmetros, μ e σ^2.

A distribuição normal é também chamada de gaussiana em homenagem ao matemático Carl Friederich Gauss (1777-1855), que a utilizou pela primeira vez na modelagem de erros de medida. A distribuição normal também funciona como uma boa aproximação para outras densidades. Por exemplo, sob algumas condições pode-se provar que a densidade binomial pode ser aproximada pela normal.

a) Densidade Normal com média μ e variância σ^2

Seja X uma variável aleatória contínua definida nos números reais. Dizemos que X tem densidade normal com média μ e variância σ^2 se a densidade de X é

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{\sigma^2}}$$

Notação: *X* ~ N(μ, σ^2)

Note que o segundo parâmetro (σ^2) nesta notação é a variância de X. A seguir, exibimos gráfico das distribuições normais com média zero e variâncias 1, 2 e 4.

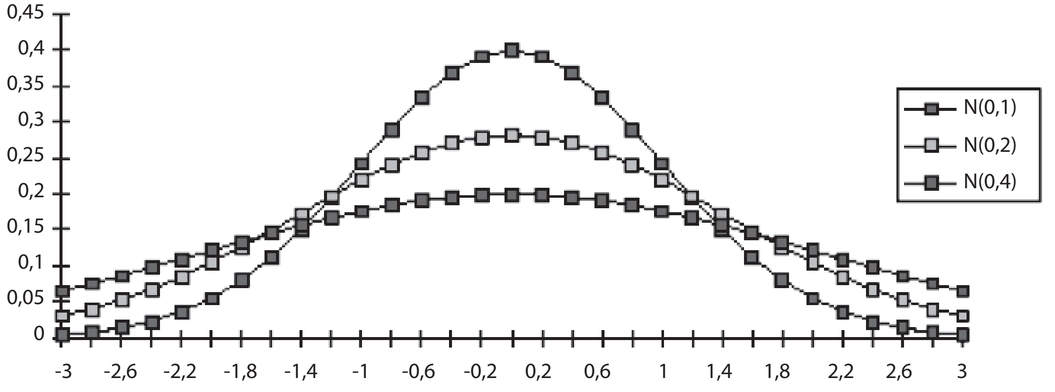

Figura 9 - Distribuições normais com média zero e variâncias 1, 2 e 4

Note que o máximo das densidades é encontrado quando x = 0, isto é, quando x é igual à média da distribuição. Isto vale para qualquer distribuição normal; o máximo de f(x) é obtido fazendo-se x = μ, onde m é a média da normal. Também, quanto maior o valor da variância σ^2, mais "espalhada" é a distribuição.

b) Propriedades da Distribuição Normal

1. f(x) dada pela expressão acima, integra a 1.
2. f(x) $^3 \geq 0$ sempre.
3. Os limites de f(x) quando x tende a $+\infty$ e $-\infty$, são iguais a zero.
4. A densidade $N(\mu, \sigma^2)$ é **simétrica em torno de** μ, ou seja,
   $f(\mu + x) = f(\mu - x)$.
5. O valor máximo de f(x) ocorre em $x = \mu$.
6. Os pontos de inflexão de f(x) são $x = \mu + \sigma$ e $x = \mu - \sigma$.

3. Momentos de uma distribuição de probabilidade

A seguir, definimos alguns dos momentos de distribuições de probabilidade contínuas e discretas. Momentos são quantidades que nos dão uma idéia da tendência central, dispersão e assimetria de uma densidade de probabilidades.

3.1 Definição (média e variância)

A média (ou valor esperado, primeiro momento de uma variável aleatória) é definida como

$$\mu = E(X) = \begin{cases} \sum_x xf(x), \text{caso discreto} \\ \int_{-\infty}^{+\infty} xf(x)dx, \text{caso contínuo} \end{cases},$$

onde o somatório refere-se a todos os valores de X quando X é uma variável discreta. Quando X é uma variável contínua a média é calculada pela integral anterior, onde f(x) representa a densidade de probabilidade da variável X.

A média de uma variável aleatória representa uma medida de tendência central da distribuição de probabilidade dessa variável aleatória.

A variância de uma variável aleatória é uma medida da dispersão da distribuição de probabilidade, definida como

$$\sigma^2 = V(X) = \begin{cases} \sum_x (x - \mu), \text{caso discreto} \\ \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx, \text{caso contínuo} \end{cases},$$

onde novamente f(x) representa a densidade de probabilidade (discreta ou contínua) da variável aleatória X e μ é a média da variável aleatória. A variância é o segundo momento em torno da média, e corresponde ao momento de inércia em Mecânica.

Da própria definição segue que a variância é uma quantidade sempre maior ou igual a zero.

3.2 Desvio padrão

O desvio padrão de uma variável aleatória é a raiz quadrada positiva da sua variância, e denotado por σ.

O desvio padrão é expresso nas mesmas unidades que a variável aleatória, e a variância é dada nas unidades da variável aleatória ao quadrado. Logo, se a variável aleatória é medida em metros, o desvio padrão também está em metros, e a variância, em metros quadrados. Um valor pequeno do desvio padrão indica que existe pouca dispersão em torno da média. Se o desvio padrão é grande, os valores da variável aleatória estão muito dispersos em torno da média.

A média e a variância são casos particulares do que chamamos de "momentos" de uma distribuição de probabilidade. Os momentos de uma distribuição servem para caracterizar esta distribuição não apenas no que se refere à sua centralidade e sua dispersão, mas também com relação a outras características, como a simetria ou assimetria da densidade de probabilidade.

3.3 Késimo momento

O *k*ésimo momento da variável aleatória *X* é definido como

$$E(X^k) = \begin{cases} \sum_x x^k f(x), \text{caso discreto} \\ \int_{-\infty}^{+\infty} x^k (x) dx, \text{caso contínuo} \end{cases}$$

Onde $k = 1, 2, 3, \ldots$

Obviamente, a definição não se aplica se algum dos $E(X^k)$ é infinito.

Definição: késimo momento central ou késimo momento em torno da média)

O késimo momento central da variável aleatória X é definido como

$$E(X^k) = \begin{cases} \sum_x (x - \mu)^k f(x), \text{caso discreto} \\ \int_{-\infty}^{+\infty} (x - \mu)^k f(x) dx, \text{caso contínuo} \end{cases},$$

onde $k = 1, 2, 3, \ldots$

Logo, a média e a variância são apenas casos particulares de momentos. A média é o primeiro momento (isto é, $\mu = E(X)$) e a variância é o segundo momento central, ou seja, $E(X - \mu)^2$.

A notação E(...) indica um valor esperado e pode ser estendida para funções mais gerais que X^k ou $(X - \mu)^k$.

3.4 Valor esperado de uma função de uma variável aleatória

Seja X uma variável aleatória com densidade f(x), e seja u(X) uma função qualquer tal que

$$E[u(X)] = \begin{cases} \sum\limits_{x} u(X)f(x), \text{caso discreto} \\ \int\limits_{-\infty}^{+\infty} u(x)dx, \text{caso contínuo} \end{cases}$$

Formula alternativa para o cálculo da variância

$$\begin{aligned} V(X) &= E[X - E(X)]^2 \\ &= E\{X^2 - 2XE(X) + [E(X)]^2\}^2 \\ &= E(X^2) - E[2XE(X)] + E\{[E(X)]^2\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

Essa fórmula é válida para qualquer variável aleatória X (contínua ou discreta), desde que a média de X seja finita.

Referências

BARBETA, P. A., REIS, M. M.,BORNIA, A. C. **Estatística para Cursos de Engenharia e Informática**, São Paulo: Atlas, 2004.

BOLFARINE, Heleno & BUSSAB, Wilton O. **Elementos de Amostragem**. 1. ed. São Paulo: Edgard Blücher, 2005.

CONOVER, W J. **Practical Nonparametric Statistics**. 3. ed. New York: John Wiley, 1998.

FELLER, W. **Teoria das probabilidades e suas aplicações**. São Paulo, Edgard Blucher, 1976

MENDENHALL, William. **Probabilidade e Estatística.** São Paulo: Campus. 1985.

MIRSHAWKA, V. **Probabilidade e Estatística para Engenharia**, São Paulo: Nobel, 1978.

MORETTIN, L. G. **Estatística Básica Probabilidade**. São Paulo: Makron Books, 1993.

SPIEGEL MR. **Estatística**. Coleção Schaum. São Paulo: Editora Afiliada, 1993.

Sobre os autores

Jorge Luiz de Castro e Silva: doutor em Ciência da Computação pela Universidade Federal de Pernambuco (2004), mestre em Ciência da Computação pela Universidade Federal do Ceará (1997), graduado em Tecnologia em Processamento de Dados pela Universidade Federal do Ceará (1978) e Bacharel em Administração pela Universidade Estadual do Ceará (1981). Professor Adjunto da Universidade Estadual do Ceará e Professor Orientador do Programa de Pós-Graduação (mestrado) em Ciências da Computação. Tem experiência na área de computação, com ênfase em Modelagem Analítica e Simulação, atuando principalmente nos seguintes temas: engenharia de tráfego de redes, tráfego auto-similar, metaheurísticas para roteamento de redes, modelagem matemática e reconhecimento de padrões em tráfego de redes.

Maria Wilda Fernandes: possui especialização em Administração de Sistemas para a Internet. Graduada em Licenciatura em Matemática pela Universidade Estadual do Ceará (1987) e Bacharel em Ciências da Computação pela Universidade Estadual do Ceará (1991). Atualmente, é servidora técnico-administrativa da Universidade Estadual do Ceará, exercendo a função de gerente da área de desenvolvimento de TI da UECE e tutora a distância do curso de Licenciatura Plena em Computação da Universidade Estadual do Ceará.

Rosa Lívia Freitas de Almeida: possui graduação em Engenharia Elétrica pela Universidade de Fortaleza (1981), especialização em Análise de Sistemas pela PUC - Rio de Janeiro (1989), mestre em Saúde Pública pela Universidade Federal do Ceará (2001). Foi professora adjunta na Faculdade Integrada do Ceará durante sete anos nos cursos de Graduação de Sistemas de Informação, Graduação em Fisioterapia, Tecnólogo em Gestão Hospitalar e nos cursos de especialização. Entre outras disciplinas, ministrou Probabilidade e Estatística e Planejamento Estatístico de Experimentos. Atua como Consultora em Tecnologia da Informação em Saúde, gerência de projetos de informática, operacionalização e suporte em Ambientes Virtuais de Aprendizagens (AVA) e tutoria em cursos de Ensino a Distância. É especialista em georeferenciamento e estatística espacial, análise de sistemas, desenvolvimento e implantação de sistema de informação.

UFRJ fax 100 ANOS 1920 | 2020
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Faculdade de Administração e Ciências Contábeis
Departamento de Biblioteconomia
CAPES
MINISTÉRIO DA EDUCAÇÃO
BRASIL GOVERNO FEDERAL
UAB UNIVERSIDADE ABERTA DO BRASIL