

Coleção UAB–UFSCar

Sistemas de Informação

Benedito Galvão Benze

Estatística aplicada a sistemas de informações



Estatística aplicada a sistemas de informações



**Reitor**

Targino de Araújo Filho

Vice-Reitor

Pedro Manoel Galetti Junior

Pró-Reitora de Graduação

Emília Freitas de Lima

**Secretária de Educação a Distância - SEaD**

Aline Maria de Medeiros Rodrigues Reali

Coordenação UAB-UFSCar

Claudia Raimundo Reyes

Daniel Mill

Denise Abreu-e-Lima

Joice Otsuka

Sandra Abib

Valéria Sperduti Lima

Coordenadora do Curso de Sistemas de Informação

Vânia Neris

UAB-UFSCar

Universidade Federal de São Carlos

Rodovia Washington Luís, km 235

13565-905 - São Carlos, SP, Brasil

Telefax (16) 3351-8420

www.uab.ufscar.br

uab@ufscar.br



EdUFSCar

Conselho Editorial

José Eduardo dos Santos

José Renato Coury

Nivaldo Nale

Paulo Reali Nunes

Oswaldo Mário Serra Truzzi (Presidente)

Secretária Executiva

Fernanda do Nascimento

EdUFSCar

Universidade Federal de São Carlos

Rodovia Washington Luís, km 235

13565-905 - São Carlos, SP, Brasil

Telefax (16) 3351-8137

www.editora.ufscar.br

edufscar@ufscar.br

Benedito Galvão Benze

Estatística aplicada a sistemas de informações

São Carlos



EdUFSCar

2009

© 2009, Benedito Galvão Benze

Concepção Pedagógica

Daniel Mill

Supervisão

Douglas Henrique Perez Pino

Equipe de Revisão Linguística

Ana Luiza Menezes Baldin

Clarissa Neves Conti

Francimeire Leme Coelho

Jorge Ialanji Filholini

Letícia Moreira Clares

Luciana Rugoni Sousa

Paula Sayuri Yanagiwara

Sara Naime Vidal Vital

Equipe de Editoração Eletrônica

Christhiano Henrique Menezes de Ávila Peres

Izís Cavalcanti

Rodrigo Rosalis da Silva

Equipe de Ilustração

Jorge Luís Alves de Oliveira

Lígia Borba Cerqueira de Oliveira

Priscila Martins de Alexandre

Capa e Projeto Gráfico

Luís Gustavo Sousa Sguissardi

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária da UFSCar

B479e	Benze, Benedito Galvão. Estatística aplicada a sistemas de informações / Benedito Galvão Benze. -- São Carlos : EdUFSCar, 2009. 232 p. -- (Coleção UAB-UFSCar). ISBN – 978-85-7600-169-0 1. Estatística matemática. 2. Pesquisa - planejamento. 3. Amostragem (Estatística). 4. Teoria das probabilidades. 5. Análise exploratória de dados. I. Título. CDD – 519.5 (20ª) CDU – 519.7
-------	---

SUMÁRIO

APRESENTAÇÃO	15
---------------------	----

UNIDADE 1: A estatística como metodologia científica

1.1 Primeiras palavras	19
1.2 Problematizando o tema	19
1.3 Texto básico para estudo	19
1.3.1 Tipos de estudos	20
1.3.2 Os objetivos, as variáveis e os dados	23
1.3.3 População e amostra	23
1.3.4 Escalas de medida, tipos de variáveis e a variação aleatória	25
1.3.5 Os instrumentos de coleta de dados	28
1.4 Considerações finais	29
1.5 Atividades de aplicação e prática	29
1.5.1 Atividades individuais	30
1.5.2 Atividades coletivas	30
1.6 Estudos complementares	30
1.6.1 Saiba mais	30
1.6.2 Referências	31

UNIDADE 2: Alguns procedimentos amostrais básicos

2.1 Primeiras palavras	35
------------------------	----

2.2	Problematizando o tema.	35
2.3	Texto básico para estudo	35
2.3.1	Sistema de referência e mecanismos de sorteio da amostra.	36
2.3.2	Principais procedimentos amostrais probabilísticos	36
2.3.2.1	Amostra aleatória simples.	37
2.3.2.2	Amostra sistemática aleatória	39
2.3.2.3	Amostra aleatória estratificada	41
2.3.2.4	Amostra aleatória por conglomerados	44
2.3.3	Considerações adicionais	45
2.4	Considerações finais	46
2.5	Atividades de aplicação e prática	46
2.5.1	Atividades individuais	46
2.5.2	Atividades coletivas.	46
2.6	Estudos complementares.	46
2.6.1	Saiba mais.	47
2.6.2	Referências	47

UNIDADE 3: Representação gráfica de dados

3.1	Primeiras palavras.	51
3.2	Problematizando o tema.	51
3.3	Elaboração de planilhas de dados	51
3.3.1	Tabelas-resumo.	53
3.3.1.1	Distribuições de frequências para dados contínuos.	56
3.3.2	Técnicas gráficas de análise descritiva	58

3.3.2.1	O gráfico de barras	.59
3.3.2.2	O diagrama circular	.59
3.3.2.3	O esquema de ramos e folhas	.61
3.3.2.4	O diagrama de pontos (ou de dispersão)	.62
3.3.2.5	O histograma	.63
3.3.2.6	O gráfico de tendência	.64
3.4	Considerações finais	.65
3.5	Atividades de aplicação e prática	.65
3.5.1	Atividades individuais	.66
3.5.2	Atividades coletivas	.66
3.6	Estudos complementares	.66
3.6.1	Saiba mais	.66
3.6.2	Referências	.66

UNIDADE 4: Análise estatística unidimensional

4.1	Primeiras palavras	.71
4.2	Problematizando o tema	.71
4.3	Texto básico para estudo	.71
4.3.1	As medidas de centralidade (ou de posição)	.71
4.3.2	As medidas de variabilidade (ou de dispersão)	.76
4.3.2.1	A variância	.77
4.3.2.2	O desvio padrão	.78
4.3.2.3	Reescrevendo as expressões da média e da variância	.79
4.3.3	Medida da assimetria de uma distribuição	.79
4.3.4	Escores padronizados	.80
4.3.5	O coeficiente de variação	.82
4.3.6	As separatrizes	.84

4.3.6.1 Os quartis	84
4.3.6.2 Outras separatrizes	85
4.3.7 O desenho esquemático (box-plot)	86
4.4 Considerações finais	90
4.5 Atividades de aplicação e prática	90
4.5.1 Atividades individuais	90
4.5.2 Atividades coletivas	90
4.6 Estudos complementares	90
4.6.1 Saiba mais	91
4.6.2 Referências	91

UNIDADE 5: Análise bidimensional conjunta

5.1 Primeiras palavras	95
5.2 Problematizando o tema	95
5.3 Texto básico para estudo	95
5.3.1 Dependência entre duas variáveis quantitativas	96
5.3.2 Duas variáveis qualitativas	100
5.3.3 Distribuição conjunta de frequências relativas ou proporções	101
5.3.4 Não associação entre duas variáveis qualitativas	103
5.4 Considerações finais	106
5.5 Atividades de aplicação e prática	106
5.5.1 Atividades individuais	106
5.5.2 Atividades coletivas	106
5.6 Estudos complementares	106

5.6.1 Saiba mais.	106
5.6.2 Referências.	107

UNIDADE 6: Introdução à Teoria de Probabilidades

6.1 Primeiras palavras.	111
6.2 Problematizando o tema.	111
6.3 Texto básico para estudo.	111
6.3.1 Definições de probabilidade.	112
6.3.1.1 Definição clássica de probabilidade.	112
6.3.1.2 Definição geométrica de probabilidade.	113
6.3.1.3 Definição frequentista de probabilidade.	115
6.3.2 Teoria axiomática de probabilidade.	116
6.3.2.1 O espaço amostral.	117
6.3.2.2 Os eventos aleatórios.	119
6.3.2.3 Outros eventos complexos.	123
6.3.2.4 Leis de Morgan.	123
6.3.3 Definição axiomática de Kolmogorov.	124
6.3.4 Consequências da definição axiomática.	126
6.3.5 Probabilidade condicional.	130
6.3.6 Regra do produto.	133
6.3.7 Regras da probabilidade total e de Bayes.	136
6.3.7.1 Regra da probabilidade total.	136
6.3.7.2 Regra de Bayes.	137
6.3.8 Independência de dois eventos.	138
6.3.8.1 Independência aos pares e independência coletiva.	140
6.4 Considerações finais.	143
6.5 Atividades de aplicação e prática.	143

6.5.1 Atividades individuais	143
6.5.2 Atividades coletivas	143
6.6 Estudos complementares.....	143
6.6.1 Saiba mais.....	143
6.6.2 Referências	144

UNIDADE 7: Variáveis aleatórias discretas

7.1 Primeiras palavras.....	147
7.2 Problematizando o tema.....	147
7.3 Texto básico para estudo	147
7.3.1 Variáveis aleatórias	147
7.3.2 Técnicas de contagem.....	149
7.3.2.1 O princípio da multiplicação	149
7.3.2.2 O número de permutações	151
7.3.2.3 O número de arranjos de n elementos tomados k a k	152
7.3.2.4 O número de combinações de n elementos tomados k a k	153
7.3.3 Distribuição de probabilidades de uma variável aleatória discreta	153
7.3.4 Determinação de probabilidades em intervalos.....	157
7.3.5 Independência entre variáveis aleatórias	159
7.3.6 Esperança e variância de variáveis aleatórias discretas.....	160
7.3.6.1 A esperança de variáveis aleatórias discretas	160
7.3.6.2 A variância de variáveis aleatórias discretas	161
7.3.7 Algumas distribuições discretas básicas.....	166
7.3.7.1 Distribuição de Bernoulli	166
7.3.7.2 Distribuição binomial.....	167
7.3.7.3 Distribuição de Poisson.....	170
7.4 Considerações finais.....	173

7.5 Atividades de aplicação e prática	173
7.5.1 Atividades individuais	173
7.5.2 Atividades coletivas	174
7.6 Estudos complementares	174
7.6.1 Saiba mais	174
7.6.2 Referências	174

UNIDADE 8: Variáveis aleatórias contínuas

8.1 Primeiras palavras	177
8.2 Problematizando o tema	177
8.3 Texto básico para estudo	177
8.3.1 Distribuição de probabilidades de uma v.a. contínua	177
8.3.1.1 Função densidade de probabilidade	177
8.3.1.2 Função distribuição acumulada de uma v.a. contínua	179
8.3.2 Determinação de probabilidades em intervalos	180
8.3.3 A distribuição uniforme	181
8.3.4 Independência entre variáveis aleatórias contínuas	181
8.3.5 Esperança e variância de variáveis aleatórias contínuas	183
8.3.5.1 A esperança de variáveis aleatórias contínuas	183
8.3.5.2 A variância de variáveis aleatórias contínuas	183
8.3.6 Relação entre a distribuição exponencial e a de Poisson	185
8.3.7 A distribuição normal	186
8.3.7.1 Cálculo de probabilidades na distribuição normal	188
8.3.7.2 Aproximação da distribuição binomial pela normal	191
8.3.8 Outras distribuições contínuas importantes	192
8.3.8.1 A distribuição gama e relacionadas	192

8.4	Considerações finais	193
8.5	Atividades de aplicação e prática	193
8.5.1	Atividades individuais	193
8.5.2	Atividades coletivas	193
8.6	Estudos complementares	193
8.6.1	Saiba mais	193
8.6.2	Referências	194

UNIDADE 9: Intervalos de confiança

9.1	Primeiras palavras	197
9.2	Problematizando o tema	197
9.3	Texto básico para estudo	197
9.3.1	Intervalos de confiança para a média populacional	197
9.3.2	Intervalos de confiança para a proporção populacional	204
9.4	Considerações finais	206
9.5	Atividades de aplicação e prática	206
9.5.1	Atividades individuais	206
9.5.2	Atividades coletivas	207
9.6	Estudos complementares	207
9.6.1	Saiba mais	207
9.6.2	Referências	207

UNIDADE 10: Noções gerais sobre testes de hipóteses

10.1	Primeiras palavras	211
10.2	Problematizando o tema	211
10.3	Texto básico para estudo	211
10.3.1	O raciocínio básico de um teste de hipóteses	211
10.3.2	Testes unilaterais	216
10.3.3	Testes bilaterais	217
10.3.4	O poder de um teste	217
10.3.5	Alguns testes de hipóteses específicos	218
10.3.5.1	Comparação das médias de dois grupos dependentes	218
10.3.5.2	Comparação das médias de dois grupos independentes	221
10.3.5.3	Comparação de duas proporções em amostras independentes	224
10.3.6	Intervalos de confiança para diferenças de parâmetros	226
10.3.7	Análise de variância	227
10.3.8	Outras aplicações dos testes de hipóteses	227
10.4	Considerações finais	228
10.5	Atividades de aplicação e prática	228
10.5.1	Atividades individuais	228
10.5.2	Atividades coletivas	228
10.6	Estudos complementares	228
10.6.1	Saiba mais	229
10.6.2	Referências	229

APRESENTAÇÃO

Com este texto você se iniciará nos chamados métodos estatísticos. Eles são essenciais para qualquer área, como a de Sistemas de Informação, em que o conhecimento é adquirido e o processo de tomada de decisões é realizado com base em alguma informação. Um aspecto importante é o planejamento das pesquisas, antes da coleta de dados. Isso pode evitar ou minimizar erros que, de alguma forma, venham a comprometer os resultados deste processo. Também, os métodos de resumo da informação facilitam a sua análise e agilizam a tomada de decisões. Com o desenvolvimento do raciocínio probabilístico, você, em sua atual ou futura profissão, poderá melhorar a qualidade do seu trabalho considerando em suas inferências a atuação da variabilidade dos dados. Não se pode ignorar que, dependendo da quantidade de dados, a maioria dos procedimentos requer algum cálculo, que realizado manualmente pode desestimular o iniciante na aplicação da Estatística em suas atividades. Felizmente, isso pode ser contornado pelo uso de sistemas e procedimentos computacionais, cada vez mais acessíveis.