Volume 2
2ª edição

Giovani Glaucio de Oliveira Costa
Juliana Di Giorgio Giannotti

Estatística Aplicada ao Turismo

Fundação
CECIERJ
Consórcio cederj
Centro de Educação Superior a Distância do Estado do Rio de Janeiro

Volume 2
2ª Edição

Giovani Glaucio de Oliveira Costa
Juliana Di Giorgio Giannotti

Estatística Aplicada ao Turismo

GOVERNO DO
Rio de Janeiro
SECRETARIA DE
CIÊNCIA E TECNOLOGIA

UAB
UNIVERSIDADE ABERTA DO BRASIL

Ministério
da Educação

BRASIL
UM PAÍS DE TODOS
GOVERNO FEDERAL

Apoio:

FAPERJ
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

Fundação Cecierj / Consórcio Cederj

Rua Visconde de Niterói, 1364 – Mangueira – Rio de Janeiro, RJ – CEP 20943-001
Tel.: (21) 2334-1569 Fax: (21) 2568-0725

Material Didático

ELABORAÇÃO DE CONTEÚDO
Giovani Glaucio de Oliveira Costa
Juliana Di Giorgio Giannotti

COORDENAÇÃO DE DESENVOLVIMENTO INSTRUCIONAL
Cristine Costa Barreto

SUPERVISÃO DE DESENVOLVIMENTO INSTRUCIONAL
Ana Paula Abreu-Fialho

DESENVOLVIMENTO INSTRUCIONAL E REVISÃO
Gustavo de Figueiredo Tarcsay
Luiz Eduardo S. Feres
Marcelo Bastos Matos
Wilson P. de Oliveira Junior

Departamento de Produção

EDITORA
Tereza Queiroz

REVISÃO TIPOGRÁFICA
Cristina Freixinho
Daniela de Souza
Diana Castellani
Elaine Bayma
Patrícia Paula

COORDENAÇÃO DE PRODUÇÃO
Katy Araújo

PROGRAMAÇÃO VISUAL
Márcia Valéria de Almeida

ILUSTRAÇÃO
André Dahmer

CAPA
André Dahmer

PRODUÇÃO GRÁFICA
Oséias Ferraz
Patricia Seabra

C837e

Costa, Giovani Glaucio de Oliveira.
Estatística Aplicada ao Turismo. v.2 / Giovani Glaucio de Oliveira Costa, Juliana Di
Giorgio Giannotti. – 2.ed. – Rio de Janeiro: Fundação CECIERJ, 2010.
136 p.; 19 x 26,5 cm.

ISBN: 978-85-7648-624-4

1. Estatística. 2. Probalidades. 3. Distribuição binomial.
4. Variáveis aleatórias. I. Giannotti, Juliana Di Giorgio. II. Título.

CDD 519.5

2010/1

Referências Bibliográficas e catalogação na fonte, de acordo com as normas da ABNT.

Governo do Estado do Rio de Janeiro

Governador
Sérgio Cabral Filho

Secretário de Estado de Ciência e Tecnologia
Alexandre Cardoso

Universidades Consorciadas

UENF - UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO
Reitor: Almy Junior Cordeiro de Carvalho

UERJ - UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO
Reitor: Ricardo Vieiralves

UFF - UNIVERSIDADE FEDERAL FLUMINENSE
Reitor: Roberto de Souza Salles

UFRJ - UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Reitor: Aloísio Teixeira

UFRRJ - UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
Reitor: Ricardo Motta Miranda

UNIRIO - UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO
Reitora: Malvina Tania Tuttman

Estatística Aplicada ao Turismo | Volume 2

SUMÁRIO

13 Noções de probabilidades – conceitos básicos

Meta da aula

Apresentar os conceitos básicos que envolvem o cálculo das probabilidades.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. identificar experimento aleatório;
2. construir espaços amostrais para diferentes experimentos aleatórios;
3. construir eventos, seus tipos, associações e realizar operações com eventos para diferentes experimentos aleatórios.

Pré-requisitos

Você deve ter consolidado os conceitos de: distribuição de frequências, tabelas e gráficos, medidas de posição, tendência central, dispersão e da forma de uma distribuição presentes nas Aulas 5, 6, 7, 8, 9, 10, 11 e 12, respectivamente.

Introdução

No estudo da Estatística, os elementos de incerteza fazem parte, muitas vezes, dos problemas a serem solucionados. Isso se dá pois, geralmente não é possível determinar de antemão as características de uma população desconhecida ou mesmo antecipar as conseqüências reais de uma tomada de decisão.

Por exemplo, muitas vezes ao nos prepararmos para ir ao trabalho, ou à escola pela manhã, temos de decidir se vestimos ou não uma blusa de frio. O que fazemos para decidir sobre esse ponto é julgar se irá ou não fazer frio naquele dia. Desse modo, estamos impondo uma "chance" à ocorrência de frio e, por meio dessa imposição, decidimos se a blusa vai ou não conosco ao trabalho ou à escola.

Em nosso cotidiano, nos deparamos com outras inúmeras situações nas quais não temos certeza sobre qual escolha tomar. Assim, para ajudar nessas decisões, dispomos de uma medida que irá exprimir numericamente essa incerteza. A essa medida damos o nome de probabilidade.

Vamos seguir o nosso estudo com algumas definições importantes.

Probabilidades

Campo do conhecimento que estuda os fenômenos ou experimentos aleatórios.

Experimentos aleatórios

São aqueles cujos resultados nem sempre são os mesmos, apresentam variações, mesmo quando repetidos indefinidamente em condições uniformes.

Exemplos:

- a experiência que consiste no lançamento de uma moeda é um fenômeno aleatório;
- a experiência que consiste no lançamento de um dado é um fenômeno aleatório;

- sorteamos uma pessoa e pedimos para ela indicar, pelo sabor (sem visualizar a marca do produto que está consumindo) a sua preferência. A preferência entre duas marcas de refrigerantes é um fenômeno aleatório.

Podemos notar por esses exemplos que, quando temos um experimento aleatório não podemos prever com certeza qual será o seu resultado. Mas podemos, sim, descrever todos os possíveis resultados desse experimento.

Atividade

Atende ao Objetivo 1

1. Seja a situação: selecionamos um cliente de uma agência de turismo e anotamos a sua opinião quanto à qualidade do roteiro da última viagem que realizou pela agência. Essa situação configura um experimento aleatório?

Resposta Comentada

Sim, pois não sabemos ao certo qual será a resposta, ou seja, a opinião que o cliente irá indicar com relação à qualidade do roteiro turístico da última viagem que realizou pela agência.

Espaço amostral (S)

É o conjunto de todos os resultados possíveis de um experimento aleatório.

Exemplos:

Vamos construir os espaços amostrais dos exemplos de experimento aleatório dados anteriormente:

- Seja a experiência que consiste no lançamento de uma moeda, o espaço amostral associado é:

{Cara, Coroa}.

- Seja a experiência que consiste no lançamento de um dado, o espaço amostral associado é:

{1, 2, 3, 4, 5, 6}.

- Quando um cliente é selecionado para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico que realizou pela agência de turismo, o espaço amostral associado é:

{Ótimo, bom, regular, ruim, péssimo}.

- Quando uma pessoa é sorteada para escolher, quando de olhos vendados, o sabor, entre duas marcas, A ou B, concorrentes de um refrigerante, o espaço amostral associado é:

{Marca A, Marca B}.

Evento (E)

É todo subconjunto finito de um espaço amostral. É um resultado ou o conjunto de resultados de interesse em uma experiência aleatória.

Exemplos:

- Seja a experiência que consiste no lançamento de uma moeda, podemos ter os seguintes eventos de interesse:

E_1= {Cara}.

E_2= {Coroa}.

- Seja a experiência que consiste no lançamento de um dado, podemos ter os seguintes eventos de interesse:

E_1={1}.

E_2={2}.

E_3={3}.

E_4={4}.

- Quando um cliente é selecionado para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico que realizou pela agência de turismo, podemos ter os seguintes eventos de interesse:

$$E_1= \{\text{Ótimo}\}.$$

$$E_2= \{\text{Bom}\}.$$

$$E_3= \{\text{Ruim}\}.$$

Tipos de eventos

Vamos estudar os quatro tipos de eventos seguintes:

a. Evento simples.

b. Evento composto.

c. Evento certo.

d. Evento impossível.

a. Evento simples

É o evento formado por um único elemento do espaço amostral associado.

Exemplos:

- Seja a experiência que consiste no lançamento de uma moeda, podemos ter os seguintes eventos simples de interesse:

$$E_1= \{\text{Cara}\}.$$

$$E_2= \{\text{Coroa}\}.$$

- Seja a experiência que consiste no lançamento de um dado, podemos ter os seguintes eventos simples de interesse:

$$E_1=\{1\}.$$

$$E_2= \{2\}.$$

$$E_3=\{3\}.$$

$$E_4=\{4\}.$$

$$E_5= \{5\}.$$

$$E_6= \{6\}.$$

- Quando uma pessoa é sorteada para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico, podemos ter os seguintes eventos simples de interesse:

$$E_1= \{\text{Ótimo}\}.$$

$$E_2= \{\text{Bom}\}.$$

b. Evento composto

É o evento formado por dois ou mais elementos do espaço amostral S associado.

Exemplos:

- Seja a experiência que consiste em vários lançamentos de uma moeda, podemos ter o seguinte evento composto de interesse:

$$E= \{\text{Cara, Coroa}\}.$$

- Seja a experiência que consiste em vários lançamentos de um dado, podemos ter os seguintes eventos compostos de interesse:

$$E_1= \{1, 2\};$$

$$E_2= \{3, 4, 5\};$$

$$E_3= \{1, 2, 3, 4, 5, 6\}.$$

- Quando pessoas são sorteadas para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico, podemos ter o seguinte evento composto de interesse:

$$E= \{\text{Ótimo, bom}\}.$$

c. Evento certo (C)

É aquele que sempre ocorre em qualquer realização da experiência aleatória. É aquele que coincide com o próprio espaço amostral. Consequentemente, a probabilidade de ocorrer o evento certo é sempre *P(C) = 1 ou P(C) = 100%*, isto é, a certeza.

Exemplos:

- Seja a experiência que consiste no lançamento de uma moeda, o evento certo associado é:

$$C = \{\text{Cara, Coroa}\} \rightarrow P(C) = 1.$$

- Seja a experiência que consiste no lançamento de um dado, o evento certo associado é:

$$C = \{1, 2, 3, 4, 5, 6\} \rightarrow P(C) = 1.$$

- Quando uma pessoa é sorteada para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico, o evento certo associado é:

$$C = \{\text{Ótimo, bom, regular, ruim, péssimo}\} \rightarrow P(C) = 1.$$

d. Evento impossível (I)

É aquele que nunca ocorre, em nenhuma realização do experimento aleatório. Assim, sempre o evento impossível $I = \varnothing$ (conjunto vazio). A probabilidade de um evento impossível é sempre igual a zero, isto é, $P(I) = 0$.

Exemplos:

- Seja a experiência que consiste no lançamento de um dado, um evento impossível associado é:

$$I = \{\text{Face} > 6\} \rightarrow I = \varnothing \rightarrow P(I) = 0.$$

- Quando uma pessoa é sorteada para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico, um evento Impossível associado é:

$$I = \{\text{Outra}\} \rightarrow I = \varnothing \rightarrow P(I) = 0.$$

Associações e operações entre eventos

Vamos estudar as seguintes associações e operações entre eventos:

1. União de dois eventos.
2. Interseção de dois eventos.
3. Eventos mutuamente exclusivos.

4. Eventos complementares.
5. Eventos independentes.

União de dois eventos

Sejam E_1 e E_2 dois eventos de um mesmo espaço amostral (S), a união de E_1 e E_2 é representada por $E_1 \cup E_2$ e é formada pelo conjunto de todos os elementos amostrais que estão em E_1, em E_2 ou em ambos. Esquematicamente, temos:

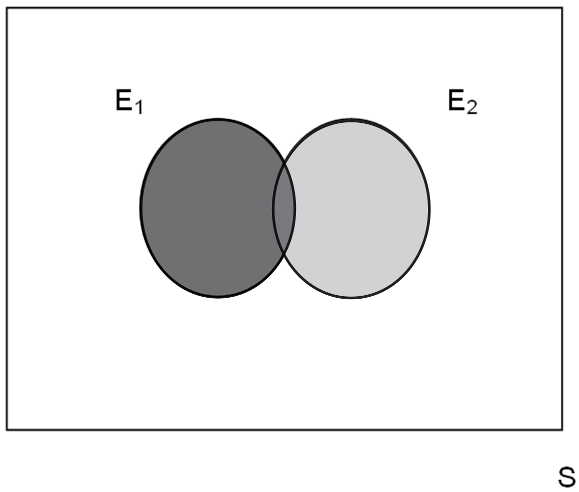


Figura 13.1: Esquema da união entre dois eventos.

Interseção de dois eventos

Sejam E_1 e E_2 dois eventos de um mesmo espaço amostral (S), a interseção de E_1 e E_2 é representada por $E_1 \cap E_2$ e é formada pelo conjunto de todos os elementos amostrais que estão em E_1 e em E_2, isto é, que são comuns aos dois eventos. Esquematicamente, temos:

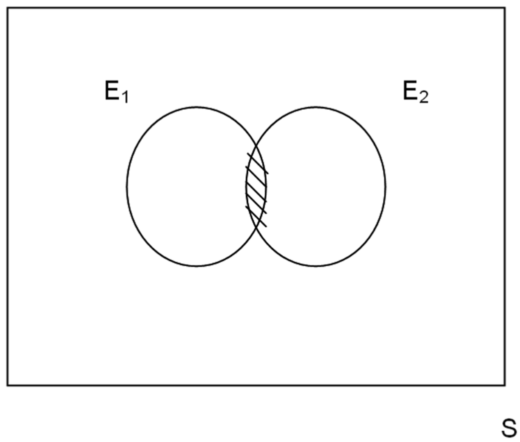


Figura 13.2: Esquema da interseção entre dois eventos.

Atividade

Atende ao Objetivo 3

2. Seja o experimento aleatório lançamento de um dado. Sejam os eventos E_1= *ocorrer número par* e E_2= *ocorrer número menor que três*. Pede-se:

a. Qual a união entre E_1 e E_2?

b. Qual a interseção entre E_1 e E_2?

Resposta Comentada

Primeiramente, deve-se realizar a construção dos eventos: $E_1 = \{2, 4, 6\}$ e $E_2 = \{1, 2\}$. Agora podemos realizar as operações entre os dois eventos em questão, que resultarão em:

a. A união entre E_1 e E_2 será $E_1 \cup E_2 = \{1, 2, 4, 6\}$, pois compreende os elementos que estão em E1 ou em E2.

b. A interseção entre E_1 e E_2 será $E_1 \cap E_2 = \{2\}$, pois compreende os elementos que estão simultaneamente em E_1 e em E_2.

Eventos mutuamente exclusivos

São aqueles que nunca podem ocorrer simultaneamente em uma mesma realização de uma experiência aleatória.

Exemplos:

- No lançamento de uma moeda, os eventos cara e coroa são mutuamente exclusivos.
- No lançamento de um dado, os eventos 1 e 4 são mutuamente exclusivos.

Lembrando da *teoria dos conjuntos*, podemos dizer que eventos mutuamente exclusivos constituem conjuntos disjuntos, isto é, a interseção é o conjunto vazio (representado por $\varnothing$). Sejam E1 e E2 dois eventos de um mesmo espaço amostral (S), assim, $E= E1 \cap E2= \varnothing$.

Atividade

Atende ao Objetivo 3

3. Seja o experimento aleatório jogar um dado e observar o resultado. Construa dois eventos mutuamente exclusivos para esse experimento.

Resposta Comentada

Sejam os eventos:

A = ocorrer número par $\Rightarrow A = \{2, 4, 6\}$, *e*

B = ocorrer número ímpar $\Rightarrow B = \{1, 3, 5\}$.

Podemos observar que os eventos A e B são mutuamente exclusivos, pois $A \cap B = \varnothing$.

Eventos complementares

Sabemos que um evento pode ocorrer ou não ocorrer. Seja A um evento qualquer de um espaço amostral (S). O complemento de A é o evento que ocorre quando A não ocorre, e a união de A e o seu complemento será sempre igual a S. O complemento de A ou A complementar é representado por A^c ou $\overline{A}$. Esquematicamente temos:

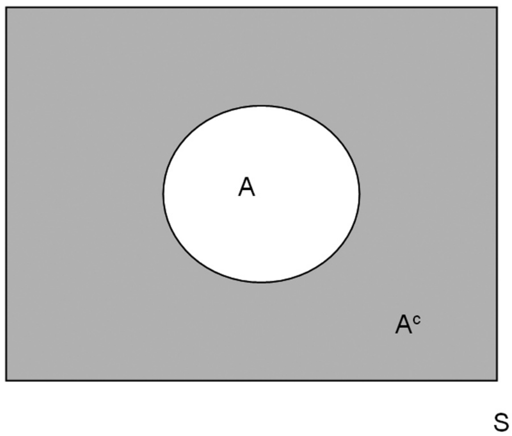


Figura 13.3: Esquema de eventos complementares.

Exemplo:

Considere o experimento aleatório *lançamento de um dado e o evento A ocorrer número maior que quatro, A = {5, 6}*. O evento complementar de A é dado pelo conjunto formado pelos elementos A^c= {1, 2, 3, 4}.

A diferença entre eventos complementares e eventos mutuamente exclusivos é que a união de eventos complementares é sempre o espaço amostral e a união de eventos mutuamente exclusivos nem sempre será o espaço amostral. Por exemplo, seja a experiência aleatória que consiste no lançamento de um dado. O espaço associado a esta experiência é S= {1, 2, 3, 4, 5, 6}. Sejam dois eventos deste espaço amostral: E_1= {1, 2, 3} e E_2= {4, 5}. Esses eventos são mutuamente exclusivos, mas não são eventos complementares, pois a união deles *não* resulta no espaço amostral.

Atividade

Atende ao Objetivo 3

4. Considere o experimento aleatório lançar duas moedas e observar a face voltada para cima e o evento A = "a face da primeira moeda é Cara". Qual é o evento complementar de A?

Resposta Comentada

O espaço amostral desse experimento é:
S = {(Cara, Cara), (Cara, Coroa), (Coroa, Cara), (Coroa, Coroa)}.
O evento A é dado por A = {(Cara, Cara), (Cara, Coroa)}. Assim, o seu complemento será: A^c *= {(Coroa, Cara), (Coroa, Coroa)}, pois* A^c *é o evento que ocorre quando A não ocorre e* $A \cup A^c = S$ *= {(Cara, Cara), (Cara, Coroa), (Coroa, Cara), (Coroa, Coroa)}.*

Eventos independentes

Dizemos que dois eventos são independentes quando a realização ou não realização de um dos eventos não afeta a probabilidade da realização do outro e vice-versa. A ocorrência de um deles, não aumenta ou diminui a ocorrência do outro. A realização de um deles não modifica a chance de realização do outro.

Exemplo:

Quando lançamos dois dados, o resultado obtido em um deles não afeta o resultado obtido no outro. O resultado que porventura deu num dos dados não afeta e nem é afetado pelo resultado que deu no outro dado. Os resultados são independentes.

Nesta aula você aprendeu os conceitos básicos de teoria dos conjuntos que auxiliarão no entendimento da teoria das probabilidades, que será vista nas aulas posteriores com maior aprofundamento.

Atividades Finais

Atendem aos Objetivos 1, 2 e 3

1. Sendo A e B dois eventos em um mesmo espaço amostral, "traduza" para a linguagem da teoria dos conjuntos as seguintes situações:

a. pelo menos um dos eventos ocorre;

b. o evento A ocorre, mas B não;

c. nenhum deles ocorre.

2. Num grupo de oito crianças em um hotel onde há monitores para conduzir brincadeiras e jogos, o conjunto da idade dessas crianças forma o seguinte espaço amostral: S= {5, 6, 7, 8, 9, 10, 11, 12}. Considere os seguintes eventos:

- A= {5, 12}.
- B= {5, 6, 7, 8}.
- C= {9, 10, 11, 12}.
- D= {9}.
- E= {5, 6, 11, 12}.

e encontre:

a. $A \cup B$; $A \cup C$; $D \cup C$.

b. $A \cap B$; $A \cap C$; $D \cap C$.

c. A^c; B^c; C^c.

d. $E \cup D \cup B$.

e. $A \cap C \cap E$.

Respostas Comentadas

1.

a. Nesta situação, pode ocorrer o evento A, o evento B ou os dois simultaneamente. Na linguagem dos conjuntos $A \cup B$ quer dizer "pelo menos um dos eventos ocorre".

b. Nesta situação significa que o evento A ocorreu, mas o B não ocorreu. Na linguagem dos conjuntos $A \cap B^c$ quer dizer "o evento A ocorre mas B não".

c. Nesta situação significa que tanto o evento A quanto o B não ocorreram. Na linguagem dos conjuntos $A^c \cap B^c$ quer dizer "nenhum deles ocorre".

2. O espaço amostra (S) já foi fornecido no enunciado da questão, assim como os eventos de interesse. Desse modo, deve-se realizar as operações que foram solicitadas da seguinte maneira:

a. $A \cup B = \{5, 6, 7, 8, 12\}$; pois compreende os elementos que estão em A ou em B.

$A \cup C = \{5, 9, 10, 11, 12\}$; pois compreende os elementos que estão em A ou em C.

$D \cup C = \{9, 10, 11, 12\}$; pois compreende os elementos que estão em D ou em C.

b. $A \cap B = \{5\}$; pois compreende os elementos que estão em A e em B, ao mesmo tempo.

$A \cap C = \{12\}$; pois compreende os elementos que estão em A e em C, ao mesmo tempo.

$D \cap C = \{9\}$, pois compreende os elementos que estão em D e em C, ao mesmo tempo.

c. $A^c = \{6, 7, 8, 9, 10, 11\}$; pois A^c é o evento que ocorre quando A não ocorre e $A \cup Ac = S$.

$B^c = \{9, 10, 11, 12\}$; pois B^c é o evento que ocorre quando B não ocorre e $B \cup B^c = S$.

C^c = {5, 6, 7, 8}, pois C^c é o evento que ocorre quando C não ocorre e $C \cup C^c = S$.

d. $E \cup D \cup B$ = {5, 6, 7, 8, 9, 11, 12}, pois compreende os elementos que estão em E ou em D ou em B.

e. $A \cap C \cap E$ = {12}, pois compreende os elementos que estão em A e em C e em E, ao mesmo tempo.

Resumo

O campo do conhecimento científico que estuda os fenômenos aleatórios ou as experiências aleatórias é o cálculo das probabilidades.

Experimentos aleatórios são aqueles em que seus resultados ou conjunto de resultados variam de forma imprevisível, ao acaso.

Eventos são resultados ou conjuntos de resultados de uma experiência aleatória. Temos eventos simples, composto, certo e impossível.

Os tipos de associações de ventos são união de eventos, interseção de eventos, eventos mutuamente exclusivos, eventos complementares e eventos independentes.

Informações sobre a próxima aula

Na próxima aula estudaremos mais conceitos básicos de probabilidades, tais como o conceito de probabilidade em função da noção de eventos, definição clássica de probabilidades, os axiomas do cálculo das probabilidades e a regra do produto.

14 Noções básicas – probabilidade

Metas da aula

Apresentar as definições frequentista e clássica de probabilidade e estudar os axiomas e os teoremas do cálculo das probabilidades.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. diferenciar as definições de probabilidades frequencial e clássica;
2. aplicar nas situações práticas os principais axiomas do cálculo das probabilidades;
3. calcular probabilidades em diferentes problemas com dados reais.

Pré-requisitos

Para o melhor acompanhamento desta aula, você deve ter aprendido os tipos, as associações e as operações com eventos da teoria dos conjuntos, estudados na Aula 13, assim como a teoria e a prática presentes na amostragem e na estatística descritivas, estudadas no decorrer do nosso curso.

Introdução

Na Aula 13, introduzimos o conceito de probabilidade por meio de exemplos e uma definição genérica, e realizamos um estudo mais aprofundado sobre espaço amostral, experimento aleatório e os tipos de eventos, suas associações e as possíveis operações entre eles.

Com esta base adquirida na aula anterior, vamos dar sequência ao nosso curso aprendendo a associar uma medida numérica à "chance" dos eventos ocorrerem. Assim, vamos, nesta aula, definir formalmente e calcular probabilidades.

Seguem as definições.

Conceito de probabilidades em função da noção de eventos

É uma medida numérica, em termos relativos ou percentuais, que expressa a chance de um evento de interesse ocorrer. É a quantificação de incertezas.

Exemplo:

Seja a experiência que consiste no lançamento de uma moeda. A medida numérica que expressa a chance de ocorrer o evento "cara", em um dado lançamento, é 50%.

Definição frequencial (intuitiva) de probabilidades – *a posteriori*

A posteriori

Vem do latim e significa "partindo daquilo que vem depois"; refere-se ao que é conhecido por meio de experiência.

Trata-se da probabilidade avaliada, empírica que tem por objetivo estabelecer um modelo adequado à interpretação de uma certa classe de fenômenos observados (não todos) em que a experiência é a base para se montar o modelo ou para ajustá-lo ao modelo ideal (teórico).

Exemplo:

Consideremos um grupo de máquinas de uma fábrica, operadas de uma certa forma, tendo uma determinada capacidade de produção. Vamos caracterizar a qualidade do produto manufaturado por essas máquinas, com um critério preestabelecido para se decidir se a peça produzida é perfeita (P) ou defeituosa (D). Tomemos seis amostras de peças produzidas pelas máquinas, sendo cada amostra constituída de 25 peças. Após a análise de qualidade, contemos as peças defeituosas e calculemos a sua porcentagem para cada amostra. Repitamos a experiência, mas aumentando o tamanho da amostra para 250 peças, inicialmente, e depois para 2.500 peças. Suponhamos que tenhamos encontrado os seguintes valores anotados na **Tabela 14.1**.

Tabela 14.1: Número de peças em cada amostra

Nº de peças tomadas para amostra (n)					
N= 25		n= 250		n= 2.500	
D	%D	D	%D	D	%D
4	16	12	4,8	157	6,28
1	4	14	5,6	151	6,08
0	0	22	8,8	136	5,44
2	8	15	6,0	160	6,40
1	4	8	3,3	153	6,12
0	0	15	6,0	157	6,28

Fonte: Dados hipotéticos.
Onde: %D = (D/n) .100.

Notemos que, em cada caso, as quantidades de peças defeituosas encontradas constituem as frequências absolutas, enquanto as porcentagens de peças defeituosas constituem as frequências relativas. Verificamos que, quando o tamanho da amostra é pequeno, as frequências relativas apresentam oscilações irregulares grandes, porém, à medida que o tamanho da amostra cresce, as oscilações tendem a ser menores e elas oscilam em torno de um valor constante hipotético.

Assim, para amostras suficientemente grandes, as frequências relativas pouco diferem entre si. É o que chamamos de *Regularidade Estatística dos Resultados*.

O valor hipotético fixo, no qual tende a haver uma estabilização da frequência relativa, denomina-se probabilidade. No exemplo, seria a probabilidade de ocorrência de peças defeituosas daquele grupo de máquinas. A frequência relativa é, portanto, considerada uma medição experimental do valor da probabilidade.

Diríamos:

$$P(E) = \lim_{n \to \infty} [F(E)/n]$$

em que:

P(E)= probabilidade de ocorrer o evento E;

F(E)= frequência absoluta do evento E;

n= tamanho da amostra.

Do ponto de vista matemático, essa definição de probabilidade apresenta dificuldades, porque um número limite real pode não existir. Assim, a formalização da definição não obedece rigorosamente à teoria matemática de limite. Isso traz, como consequência, dificuldades em demonstrar os teoremas de probabilidades, muito embora essa definição seja bastante intuitiva.

A denominação *a posteriori* resulta do fato de termos de repetir a experiência várias vezes para podermos calcular a probabilidade.

Definição matemática de probabilidades *a priori*

A priori

Vem do latim e significa "partindo daquilo que vem antes", a princípio.

Seja uma experiência aleatória onde todos os elementos de um espaço amostral S associado a uma experiência aleatória tenham a mesma chance de ocorrer, e seja E um evento de interesse do espaço amostral S, então a probabilidade de ocorrer o evento E pode ser assim definida:

$$\boxed{P(E) = \frac{n(E)}{n(S)}}$$

onde,

n(E) é o número de elementos do evento de interesse (E), e

n(S) é o número de elementos do espaço amostral (S).

Exemplos:

a. Uma pessoa tem três notas de R$ 2,00 e uma nota de R$ 5,00 no bolso. Essa pessoa entra apressadamente no ônibus e retira uma nota do seu bolso. Qual a probabilidade de ter retirado uma nota de R$ 2,00?

E= retirar uma nota de R$2,00 do bolso

n(E)= 3

n(S)= 4, então:

$$P(E) = \frac{3}{4} = 0{,}75 \text{ ou } 75\% \text{ de chance.}$$

b. Um banco de dados de clientes cadastrados de uma agência de turismo possui 40 pessoas do sexo masculino e 60 pessoas do sexo feminino. Seja a experiência de selecionar uma pessoa do cadastro aleatoriamente. Qual a probabilidade de essa pessoa ser do sexo masculino?

E= pessoa selecionada do cadastro de clientes ser do sexo masculino.

n(E)= 40

n(S)= 100

$$P(E) = \frac{40}{100} = 0{,}40 \text{ ou } 40\% \text{ de chance.}$$

Atividade

Atende ao Objetivo 3

1.

a. Em uma loja de departamento existem 70 calças de couro vermelho e 90, de couro preto. Selecionando uma calça aleatoriamente dentre as 160 existentes, qual a probabilidade de a calça selecionada ser de couro preto?

b. Quando uma pessoa é sorteada para avaliar como ótimo, bom, regular, ruim ou péssimo o roteiro de determinado pacote turístico que realizou pela agência de turismo, qual a probabilidade de a pessoa avaliá-lo positivamente?

Respostas Comentadas

a. O evento em estudo é: E= calça selecionada ser de couro preto.
Como foi dito no enunciado, têm-se 90 calças de couro preto de um total de 160 calças, o que resulta em:
n(E)= 90
n(S)= 160
Assim, você pode calcular a probabilidade do evento E ocorrer por meio de:

$$P(E) = \frac{90}{160} = 56{,}25\% \text{ de probabilidade.}$$

b. O evento em estudo é: E= a pessoa avaliar positivamente o referido roteiro turístico.
Assim, como foi dito no enunciado, há duas possibilidades de o turista avaliar positivamente o roteiro e, deste modo, você pode montar o conjunto E= {ótimo, bom}, e há cinco possibilidades no total, o que resulta em:
n(E)= 2
n(S)= 5
Portanto, você pode calcular a probabilidade de o evento E ocorrer por meio de:

$$P(E) = \frac{2}{5} = 40\% \text{ de probabilidade.}$$

Axiomas do cálculo das probabilidades

Pelos conceitos que acabamos de estudar, podemos concluir que:

1. $0 \leq P(E) \leq 1$;
2. $P(S)= 1$;
3. Se E_1 e E_2 forem eventos mutuamente exclusivos, ou seja, se $E_1 \cap E_2= \varnothing$, então: $P(E_1 \cup E_2)= P(E_1) + P(E_2)$.

Se $E_1 \cap E_2 \neq \varnothing$, então $P(E)= P(E_1 + E_2)= P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

Exemplos:

a. No lançamento de um dado, qual a probabilidade de sair a face 1 ou a face 4?

E_1 = sair a face 1 $\rightarrow$ $\mathbf{P(E_1)= 1/6}$

E_2 = sair a face 4 $\rightarrow$ $\mathbf{P(E_2)=1/6}$

E= sair a face 1 ou a face 4.

Em probabilidade, a chance de sair um evento ou outro é igual à soma das probabilidades dos eventos envolvidos, então a probabilidade pedida é:

$$\boxed{P(E)= P(E_1 + E_2)}$$

Como E_1 e E_2 são mutuamente exclusivos, então:

$$\mathbf{P(E)= P(E_1 + E_2)= P(E_1) + P(E_2)= 1/6 + 1/6= 2/6= 1/3.}$$

b. Uma população é formada de 20 pessoas que consomem o produto A, 30 pessoas que consomem o produto B e 50 pessoas que consomem o produto C. Um pesquisador de mercado seleciona uma pessoa dessa população. Sabendo que uma pessoa não consome mais de um produto ao mesmo tempo, qual a probabilidade de ter sido selecionada uma pessoa que consome o produto A ou C?

E_1 = consumir o produto A → P**(E_1)= 20/100= 0,2.**

E_2 = consumir o produto C → P**(E_2)= 50/100= 0,5.**

E= consumir o produto A ou C.

Como E_1 e E_2 são mutuamente exclusivos, então:

P(E)= P(E_1 + E_2)=P(E_1) + P (E_2)= 0,2 + 0,5= 0,7 ou 70% de chances.

Atividade

Atende aos Objetivos 2 e 3

2. Em uma agência de turismo, o departamento de recursos humanos ofereceu a oportunidade de seus funcionários escolherem pelo menos dois cursos de língua estrangeira para aperfeiçoamento: inglês ou espanhol. A probabilidade de optarem pelo curso de inglês é de 30%, pelo curso de espanhol é de 40% e, por ambos, 10%. Qual a probabilidade de um funcionário selecionado aleatoriamente do banco de dados dos empregados da agência escolher um ou outro curso?

Resposta Comentada

O evento de interesse do problema é:
E= funcionário selecionado aleatoriamente do banco de dados dos empregados da agência de turismo escolher um ou outro curso.
Identificado o evento, podemos calcular a probabilidade de ele ocorrer por meio de:
P(E)= 0,30 + 0,40-0,10= 0,60 ou 60% de chances.

Regra do produto

Se dois eventos são independentes, a probabilidade de que eles se realizem simultaneamente é igual ao produto das probabilidades de realização dos dois eventos.

Sendo $\mathbf{p}_1$ a probabilidade de realização do primeiro evento e $\mathbf{p}_2$ a probabilidade de realização do segundo evento, a probabilidade de que tais eventos se realizem simultaneamente é dada por:

$$\boxed{\mathbf{p_1 \times p_2}}$$

Ou: $\boxed{\mathbf{P(E_1 \cap E_2) = P(E_1) \times P(E_2)}}$

Exemplo:

a. Seja o experimento aleatório o lançamento de dois dados.

Sabemos que a probabilidade de obtermos 1 no primeiro dado é:

$p_1 = 1/6$

e a probabilidade de obtermos 5 no segundo dado é:

$p_2 = 1/6$.

Logo, a probabilidade de obtermos, simultaneamente, 1 no primeiro e 5 no segundo é:

P= 1/6 x 1/6 = 1/36.

Atividade

Atende aos Objetivos 2 e 3

3. A probabilidade de um turista ficar satisfeito com o atendimento no restaurante de um hotel é de 25%. A probabilidade de um outro turista ficar satisfeito com o atendimento no restaurante do mesmo hotel é de 40%. Suponhamos que os dois turistas vão frequentar o restaurante do hotel num mesmo momento e de forma independente. Qual a probabilidade de os dois turistas ficarem satisfeitos simultaneamente?

Resposta Comentada

Pelo enunciado da Atividade 3, você tem:
A probabilidade de o turista 1 ficar satisfeito é:
$p_1 = 0{,}25$,
A probabilidade de o turista 2 ficar satisfeito é:
$p_2 = 0{,}40$.
Logo, a probabilidade de simultaneamente os dois turistas ficarem satisfeitos é calculada pela regra do produto e resultará em:
$P = 0{,}25 \times 0{,}40 = 0{,}10$ *ou 10%*

Vamos tratar agora de um tipo de problema que nos coloca a reposição ou não de elementos. Seja o seguinte exemplo:

Suponhamos que um setor de uma empresa tenha cinco funcionários operacionais e dois gerentes, e que a diretoria irá selecionar dois funcionários desse setor, um após o outro, para obtenção de um prêmio de final de ano: uma passagem de ida e volta para os EUA para cada um. Suponha que o primeiro funcionário selecionado aleatoriamente seja operacional. Será que a probabilidade de o segundo funcionário selecionado também ser operacional é influenciada pela retirada do primeiro funcionário? Qual é a chance de a mesma pessoa ser selecionada na segunda seleção?

Assim, temos duas situações a considerar:

1. se houver reposição do primeiro funcionário, o setor vai ter a mesma configuração inicial, e, então, a 1ª retirada em nada influenciará na 2ª retirada, ou seja, temos **eventos independentes**;
2. se não houver a reposição da 1ª retirada, o setor conterá um funcionário a menos, isto é, diminui a probabilidade de sair um funcionário operacional na 2ª retirada, ou seja, temos "**eventos condicionados**".

"Eventos condicionados" ($E_1|E_2$) – Dois eventos associados a uma mesma experiência aleatória são ditos condicionados quando a ocorrência prévia de um deles aumenta ou diminui a ocorrência do outro. Já a ocorrência de um deles modifica a ocorrência do outro.

Exemplos:

a. Suponhamos que uma pessoa que está saindo para trabalhar de manhã tem dúvida se leva guarda-chuva ou não ao sair. Ele vai à janela ver o tempo. A chance de sair com guarda-chuva depende da informação que obtiver ao olhar o tempo: se o tempo estiver "ruim", a probabilidade de sair com guarda-chuva aumenta, ou seja, os eventos "tempo ruim" e "sair com guarda-chuva" são condicionados.

b. Seja o evento E_1 = "a letra *u* ocorre na palavra" e o vento E_2 = "a letra *q* ocorre na palavra". Certamente o evento E_1 tem uma probabilidade, mas ao saber que o evento E_2 ocorre, fica mais certo de que E_1 deve também ocorrer, uma vez que *q* raramente ocorre em uma palavra sem vir seguido de *u*.

c. Se for sabido que os ônibus de uma certa linha passam em um ponto em intervalos de, aproximadamente, 10 minutos, a probabilidade de passar um ônibus dessa linha no próximo minuto será fortemente influenciada pelo conhecimento que se tem da passagem de um ônibus da linha nos últimos cinco minutos.

Atividades Finais

Atendem aos Objetivos 2 e 3

1. Considere o experimento aleatório lançar duas moedas e observar a face voltada para cima. Calcule a probabilidade de ocorrência do evento A = ocorre pelo menos uma cara.

2. Uma rosa é retirada de uma caixa contendo: 10 rosas vermelhas (V); 30 rosas brancas (B); 20 rosas amarelas (A); 15 rosas cor-de-rosa (CR). Encontre a probabilidade de a rosa retirada da caixa ser:

a. cor-de-rosa ou vermelha;

b. não amarela;

c. branca;

d. vermelha ou branca ou amarela;

e. azul.

3. Um hotel possui 100 suítes tipo executivo (E), 8, superior (S) e 3, *master* (M). Uma suíte é escolhida ao acaso. Calcule a probabilidade de que:

a. a suíte não seja tipo executivo;

b. a suíte seja tipo executivo ou *master*;

c. a suíte seja tipo executivo e superior.

4. Uma operadora de turismo irá recepcionar 300 turistas em determinada capital turística. Desses turistas, 100 irão visitar museus (M) e 50 irão visitar reservas ecológicas (RE). Considerando ainda que 30 turistas farão os dois passeios, qual a probabilidade de que, selecionando aleatoriamente um turista recepcionado por esta operadora, ele:

a. Visite somente reservas ecológicas?

b. Visite museus ou reservas ecológicas?

c. Visite somente museus?

d. Não visite reservas ecológicas?

5. As chances de dois atacantes, A e B, de um time de futebol marcarem um gol de falta são: P(A)= 3/5 e P(B)= 5/7, respectivamente. Se cada um deles bate uma falta no jogo, qual a probabilidade de:

a. Não marcarem nenhum gol?

b. Marcarem dois gols?

Respostas Comentadas

1. Para este experimento, o espaço amostral é S= {(Cara Cara), (Cara Coroa), (Coroa Cara), (Coroa Coroa)}, assim, n(S)= 4.
O evento é dado por A= {(Cara Cara), (Cara Coroa), (Coroa Cara)}, assim, n(A)= 3.

Portanto:

$P(A)=\dfrac{3}{4} = 0{,}75 = 75\%$ *de probabilidade.*

2. Para este experimento, o tamanho do espaço amostral é n(S)= 75, assim, para calcular as probabilidades, vamos considerar as situações:

a. cor-de-rosa ou vermelha $\Rightarrow P(CR \cup V)$

$P(CR \cup V)= P(CR) + P(V)= \dfrac{15}{75}+\dfrac{10}{75}=\dfrac{25}{75}=\dfrac{1}{3}=0{,}33=33\%$;

b. não amarela $\Rightarrow P(A^c)$

$P(A^c)= 1 - P(A)= 1 - \dfrac{20}{75}=\dfrac{55}{75}=\dfrac{11}{15}=0{,}73=73\%$;

c. branca $\Rightarrow P(B)$

$P(B)= \dfrac{30}{75}=\dfrac{2}{5}=\dfrac{11}{15}=0{,}40=40\%$;

d. vermelha ou branca ou amarela $\Rightarrow P(V \cup B \cup A)$

$P(V \cup B \cup A)= P(V)+P(B)+P(A)$

$=\dfrac{10}{75}+\dfrac{30}{75}+\dfrac{20}{75}=\dfrac{60}{75}=\dfrac{4}{5}=0{,}80=80\%$;

e. azul $\Rightarrow$ *(azul)=0, pois este é o evento impossível de ocorrer no experimento aleatório em questão. Ou seja, azul= I =* $\varnothing$.

3. Para este experimento, o tamanho do espaço amostral é n(S)= 111, assim, para calcular as probabilidades, vamos considerar as situações:

a. A suíte não ser tipo executivo $\Rightarrow P(S \cup M)$

$P(S \cup M)= P(S) + P(M)= \dfrac{8}{111}+\dfrac{3}{111}=\dfrac{11}{111}=0{,}099=9{,}9\%$

b. A suíte ser tipo executivo ou master $\Rightarrow P(E \cup M)$

$P(E \cup M)= P(E) + P(M)= \dfrac{100}{111}+\dfrac{3}{111}=\dfrac{103}{111}=0{,}93=93\%$

c. A suíte ser tipo executivo e superior $\Rightarrow P(E \cap S)= 0$*, pois este é o evento impossível de ocorrer no experimento aleatório em questão. Ou seja,* $E \cap S= I= \varnothing$.

4. Para este experimento, o tamanho do espaço amostral é n(S)= 300, assim, para calcular as probabilidades, vamos considerar as situações:

a. visitar somente reservas ecológicas $\Rightarrow P(RE)$

$$P(RE) = \frac{50}{300} + \frac{5}{30} = 0{,}17 = 17\%\ ;$$

b. visitar museus ou reservas ecológicas $\Rightarrow P(M \cup RE)$

$$P(M \cup RE) = P(M) + P(RE) = \frac{100}{300} + \frac{50}{300} = \frac{15}{30} = 0{,}50 = 50\%\ ;$$

c. visitar somente museus $\Rightarrow P(M)$

$$P(M) = \frac{100}{300} = \frac{1}{3} = 0{,}33 = 33\%\ ;$$

d. não visitar reservas ecológicas $\Rightarrow P(RE^c)$

$$P(RE^c) = 1 - P(RE) = 1 - \frac{50}{300} = \frac{250}{300} = 0{,}83 = 83\%\ .$$

5. Neste experimento aleatório, estamos trabalhando com dois eventos independentes, ou seja, o atacante A fazer, ou não, um gol quando bate uma falta não depende do atacante B fazer, ou não, um gol quando bate uma falta. Assim, para calcular as probabilidades, vamos considerar as situações:

a. não marcarem nenhum gol $\Rightarrow P(A^c \cap B^c) = P(A^c)\, P(B^c)$

$$P(A^c \cap B^c) = \left(1 - \frac{3}{5}\right)\left(1 - \frac{5}{7}\right) = \frac{2}{5} - \frac{2}{7} = 0{,}40 - 0{,}29 = 0{,}11 = 11\%$$

b. marcarem dois gols $\Rightarrow P(A \cap B) = P(A)\, P(B)$

$$P(A \cap B) = \left(\frac{3}{5}\right)\left(\frac{5}{7}\right) = \frac{3}{7} = 0{,}43 = 43\%.$$

Resumo

Probabilidade é uma medida relativa que indica a chance de um evento ocorrer.

O cálculo das probabilidades pode ser realizado *a posteriori*, com base na experiência. Essa é a definição frequentista de probabilidades, que não engloba todos os casos práticos de cálculos de probabilidades. A definição clássica de probabilidades permite o cálculo das probabilidades *a priori*. A definição clássica é calculada pela divisão do número de casos favoráveis sobre todos os possíveis.

Os axiomas do cálculo das probabilidades permitem medir a chance da união de eventos e a regra do produto.

Informação sobre a próxima aula

Na próxima aula, abordaremos a noção de probabilidade condicionada.

15 Noções de probabilidades – probabilidade condicional

Meta da aula

Apresentar os conceitos básicos que envolvem o cálculo das probabilidades condicionadas.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. conceituar probabilidade condicionada;
2. realizar cálculos que envolvam probabilidades condicionadas;
3. diferenciar probabilidade condicionada de regra do produto.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, são necessários os conceitos básicos de probabilidades das Aulas 13 e 14.

Introdução

Ocorrem em muitas situações práticas de você ter uma informação extra sobre a ocorrência de determinado evento e de essa informação ser útil para o cálculo da probabilidade do evento do seu interesse – por exemplo, no lançamento de um dado, o evento cuja probabilidade é ocorrer o número três. Você viu, na Aula 14, que o valor dessa probabilidade é de 1/6, porém, se você for informado que nesse lançamento ocorreu um número menor do que quatro, essa informação poderá entrar no cálculo da probabilidade de ocorrer o número 3. A forma como isso pode ser feito é o assunto desta aula.

Definição: probabilidade condicional

Muitas vezes, o fato de ficarmos sabendo que certo evento ocorreu faz com que se modifique a probabilidade que atribuímos a outro evento. Denotaremos por $P(E_2/E_1)$ a probabilidade do evento E_2 sabendo-se que E_1 ocorreu ou, simplesmente, probabilidade de E_2 condicionada a E_1.

Com base na definição intuitiva de probabilidade, pode-se calcular a probabilidade condicional do evento E_2 ocorrer, dado que o evento E_1 já ocorreu (ou de que já se tenha conhecimento de E_1), pela fórmula:

$$P(E_2 / E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)}$$

ou

$$P(E_2 / E_1) = \frac{n(E_2 \cap E_1)}{n(E_1)}$$

Obs.: Com $P(E_1) \neq 0$.

Regra do produto para eventos condicionados

A regra do produto para eventos condicionados é útil, pois permite que você calcule a probabilidade de ocorrência da interseção de dois eventos, quando se tem conhecimento da probabilidade condicional.

Assim, se dois eventos são condicionados, pode-se obter das expressões de probabilidades condicionadas:

$$P(E_1 \cap E_2) = P(E_1) \,.\, P(E_2/E_1)$$

Exemplo:

Uma pesquisa de perfil demográfico feita junto a vinte turistas adultos, cadastrados em uma agência de viagens, revelou a base de dados na tabela a seguir:

Tabela 15.1: Perfil de turistas

Turista	Sexo	Idade	Nível escolar	Número de filhos	Classe social
1	M	35	2	2	B
2	M	25	2	1	B
3	F	40	3	1	C
4	M	25	2	3	B
5	M	32	2	2	C
6	F	22	2	0	C
7	M	37	3	2	B
8	M	28	2	0	B
9	F	25	2	1	B
10	F	39	3	2	C
11	M	35	1	1	B
12	F	21	1	0	A
13	F	27	0	0	A
14	F	45	2	2	C
15	M	57	4	4	C
16	F	33	2	2	A
17	M	36	1	0	B
18	M	35	2	2	C
19	M	33	2	2	B
20	F	22	3	0	C

Os códigos usados para montar a base de dados foram:

Variável sexo: M-masculino e F-feminino;
Variável idade: idade em anos (em dois dígitos);
Variável nível escolar: 0 - ausência de nível escolar; 1 - Ensino Fundamental; 2 - Ensino Médio; 3 - Ensino Superior; 4 - Pós-graduação;
Variável nº filhos: número de filhos do morador;
Variável classe social: A - alta; B - média; C - baixa.

Qual a probabilidade de que, ao selecionar aleatoriamente um turista dessa base de dados, os eventos a seguir ocorram?

a) Dado que é mulher, ter menos de 2 filhos.

b) Dado que é homem, ser da classe social C.

c) Dado que é da classe social B, não ter mais que 3 filhos.

d) Que seja um homem, sabendo-se que tem nível de escolaridade médio e classe social baixa.

e) Que seja um turista de Ensino Médio, com 2 ou menos filhos, sabendo-se que tem 30 anos ou mais.

Solução:

Antes de partir para a resposta, lembre-se de que é necessário sempre fazer a representação do espaço amostral (S). Neste caso, não precisaremos, pois a **Tabela 15.1** já é essa representação.

a) A probabilidade de ter menos de 2 filhos, dado que é mulher, é calculada pela condicional a seguir:

$P(<2filhos|F) = 6/9$.

Primeiro verifique a quantidade de mulheres (note que essa é a condição da probabilidade). O resultado será 9 mulheres. Atendida a condição, contam-se as mulheres que têm menos de 2 filhos (isso quer dizer que as mulheres com 2 ou mais filhos estão excluídas). O resultado dessa contagem será 6, de onde chegamos ao resultado $P(<2filhos|F) = 6/9$.

Pronto. Utilizando-se dessa lógica, você já adquiriu os subsídios necessários para entender e fazer os outros exemplos e atividades.

b) A probabilidade de ser da classe social C, dado que é homem, é calculada pela condicional a seguir:

$P(C|H) = 3/11$

c) A probabilidade de não ter mais que 3 filhos, dado que é da classe social B, é calculada pela condicional a seguir:

$P(<3\text{filhos}|B) = 9/9$

d) A probabilidade de que seja um homem, sabendo-se que tem nível de escolaridade médio e classe social baixa, é calculada pela condicional a seguir:

$P(H|\ 2 \cap C) = 2/4$

e) A probabilidade de que seja um turista com Ensino Médio, com 2 ou menos filhos, e sabendo-se que tem 30 anos ou mais, é calculada pela condicional a seguir:

$P(2 \cap \leq 2|\geq 30 \text{ anos}) = 6/11$

Atividade

Atende aos Objetivos 1, 2 e 3

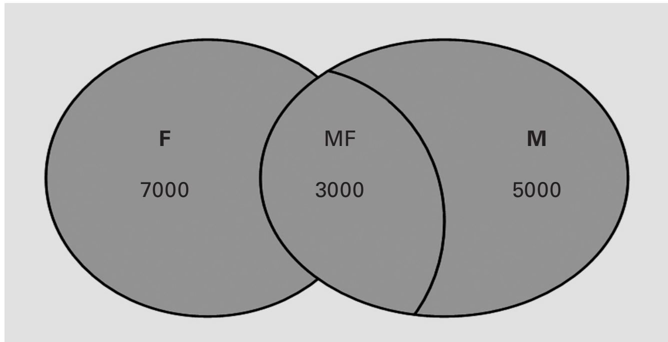
1. Em uma cidade existem 15.000 usuários de telefonia, dos quais 10.000 possuem telefones fixos, 8.000 deles, telefones móveis e 3.000, telefones fixos e móveis. Seja a experiência aleatória de uma operadora de telefone móvel, selecionar uma pessoa da cidade para oferecer uma promoção do tipo "*Fale grátis de seu móvel para seu fixo*", pergunta-se:

a. Já sabendo que ela tem telefone móvel, qual a probabilidade de ela ter telefone fixo também?

b. Já sabendo que ela tem telefone fixo, qual a probabilidade de ela ter telefone móvel também?

Respostas Comentadas

O espaço amostral S desse problema é dado por:



Nomeando os eventos de interesse, temos:
F = pessoa com telefone fixo.
M = pessoa com telefone móvel.
MF = pessoa com telefone fixo e móvel.

Assim, você pode calcular as probabilidades condicionais solicitadas por meio de:

a.

$$P(F|M) = \frac{n(MF)}{n(M)} = \frac{3.000}{8.000} = 3/8 = 0{,}375.$$

b.

$$P(M|F) = \frac{n(MF)}{n(M)} = \frac{3.000}{10.000} = 3/10 = 0{,}3.$$

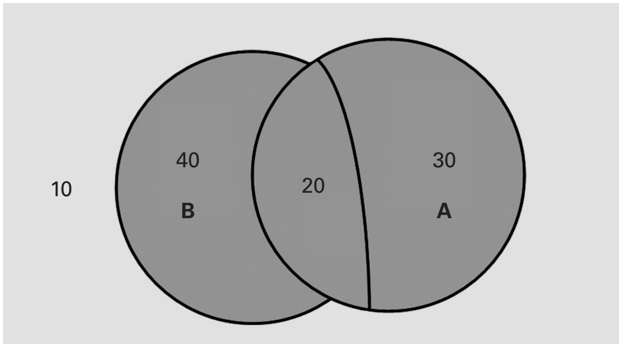
Atividade

Atende aos Objetivos 1 e 2

2. Em uma escola com 100 alunos, 40 estudam só Biologia, 30 estudam só Alemão, e 20 estudam Biologia e Alemão. Qual é a probabilidade de um aluno que já estuda Biologia estudar também Alemão?

Resposta Comentada

Espaço Amostral S



E_1 = aluno estudar Biologia.

E_2 = aluno estudar Alemão.

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{20/100}{60/100} = 20/60 = 33\%.$$

Você pode concluir, com o que foi estudado nesta aula, que quando a ocorrência de um evento está condicionada à ocorrência de outro evento, essa informação agregará confiabilidade ao ser incluída no cálculo da probabilidade da ocorrência do evento de interesse.

Atividade

Atende aos Objetivos 1, 2 e 3

3. Suponha que a seguinte tabela represente uma possível divisão dos alunos matriculados em um dado instituto de Matemática, clientes de uma viagem de excursão para um passeio turístico:

Curso	Sexo		Total
	M	F	
Matemática Pura	70	40	110
Matemática Aplicada	15	15	30
Estatística	10	20	30
Computação	20	10	30
Total	115	85	200

M = masculino; F = feminino.

Seleciona-se aleatoriamente um aluno presente na excursão. Foi constatado que ele é do curso de Estatística. Qual a probabilidade de ele ser homem?

Resposta Comentada

Sejam os eventos:
E_1 = aluno do curso de Estatística;
E_2 = aluno do sexo masculino.

Assim, a probabilidade condicional de ocorrer E_2, dado que E_1 ocorreu, é dada por:

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{10/200}{30/200} = 10/30 = 33\%.$$

Atividade

Atende aos Objetivos 1, 2 e 3

4. Considere o lançamento de um dado e a observação da face superior. Considere os seguintes eventos:

$E_1 = \{2,3,4,5\}$ e $E_2 = \{1,3,4\}$;

$E_1 = \{1,3,5,6\}$ e $E_2 = \{1,3,6\}$;

$E_1 = \{2,3,5,6\}$ e $E_2 = \{1,2\}$;

Em cada caso, obtenha a probabilidade condicional $P(E_2|E_1)$ e indique se os eventos E_1 e E_2 são independentes ou condicionados.

Respostas Comentadas

a) Considerando os eventos $E_1 = \{2,3,4,5\}$ e $E_2 = \{1,3,4\}$, temos que a sua interseção será dada por: $E_2 \cap E_1 = \{3,4\}$.

Com isso, você pode calcular as probabilidades: $P(E_2 \cap E_1) = 2/6$ e $P(E_1) = 4/6$. Assim, a probabilidade condicional de E_2, dado que E_1 ocorreu, será dada por:

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{2/6}{4/6} = 1/2 = 50\%.$$

$P(E_2) = 3/6 = 50\%$ (probabilidade incondicional).

Desse modo, conclui-se que a informação adicional de que E_1 já ocorreu não altera a ocorrência de E_2; portanto, são independentes.

b) Sejam os seguintes eventos de interesse: $E_1 = \{1,3,5,6\}$ e $E_2 = \{1,3,6\}$.

A interseção desses eventos é dada por: $E_2 \cap E_1 = \{1,3,6\}$.

Você, dessa maneira, pode calcular as probabilidades: $P(E_2 \cap E_1) = 3/6$ e $P(E_1) = 4/6$.

Assim, a probabilidade condicional de E_2, dado que E_1 ocorreu, será dada por:

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{3/6}{4/6} = 3/4 = 75\%.$$

$P(E_2) = 3/6 = 50\%$ (probabilidade incondicional).

Esse resultado permite você concluir que a informação adicional de que E_1 já ocorreu altera a ocorrência de E_2. A chance de ocorrer E_2 fica mais certa; portanto, são condicionados.

c) Sejam os eventos de interesse: $E_1 = \{2,3,5,6\}$ e $E_2 = \{1,2\}$.

A interseção desses eventos é dada por: $E_2 \cap E_1 = \{2\}$.

Você pode, então, calcular as probabilidades: $P(E_2 \cap E_1) = 1/6$ e $P(E_1) = 4/6$.

Assim, a probabilidade condicional de E_2, dado que E_1 ocorreu, será dada por:

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{1/6}{4/6} = 1/4 = 25\%.$$

$P(E_2) = 2/6 = 33\%$ (probabilidade incondicional).

Conclui-se que a informação adicional de que E_1 já ocorreu altera a ocorrência de E_2. A chance de ocorrer E_2 fica menos certa; portanto, são condicionados.

A essa altura deve estar claro que as noções de independência e probabilidade condicionada estão intimamente ligadas. O efeito da independência de E_1 e E_2 é tornar as probabilidades incondicionais iguais às respectivas probabilidades condicionais.

Atividade Final

Atende ao Objetivo 3

Numa grande agência de turismo, existem 4 turismólogos e 5 administradores de empresas. Seja a experiência aleatória de selecionar quatro desses profissionais, sem reposição, para formar uma comissão de elaboração de um novo plano turístico para uma cidade em alta temporada. Qual a probabilidade do seguinte evento:

{turismólogo $\cap$ administrador $\cap$ turismólogo $\cap$ administrador}?

Resposta Comentada

Vamos chamar:

TUR – o evento selecionar um turismólogo;

ADM – o evento selecionar um administrador.

Logo, a probabilidade pedida é: $P(TUR \cap ADM \cap TUR \cap ADM) = ?$

Como esses eventos são independentes, você irá calcular essa probabilidade utilizando a regra do produto, o que resultará em:

$P(TUR \cap ADM \cap TUR \cap ADM) =$

$$\frac{4}{9} \times \frac{5}{8} \times \frac{3}{7} \times \frac{4}{6} = 240/3.024 = 0,08.$$

Resumo

Probabilidade condicionada consiste em calcular a probabilidade de um evento ocorrer sabendo que um outro evento já ocorreu. No cálculo da probabilidade de um evento ocorrer, é incluída a informação adicional de que um outro evento já ocorreu. Probabilidade condicionada não é teorema. Seu cálculo se baseia na definição clássica de probabilidades. O número total de casos possíveis da probabilidade condicionada fica restrito ao evento que já ocorreu, e o número de casos favoráveis é o número de casos da interseção dos dois eventos envolvidos na probabilidade condicionada.

Informação sobre a próxima aula

A próxima aula é um dos temas mais importantes da teoria das probabilidades, e generaliza a regra do produto para eventos condicionados, permitindo que se "invertam" probabilidades dada uma partição do espaço amostral.

16 Noções de probabilidades – teorema de Bayes

Meta da aula

Apresentar os conceitos básicos que envolvem o teorema de Bayes.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. realizar cálculos que envolvam probabilidades de Bayes;
2. diferenciar probabilidade condicionada geral da definição específica do teorema de Bayes.

Pré-requisito

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos básicos de probabilidades das Aulas 13, 14 e 15.

Introdução

O teorema de Bayes é importante porque nos permite inverter probabilidades condicionais. Às vezes é fácil calcular $P(B/E_i)$, mas o que desejamos conhecer é $P(E_i/B)$. O teorema de Bayes nos permite calcular $P(E_i/B)$ em termos de $P(B/E_i)$. O teorema de Bayes nada mais é do que a "mistura" dos teoremas da probabilidade total com a regra do produto.
Como esse teorema é de suma importância dentro da teoria estatística, esta aula será dedicada à sua exaustiva aplicação.

Teorema de Bayes

Suponhamos um evento B que só pode ocorrer devido a uma das causas complementares E_1, E_2, E_3,... , E_n , eventos de um mesmo espaço amostral S. Dado que o evento B tenha ocorrido, a probabilidade que tenha se manifestado devido à causa E_1 ou E_2 ou E_3 , ... , ou E_n é calculada pela fórmula a seguir, denominada fórmula da probabilidade das causas ou dos antecedentes.

$$\mathbf{P(E_i / B) = \frac{P(E_i).\, P(B / E_i)}{\sum_{i=1}^{n}[P(E_i).\, P(B / E_i)]}}$$

O denominador do teorema de Bayes é o teorema da probabilidade total. Esse resultado pode ser demonstrado considerando-se o evento B subdividido em suas intersecções com os eventos E_i e aplicando-se as propriedades da regra do produto. Na prática, essa é a probabilidade de ocorrer o evento B.

Exemplo:

Em uma cidade, durante um período de observação, verificou-se que o trânsito ficou engarrafado 30% das vezes no horário do *rush* da manhã. Nos dias em que o trânsito ficou engarrafado, um funcionário chegou atrasado 10% das vezes e, nos dias de trânsito bom, ele chegou atrasado com uma freqüência de 1%. Certo dia, o funcionário chegou atrasado. Qual a probabilidade de ter sido em um dia de trânsito engarrafado?

Solução:

Evento efeito B: chegar atrasado.

Eventos causais (E_i): trânsito engarrafado(E_1) e trânsito não engarrafado(E_2).

Elementos da fórmula (modelagem):

$P(E_1)=0,3 \qquad P(B|E_1)=0,10$

$P(E_2)=0,7 \qquad P(B|E_2)=0,01$

$$P(E_i/B) = \frac{P(E_i).P(B/E_i)}{\sum_{i=1}^{2}[P(E_i).P(B/E_i)]}$$

$$P(E_1 / B) = \frac{0,3 . 0,10}{[(0,3 . 0,10) + (0,7 . 0,01)]} = 0,81$$

A probabilidade de o atraso ter ocorrido num dia de trânsito engarrafado é de 81%.

Atividade

Atende aos Objetivos 1, 2

1. Um indivíduo pode chegar atrasado ao emprego utilizando-se apenas de um destes meios de locomoção: bicicleta, motocicleta ou carro. Sabe-se, por experiência, que a probabilidade de ele utilizar o carro é de 0,6; bicicleta 0,1 e a motocicleta 0,3. A probabilidade de chegar atrasado, dado que se utilizou do carro, é de 0,05; de bicicleta, 0,02 e de motocicleta, 0,08. Certo dia ele chegou atrasado, qual a probabilidade de ter sido devido ao uso do carro?

Resposta Comentada

Inicialmente, você deve considerar os eventos de interesse que são:
Evento efeito B: chegar atrasado.
Eventos causais (E_i): utilizar carro (E_1), bicicleta (E_2) ou motocicleta (E_3);
Assim, você pode agora partir para os elementos requeridos pela fórmula para a modelagem, que são as probabilidades dos eventos de interesse e as probabilidades condicionais, dadas por:

$P(E_1)= 0,6$ $\quad$ $P(B|E_1)= 0,05$

$P(E_2)= 0,1$ $\quad$ $P(B|E_2)= 0,02$

$P(E_3)= 0,3$ $\quad$ $P(B|E_3)= 0,08$

Isso posto, você tem, agora, elementos suficientes para alocar as probabilidades na fórmula de Bayes que fica:

$$P(E_i/B) = \frac{P(E_i).P(B/E_i)}{\sum_{i=1}^{3}[P(E_i).P(B/E_i)]} =$$

$$P(E_1\ /\ B) = \frac{0,6\ .\ 0,05}{[(0,6\ .\ 0,05) + (0,1\ .\ 0,02) + (0,3\ .\ 0,08)]} =$$

$$P(E_1\ /\ B) = \frac{0,03}{(0,03 + 0,002 + 0,024)} = 54\%$$

Com esse resultado, você pode concluir que a probabilidade de o indivíduo ter ido de carro, sabendo-se que ele chegou atrasado ao emprego, é de 0,54 ou 54%.

Conclusão

Pelo que foi exposto nesta aula, você pode concluir que quando se tem uma informação adicional a respeito de eventos de interesse, essa nova informação pode ser utilizada para o cálculo das probabilidades posteriores, prestando um grande auxílio no trabalho de pesquisador.

Atividades Finais

Atende aos Objetivos 1, 2

1. Ficou constatado que o aumento nas vendas de um produto comercializado por certa empresa num determinado mês pode ocorrer somente por uma das causas mutuamente exclusivas: ação de *marketing*, publicidade/propaganda, oscilações econômicas do país e sazonalidade. A probabilidade de haver uma ação de *marketing* eficaz no mês é de 40%, de publicidade/propaganda 30%, oscilações econômicas 20% e sazonalidade 10%. Uma pesquisa mostrou que a probabilidade de haver aumento nas vendas do produto devido a uma ação de *marketing* eficaz é de 7%, de publicidade/propaganda é de 7,5%, de oscilações econômicas no país é de 3% e de sazonalidade é de 2%. Em um dado mês, o incremento nas vendas foi considerável. Indique a causa mais provável.

2. Considerando a Atividade 1, qual a probabilidade de aumento nas vendas em dado mês?

Respostas Comentadas

1. Inicialmente, você deve considerar os eventos de interesse que são:
Evento efeito B: aumento nas vendas.
Eventos causais (E_i): ação de marketing *(E_1), publicidade/propaganda (E_2), oscilações econômicas no país (E_3) e sazonalidade (E_4).*
Assim, você pode agora partir para os elementos requeridos pela fórmula para a modelagem, que são as probabilidades dos eventos de interesse e as probabilidades condicionais, dadas por:

$P(E_1)$= 0,4 $\quad$ $P(B/E_1)$= 0,070

$P(E_2)$= 0,3 $\quad$ $P(B/E_2)$= 0,075

$P(E_3)$= 0,2 $\quad$ $P(B/E_3)$= 0,030

$P(E_4)$= 0,1 $\quad$ $P(B/E_4)$= 0,020

Agora você tem elementos suficientes para alocar as probabilidades na fórmula de Bayes, que fica:

$$P(E_i/B) = \frac{P(E_i).P(B/E_i)}{\sum_{i=1}^{n}[P(E_i).P(B/E_i)]} =$$

Dado que ocorreu um aumento das vendas, a probabilidade desse aumento ter sido devido à ação de marketing *é dada por:*

$$P(E_1 / B) = \frac{0,4 \,.\, 0,07}{[(0,4.\,0,07) + (0,3\,.\,0,075) + (0,2\,.\,0,03) + (0,1\,.\,0,02)]}$$

$$P(E_1 / B) = \frac{0,028}{(0,028 + 0,0225 + 0,006 + 0,002)} = 47,8\%$$

Dado que ocorreu um aumento das vendas, a probabilidade desse aumento ter sido devido à publicidade/propaganda é dada por:

$$P(E_1 / B) = \frac{0,03 \,.\, 0,075}{[(0,4.\,0,07) + (0,3\,.\,0,075) + (0,2\,.\,0,03) + (0,1\,.\,0,02)]}$$

$$P(E_1 / B) = \frac{0,0225}{(0,028 + 0,0225 + 0,006 + 0,002)} = 38,5\%$$

Dado que ocorreu um aumento das vendas, a probabilidade desse aumento ter sido devido às oscilações econômicas é dada por:

$$P(E_1 / B) = \frac{0,2 \,.\, 0,03}{[(0,4.\,0,07) + (0,3\,.\,0,075) + (0,2\,.\,0,03) + (0,1\,.\,0,02)]}$$

$$P(E_1 / B) = \frac{0,006}{(0,028 + 0,00225 + 0,006 + 0,002)} = 10,3\%$$

Dado que ocorreu um aumento das vendas, a probabilidade desse aumento ter sido devido à sazonalidade é dada por:

$$P(E_1 / B) = \frac{0,1 \,.\, 0,02}{[(0,4.\,0,07) + (0,3\,.\,0,075) + (0,2\,.\,0,03) + (0,1\,.\,0,02)]}$$

$$P(E_1 / B) = \frac{0,002}{(0,028 + 0,00225 + 0,006 + 0,002)} = 3,4\%$$

Pelas probabilidades calculadas anteriormente, você tem elementos para concluir que a causa mais provável para o aumento das vendas naquele mês foi a ação de marketing.

2. A probabilidade de aumento nas vendas em dado mês é dada pelo denominador do teorema de Bayes; probabilidade total, portanto, 5,8% .

Resumo

O teorema de Bayes é uma combinação dos teoremas da probabilidade total e da regra do produto.
O teorema de Bayes nos permite calcular $P(E_i/B)$ em termos de $P(B/E_i)$, invertendo probabilidades.
O teorema de Bayes tem aplicações importantes dentro da estatística moderna.

Informação sobre a próxima aula

A próxima aula aborda o conceito de variáveis aleatórias, isto é, aquelas que sofrem variações devido ao acaso, a fatores não controlados.

17

Variáveis aleatórias discretas

Meta da aula

Apresentar o conceito de variáveis aleatórias discretas com enfoque sobre o cálculo de probabilidades e estatísticas descritivas.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. reconhecer o conceito variável aleatória;
2. calcular as probabilidades de ocorrência de uma variável aleatória discreta;
3. calcular a esperança e a variância de uma variável aleatória discreta para um conjunto de dados reais.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos de variáveis e seus tipos (Aula 2 do Volume 1), medidas de tendência central e dispersão (Aulas 9 e 10 do Volume 1) e probabilidades (13, 14 e 15 do Volume 2).

Introdução

Para descrever um experimento aleatório, é conveniente associar valores numéricos aos seus resultados. Como os eventos que ocorrem quando se realizam experimentos aleatórios variam a cada realização dos mesmos, os valores numéricos que lhes são associados também irão variar.
Em probabilidade, a função que associa a cada ponto do espaço amostral um número real é denominada variável aleatória (v.a.).

Variável aleatória

Variável aleatória é uma função definida num espaço amostral que assume valores reais, ou seja, é uma função que associa cada elemento do espaço amostral a um número Real.

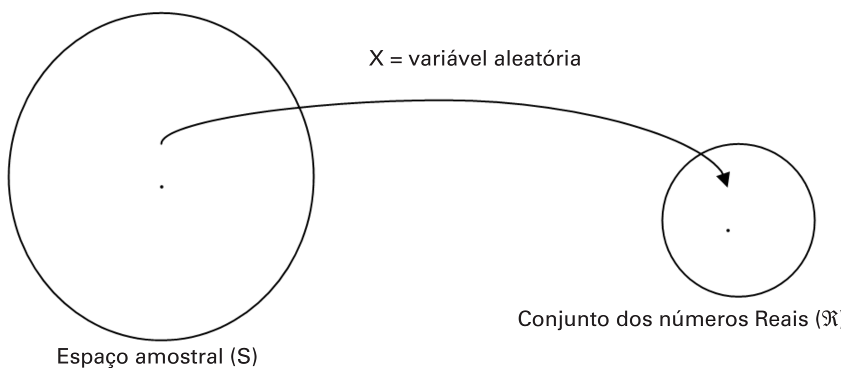


Figura 17.1: Esquema conceitual da variável aleatória.

Exemplo:

Considere o experimento aleatório de lançar duas moedas e observar a face voltada para cima. Represente por X a variável aleatória número de caras. A cada ponto do espaço amostral você pode associar um número Real. Descreva o resultado.

Solução:

O espaço amostral do problema é dado pelo conjunto constituído dos seguintes pontos amostrais:

S={Cara, Cara; Cara, Coroa; Coroa, Cara; Coroa, Coroa}

Podemos colocar esse resultado em uma tabela, associando a cada ponto amostral um número real; aí vamos ter a descrição da variável aleatória da seguinte maneira:

Tabela 17.1: Pontos amostrais e valores possíveis da variável aleatória

Pontos amostrais	Número Real: x_i = número de caras no ponto amostral i
Cara/Cara	2
Cara/Coroa	1
Coroa/Cara	1
Coroa/Coroa	0

Você pode ver, assim, que a cada ponto do espaço amostral associou-se um número Real.

Atividade

Atende ao Objetivo 1

1. Uma urna contém duas bolas brancas (B) e três bolas vermelhas (V). Represente por X a variável aleatória número de bolas vermelhas, ao selecionar aleatoriamente sucessivamente, sem reposição, duas bolas dessa urna. A cada ponto amostral você pode associar um número Real. Represente esse resultado.

Resposta Comentada

O espaço amostral do problema é dado pelo conjunto constituído dos seguintes pontos amostrais:

S={VV; VB; BV; BB}

Podemos colocar esse resultado em uma tabela, associando a cada ponto amostral um número real; aí vamos ter a descrição da variável aleatória da seguinte maneira:

Pontos amostrais	Número Real: x_i = número de bolas vermelhas no ponto amostral i
VV	2
VB	1
BV	1
BB	0

Você pode ver, assim, que a cada ponto do espaço amostral associou-se um número Real.

Uma variável aleatória pode ser discreta ou contínua. Nesse tópico, você vai aprender a calcular probabilidades das variáveis aleatórias discretas.

Variável aleatória discreta

Variável aleatória discreta é aquela que assume valores num conjunto finito ou enumerável de valores. Por exemplo: número de turistas em uma visita; número de leitos ocupados num hotel etc...

Para entendermos uma variável aleatória (v.a.) discreta, além de conhecermos os seus valores, é preciso também conhecermos as probabilidades associadas aos mesmos.

Assim, vamos definir função de distribuição de probabilidade.

Distribuição de probabilidades

Distribuição de probabilidades é uma função que associa a cada valor da variável sua probabilidade.

Seja X uma v.a. discreta qualquer, e x_1, x_2, ..., x_n, seus possíveis valores.

A distribuição de probabilidade de X é definida como:

$P(X=x_i) = p(x_i) = p_i$ onde $i= 1, 2, \ldots, n$

Ou

X	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

em que: p_i é a probabilidade do i-ésimo valor de x ocorrer, isto é, $P(X=x_1)$, $P(X=x_2)$, $P(X=x_3)$, ..., $P(X=x_n)$.

A distribuição de probabilidade da variável aleatória discreta deve satisfazer:

I. $\sum_{i=1}^{n} p(x_1) = 1$

II. $0 \leq p(x_1) \leq 1$

Exemplo:

Considere o experimento aleatório de lançar duas moedas e observar a face voltada para cima. Represente por X a variável aleatória número de caras. Encontre a função de distribuição de probabilidade da variável aleatória X.

Solução:

O espaço amostral do problema é dado pelo conjunto constituído dos seguintes pontos amostrais:

S= {Cara, Cara; Cara, Coroa; Coroa, Cara; Coroa, Coroa}.

E a variável aleatória em estudo é dada por:

X= número de caras.

Assim, você pode calcular a probabilidade de cada ponto amostral ocorrer, da seguinte maneira:

$P(X=0) = P(\text{Coroa, Coroa}) = 1/4$

$P(X=1) = P(\text{Cara, Coroa} \cup \text{Coroa, Cara}) = 1/2$

$P(X=2) = P(\text{Cara, Cara}) = 1/4$.

Então, a distribuição de distribuição de probabilidade da variável aleatória X será:

X	0	1	2
p_i	1/4	1/2	1/4

Que satisfaz as condições:

I. $\sum_{i=1}^{n} p(x_1) = 1$

II. $0 \leq p(x_1) \leq 1$

Atividade

Atende ao Objetivo 2

2. Uma urna contém duas bolas brancas (B) e três bolas vermelhas (V). Represente por X a variável aleatória número de bolas vermelhas. Suponha que são sorteadas ao acaso duas bolas, uma após a outra, sem reposição. Encontre a função de distribuição de probabilidade da variável aleatória X.

Resposta Comentada

O espaço amostral do problema é dado pelo conjunto constituído dos seguintes pontos amostrais:

S={VV; VB; BV; BB}

E a variável aleatória em estudo é dada por:

X= número de bolas vermelhas

Assim, você pode calcular a probabilidade de cada ponto amos-tral ocorrer, da seguinte maneira:

$P(X=0) = P(B \cap B) = P(B)\,P(B) = (2/5)\,(1/4) = 1/10$

$P(X=1) = P(V \cap B) + P(B \cap V) = P(V)\,P(B) + P(B)\,P(V) = (3/5)\,(2/4) + (2/5)\,(3/4) = 6/10$

$P(X=2) = P(V \cap V) = P(V)\,P(V) = (3/5)\,(2/4) = 3/10.$

Então, a função de distribuição de probabilidade da variável aleatória X será:

X	*0*	*1*	*2*
p_i	*1/10*	*6/10*	*3/10*

Que satisfaz as condições:

I. $\sum_{i=1}^{n} p(x_1) = 1$

II. $0 \leq p(x_1) \leq 1$

Esperança e variância da variável aleatória discreta

A esperança, ou média, de uma variável aleatória discreta X é dada por:

$E(X) = \sum_{i=1}^{n} x_i\, p(x_i)$, onde:

x_i é o valor que a v.a. pode assumir;

$p(x_i)$ é a probabilidade da v.a. assumir o valor x_i.

A variância de uma variável aleatória discreta X é dada por:

$Var\,(X) = \sum_{i=1}^{n} [x_i - E(X)]^2\, p(x_i)$, onde:

x_i é o valor que a v.a. pode assumir;

$p(x_i)$ é a probabilidade da v.a. assumir o valor x_i.

E(x) é a esperança matemática da v.a. X.

Note que tanto E(X) quanto Var(X) são ponderadas pelos valores de probabilidade.

Exemplo:

Considere o experimento aleatório de lançar duas moedas e observar a face voltada para cima. Represente por X a variável aleatória número de caras. Calcule a esperança e a variância de X.

Solução:

Você sabe, do exemplo anterior, que a função de distribuição de probabilidade da variável aleatória X é:

X	0	1	2
p_i	1/4	1/2	1/4

Assim, pode calcular E[X] e Var[X], aplicando as expressões:

$$E(X) = \sum_{i=1}^{n} x_i\, p(x_i) = 0\ (1/4) + 1\ (1/2) + 2\ (1/4) = 1$$

$$\text{Var}\ (X) = \sum_{i=1}^{n} [x_i - E(X)]^2\ p(x_i) = \{[0\text{-}(1)]^2.(1/4) + [1\text{-}(1)]^2.(1/2) + [2\text{-}1]^2\ (1/4)\} = 1/2.$$

Atividade

Atende ao Objetivo 3

3. Uma urna contém duas bolas brancas (B) e três bolas vermelhas (V). Represente por X a variável aleatória número de bolas vermelhas. Suponha que são sorteadas ao acaso duas bolas, uma após a outra, sem reposição. Encontre a esperança e a variância da variável aleatória X.

Resposta Comentada

Você sabe, pelo exemplo anterior, que a função de distribuição de probabilidade da variável aleatória X é:

X	*0*	*1*	*2*
p_i	*1/10*	*6/10*	*3/10*

Assim, pode calcular E[X] e Var[X], aplicando as expressões:

$$E(X) = \sum_{i=1}^{n} x_i\, p(x_i) = 0\ (1/10) + 1\ (6/10) + 2\ (3/10) = 1{,}2$$

$$\text{Var}\ (X) = \sum_{i=1}^{n} [x_i - E(X)]^2\ p(x_i) = \{[0\text{-}(1{,}2)]^2\ (1/10)\} + \{[1\text{-}(1{,}2)]^2(6/10) + [2\text{-}(1{,}2)]^2\ (3/10)\} = 0{,}36.$$

Em várias situações é útil calcular a probabilidade acumulada até certo valor da v.a. X. Isto é feito pela função de distribuição acumulada.

Função de distribuição acumulada

Dada uma v.a. X discreta, sua função de distribuição acumulada será:

$F(x) = P(X \leq x), \forall x \in \Re$

Exemplo:

Encontre a função de distribuição acumulada da variável aletória, número de caras (X), do experimento de lançar duas moedas.

Solução:

Relembrando, a função de distribuição de probabilidade de X é:

X	0	1	2
p_i	1/4	1/2	1/4

Assim, sua função de distribuição acumulada será obtida da seguinte maneira:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x \leq 1 \\ \frac{1}{4} + \frac{1}{2} & 1 \leq x \leq 2 \\ \frac{1}{4} + \frac{1}{2} + \frac{1}{4} & x \geq 2 \end{cases}$$

Nesta aula de hoje, você aprendeu a identificar uma variável aleatória discreta, calcular as probabilidades de ocorrência dos pontos amostrais, montar sua função de distribuição de probabilidade e a acumulada.

Atividade Final

Atende aos Objetivos 1, 2 e 3

Considere o experimento aleatório de lançar uma moeda três vezes e observar as faces voltadas para cima. Represente por X a variável aleatória número de caras. Pede-se:

a. Você pode associar um número Real a cada ponto do espaço amostral. Descreva esse resultado.

b. Encontre a função de distribuição de probabilidade da variável aleatória X.

Resposta Comentada

O espaço amostral do problema é dado pelo conjunto constituído dos seguintes pontos amostrais, sendo C=Cara e K=Coroa:

S= {CCC; CCK; CKC; KCC; KKC; KCK; CKK; KKK }

a. Podemos colocar esse resultado em uma tabela, associando a cada ponto amostral um número real; vamos ter a descrição da variável aleatória da seguinte maneira:

Ponto amostral	**Número Real: x_i = número de caras no ponto amostral i.**
CCC	3
CCK	2
CKC	2
KCC	2
Ponto Amostral	Número Real: x_i = número de caras no ponto amostral i.
KKC	1
KCK	1
CKK	1
KKK	0

Você pode ver assim que a cada ponto do espaço amostral associou-se um número Real.
b. Sabendo-se o S e a variável aleatória, você pode calcular a probabilidade de cada ponto amostral ocorrer da seguinte maneira:

$P(X=0) = P(KKK) = 1/8$
$P(X=1) = P(KKC \cup KCK \cup CKK) = P(KKC) + P(KCK) + P(CKK) = 3/8$
$P(X=2) = (CCK \cup CKC \cup KCC) = P(CCK) + P(CKC) + P(KCC) = 3/8$
$P(X=3) = P(CCC) = 1/8$

Assim, a função de distribuição de probabilidade da variável aleatória X será:

X	0	1	2	3
p_i	1/8	3/8	3/8	1/8

Que satisfaz as condições:

I. $\sum_{i=1}^{n} p(x_i) = 1$

II. $0 \leq p(x_i) \leq 1$

Resumo

Variável aleatória é um valor que varia ao acaso, de forma imprevisível.
Variável aleatória discreta é aquela que resulta de processos de contagem.
Distribuição de probabilidade é uma função que associa a cada valor da variável aleatória sua probabilidade.
As distribuições de probabilidades também possuem seus parâmetros característicos: esperança matemática ou média e a variância, por exemplo.
Função de distribuição acumulada é uma outra forma de se dispor o comportamento probabilístico da variável aleatória X. Na prática, é a probabilidade da variável aleatória X assumir valor menor ou igual a um valor específico x.
Esperança matemática é o valor médio dos resultados da variável aleatória quando a experiência aleatória a ela associada é realizada um número muito grande de vezes.
A variância é o índice de dispersão que a variável aleatória apresenta em uma experiência realizada um número muito grande de vezes.

Informação sobre a próxima aula

Na próxima aula, estudaremos um modelo de distribuição de probabilidades para representar uma determinada classe de variáveis aleatórias discretas: a distribuição binomial.

18 Distribuição Binominal

Metas da aula

Apresentar o conceito de variáveis aleatórias discretas que seguem o Modelo Binomial com enfoque sobre o cálculo de probabilidades e cálculo da esperança e variância da variável aleatória binominal.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. reconhecer o conceito de variável aleatória;
2. calcular as probabilidades de ocorrência de uma variável aleatória discreta com Distribuição Binomial em uma situação real;
3. calcular a esperança e variância de uma variável aleatória discreta com Distribuição Binomial para um conjunto de dados reais.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos de variáveis e seus tipos, da Aula 2, medidas de posição e dispersão, das Aulas 9 e 10, e de probabilidades, das Aulas 13, 14, 15 e 17.

Introdução

Algumas variáveis aleatórias (v.a.) se adaptam muito bem a uma série de problemas práticos. Portanto, um estudo mais detalhado dessas variáveis aleatórias é de grande importância para a construção de modelos probabilísticos e estimação dos seus parâmetros.

Assim, será apresentado na sequência uma das principais distribuições para variáveis aleatórias discretas, que é a Distribuição Binomial, bem como a: função de distribuição de probabilidade; os parâmetros; a esperança; a variância.

Distribuição Binomial

A Distribuição Binominal é uma distribuição de probabilidade adequada aos experimentos que apresentam apenas dois resultados: sucesso ou fracasso.

Alguns exemplos de experimentos aleatórios desse tipo são:

1. Vender pacotes turísticos para a Europa (resultado: vender = sucesso ou não vender = fracasso).
2. Aprovação dos alunos de uma turma de cursinho em determinado vestibular (resultado: aprovado = sucesso ou não aprovado = fracasso).
3. Um time de futebol jogar um campeonato vencer as partidas (resultado: vencer = sucesso ou não vencer = fracasso).

Assim, a distribuição de probabilidades da variável aleatória que conta o número de sucessos em "n" experimentos aleatórios independentes é dada por:

$$P(X = x) = \binom{n}{x} p^{x} (1- p)^{n-x}, \quad x = 0, 1, \ldots$$

com parâmetros n e p.

Em que:

x = número de sucessos em n repetições independentes do experimento aleatório;

n = número de experimentos aleatórios (número de tentativas);

p = probabilidade de ocorrer sucesso em um experimento aleatório;

1 - p = probabilidade de ocorrer fracasso em um experimento aleatório.

Esperança e variância

O valor esperado e a variância de X são, respectivamente:

$E(X) = np \quad Var(X) = np(1 - p)$.

Exemplo

Sabe-se que 10% dos pacotes de viagem vendidos por uma agência de turismo são para a Europa. Determine as probabilidades de que, dentre três pacotes vendidos:

a. dois sejam para a Europa;

b. nenhum seja para a Europa.

Solução:

Vamos identificar qual é a variável aleatória em estudo, e quais são o sucesso e o fracasso associados a essa variável:

X: variável aleatória número de pacotes vendidos.

1 pacote vendido – para a Europa = sucesso;

para outro destino = fracasso.

Assim, você pode agora identificar os elementos da expressão da Distribuição Binomial.

a. x = número de sucessos, x = 2.

n = número de experimentos aleatórios, n = 3:

p = probabilidade de sucesso, p = 0,10 e 1-p = 0,90.

Com estes elementos podemos calcular a probabilidade:

$$P(X = 2) = \binom{3}{2} 0,10^2 (0,90)^{3\text{-}2} = 3(0,10)^2 0,90 \cong 1,03$$

Pode-se concluir que, dentre três pacotes vendidos, a probabilidade de dois serem para a Europa é de 0,03 ou, aproximadamente, 3%.

b. x= número de sucessos, x = 0.

n = número de experimentos aleatórios, n = 3;
p = probabilidade de sucesso, p = 0,10 e 1-p = 0,90.

Com esses elementos, podemos calcular a probabilidade:

$$P(X = 2) = \binom{3}{0} 0,10^0 (0,90)^{3\text{-}0} = 0,90^3 \cong 0,73$$

Pode-se concluir que, dentre três pacotes vendidos, a probabilidade de nenhum ser para a Europa é de 0,73 ou aproximadamente 73%.

Atividade

Atende aos Objetivos 1 e 2

1. Se 20% dos alunos de um curso pré-vestibular são aprovados para universidades federais, determine a probabilidade de que, em 4 alunos escolhidos aleatoriamente:

a. um seja aprovado em universidades federais;

b. nenhum seja aprovado em universidades federais;

c. menos que dois sejam aprovados em universidades federais.

Resposta Comentada

Primeiramente você irá identificar qual é a variável aleatória em estudo e quais são o sucesso e o fracasso associados a essa variável:
X: variável aleatória aprovação em universidades federais.
1 aluno realiza o vestibular: aprovado em universidade federal = sucesso, e reprovado em universidade federal = fracasso.
Assim, você pode agora identificar os elementos da expressão da Distribuição Binomial.

a. x= número de sucessos, x =1.
n = número de experimentos aleatórios, n = 4;
p = probabilidade de sucesso, p = 0,20 e 1-p = 0,80.
Com esses elementos, podemos calcular a probabilidade:

$$P(X=1)=\binom{4}{1}0,20^{0}(0,80)^{4-1}=4\,(0,20)^{1}(0,80)^{3}\cong 0,4096$$

Pode-se concluir que, dentre quatro alunos desse cursinho, a probabilidade de um ser aprovado em universidade federal é de 0,4096, aproximadamente.

b. x = número de sucessos, x = 0.
n = número de experimentos aleatórios, n = 4;
p = probabilidade de sucesso, p = 0,20 e 1-p = 0,80.
Com esses elementos podemos calcular a probabilidade:

$$P(X=1)=\binom{4}{0}0,20^{0}(0,80)^{4-0}=1(1)(0,80)^{4}\cong 0,4096$$

Pode-se concluir que, dentre quatro alunos desse cursinho, a probabilidade de um ser aprovado em universidade federal é de 0,4096, aproximadamente.

c. x= número de sucessos, x = 0 e 1.
n = número de experimentos aleatórios, n = 4;
p = probabilidade de sucesso, p = 0,20 e 1-p = 0,80.

Com esses elementos, podemos calcular a probabilidade:

$$P(X \langle\, 2) = P(X = 0) + P(X = 1) \cong 0,4096 + 0,4096 \cong 0,8192$$

Pode-se concluir que, dentre quatro alunos desse cursinho, a probabilidade de menos que dois, ou seja, nenhum ou um, serem aprovados em universidade federal é de 0,8192, aproximadamente.

Conclusão

Com o que foi visto nesta aula, você pode concluir que a Distribuição, ou modelo Binomial, pode ser utilizada para calcular a probabilidade de determinado número de sucesso (x) em n experimentos ou ensaios idênticos e independentes, cujos resultados possíveis sejam sucesso ou fracasso.

Atividade Final

Atende aos Objetivos 1, 2, 3

Numa agência de viagens, de cada 500 passagens vendidas, 60 são para a Bahia, ou seja, 12% das passagens vendidas são para a Bahia. Na venda de cinco passagens, qual a probabilidade de que:

a. uma seja para a Bahia?

b. duas ou mais sejam para a Bahia?

c. nenhuma seja para a Bahia?

d. Calcule a esperança e a variância dessa variável aleatória.

Resposta Comentada

Primeiramente, você irá identificar qual é a variável aleatória em estudo e quais são o sucesso e o fracasso associados a essa variável:
X: variável aleatória vender passagem aérea para a Bahia.
1 venda de passagem é realizada – o destino é a Bahia = sucesso; o destino não é a Bahia = fracasso.
Assim, você pode agora identificar os elementos da expressão da Distribuição Binomial.

a. x = número de sucessos, x = 3;
n = número de experimentos aleatórios, n = 5;
p = probabilidade de sucesso, p = 0,12 e 1-p = 0,88.
Com esses elementos, podemos calcular a probabilidade:

$$P(X = 3) = \binom{5}{3} 0,12^3 (0,88)^{5-3} = 10\,(0,12)^3 (0,88)^2 \cong 0,013$$

Pode-se concluir que, dentre cinco passagens aéreas vendidas nessa agência de turismo, a probabilidade de uma ser para a Bahia é de 0,013, aproximadamente.

b. x = número de sucessos, x ≥ 2;

n = número de experimentos aleatórios, n = 5;

p = probabilidade de sucesso, p = 0,12 e 1-p = 0,88.

Com esses elementos, podemos calcular a probabilidade:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = 1 - [P(X = 0) + P(X = 1)] =$$

$$1 - \left[\binom{5}{0} 0,12^0 (0,88)^{5-0} + \binom{5}{1} 0,12^1 (0,88)^{5-1}\right] = 1 - [0,88^2 + 5\,(0,12)\,(0,88)^4 = 1 - 0,8875 \cong 0,112$$

Pode-se concluir que, dentre cinco passagens aéreas vendidas nessa agência de turismo, a probabilidade de duas ou mais serem para a Bahia é de 0,112, aproximadamente.

c. x = número de sucessos, x = 0;

n = número de experimentos aleatórios, n = 5;

p = probabilidade de sucesso, p = 0,12 e 1-p = 0,88.

Com esses elementos, podemos calcular a probabilidade:

$$P(X = 0) = \binom{5}{0} 0,12^{0} (0,88)^{5-0} \cong 0,528$$

Pode-se concluir que, dentre cinco passagens aéreas vendidas nessa agência de turismo, a probabilidade de nenhuma ser para a Bahia é de 0,528, aproximadamente.

d. Os valores da esperança e da variância de uma variável aleatória com Distribuição Binomial são dados por, $E(X) = np \quad Var(X) = np(1-p)$, *assim, para essa atividade você terá:*

Esperança: E(X) = 5(0,12) = 0,60.

Variância: Var(X) = 5(0,12)(0,88) = 0,528.

Resumo

A Distribuição Binomial é um modelo de distribuição de probabilidades para variáveis aleatórias discretas, do tipo contagem.

Uma variável aleatória discreta tem Distribuição Binomial quando representa o número total de sucessos em "*n*" provas independentes e cada prova somente admite dois resultados possíveis.

Se ocorrer o evento de interesse, ocorre o sucesso (x = 1) e se não ocorrer o sucesso, ocorre o fracasso (X = 0). A soma dos sucessos nas *n* experiências tem, então, distribuição binomial.

A probabilidade de sucesso, "*p*", é fixa em cada prova em que as "*n*" experiências são realizadas.

Informação sobre a próxima aula

Na próxima aula, você irá estudar as variáveis aleatórias contínuas, aquelas que resultam de processos de medição e que podem ser modeladas por importantes distribuições de probabilidades contínuas, como a Curva Normal.

19 Variáveis aleatórias contínuas e a distribuição normal – conceitos básicos

Metas da aula

Apresentar o conceito de variáveis aleatórias contínuas com enfoque sobre o cálculo de probabilidades e estatísticas descritivas; apresentar os conceitos básicos que envolvem distribuição normal e o uso da distribuição normal reduzida.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. conceituar variáveis aleatórias contínuas;
2. conceituar a distribuição normal;
3. identificar as características da distribuição normal;
4. conceituar a distribuição normal reduzida;
5. interpretar e utilizar a tabela de probabilidades da distribuição normal reduzida;
6. identificar situações reais e práticas da aplicação da distribuição normal reduzida.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos básicos de probabilidades das Aulas 13, 14, 15 e 16.

Introdução

Você estudou, nas aulas anteriores, variáveis aleatórias discretas e o modelo binomial. A partir da aula de hoje, você vai estudar as variáveis aleatórias contínuas, que são aquelas provenientes de medidas e seus possíveis valores pertencentes a um intervalo de números reais.

Os modelos contínuos encontram importantes aplicações em diversas áreas da ciência. São exemplos de fenômenos aleatórios contínuos: velocidade do carro, altura do voo do avião, tempo de viagem, entre outros.

Assim, como vimos, uma variável aleatória pode ser discreta ou contínua. Neste tópico você vai aprender a calcular probabilidades das variáveis aleatórias contínuas por meio do modelo normal.

Conceitos de variáveis aleatórias

O pesquisador estuda variáveis. Ele diz que essas variáveis são aleatórias porque elas têm um componente que varia ao acaso. Por exemplo, a variabilidade dos pesos, ao nascer, de nascidos vivos de mesmo sexo, da mesma raça, da mesma idade gestacional e de filhos de mães em condições similares de saúde e alimentação é explicada pelo acaso. Então, o peso ao nascer é uma variável aleatória.

As grandes amostras de certas variáveis aleatórias permitem construir gráficos que têm aparência típica. Como exemplo, observe a **Figura 19.1**, que apresenta uma distribuição de pesos, ao nascer, de nascidos vivos brancos de sexo masculino, com cerca de 40 semanas de gestação. Gráficos com esse tipo de configuração são obtidos, por exemplo, quando se analisa o peso ao nascer de cerca de 2.000 nascidos com características iguais.

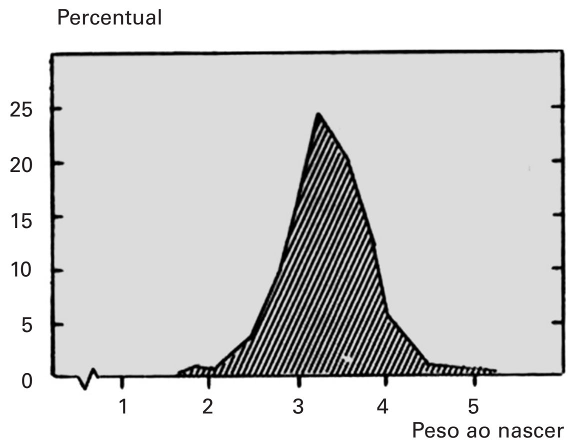


Figura 19.1: Peso ao nascer de nascidos vivos brancos do sexo masculino com cerca de 40 semanas de gestação.
Fonte: Do autor.

As medidas estatísticas que quando são observadas em grandes amostras geram o gráfico da **Figura 19.1** têm distribuição normal de probabilidades.

Variáveis aleatórias contínuas

As variáveis aleatórias contínuas são aquelas que assumem valores num intervalo do conjunto dos números reais. Por exemplo: peso (em kg) da mala dos turistas; preço (em reais) da diária de um hotel; etc.

Para entendermos uma variável aleatória (v.a.) contínua, além de conhecer o intervalo de valores a que ela pertence, é preciso atribuir probabilidades a esse intervalo de valores, o que é feito pela função de densidade de probabilidade, nome dado às distribuições de probabilidades de variáveis aleatórias contínuas.

A distribuição de probabilidade de variáveis aleatórias contínuas chama-se função densidade de probabilidade. A distribuição acumulada denomina-se função repartição de probabilidades. Essas formas de apresentação do comportamento probabilístico de variáveis aleatórias contínuas objetivam fornecer informações úteis para o pesquisador sobre as variáveis.

As distribuições de probabilidades de variáveis aleatórias contínuas podem ser representadas pelos seus principais parâmetros característicos: esperança matemática (média) e variância.

Muitas variáveis aleatórias contínuas seguem modelos típicos de distribuições de probabilidades de variáveis aleatórias, como a distribuição normal.

Distribuição normal

As variáveis observadas na prática são, quase sempre, resultados da soma de inúmeras outras variáveis aleatórias independentes. Assim, por exemplo: o tempo gasto na realização de uma viagem turística é a soma dos tempos despendidos nos diversos momentos até se chegar ao destino; as características biológicas dos indivíduos é a soma das hereditariedades de seus ancestrais etc.

Tais variáveis dão uma distribuição que, por ser muito freqüentemente encontrada na prática, é denominada distribuição normal.

Características gerais

As medidas biológicas, as medidas de produtos fabricados em série e os erros de medidas dão origem a gráficos semelhantes ao apresentado na **Figura 19.1**. Todas essas medidas são variáveis que têm distribuições que se aproximam da distribuição normal, apresentada na **Figura 19.2**:

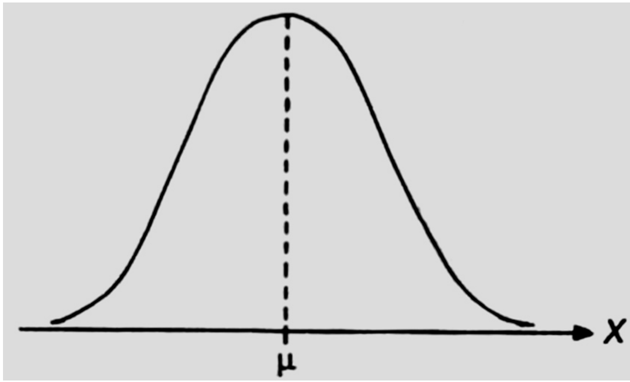


Figura 19.2: Gráfico da distribuição normal.
Fonte: Do autor.

A distribuição normal tem as seguintes características:

a. a variável aleatória pode assumir qualquer valor real;
b. o gráfico da distribuição normal é uma curva em forma de sino, simétrica em torno da média μ (lê-se mi), como mostra a **Figura 19.2**;
c. a área total sob a curva vale 1, porque essa área corresponde à probabilidade de a variável aleatória assumir qualquer valor real;
d. como a curva é simétrica em torno da média, os valores maiores do que a média e os valores menores do que a média ocorrem com igual probabilidade;
e. a configuração da curva é dada por dois parâmetros: a média μ e a variância σ^2. Mudando-se a média, muda-se a posição da distribuição, como mostra a **Figura 19.3**. Mudando-se a variância, muda-se a dispersão da distribuição, como mostra a **Figura 19.4**.

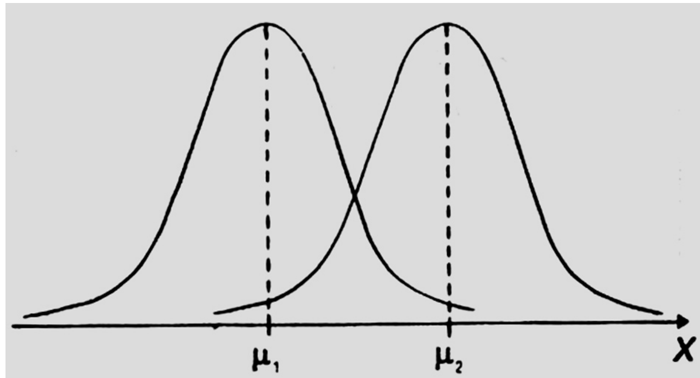


Figura 19.3: Duas distribuições normais de mesma variância e com médias diferentes.
Fonte: Do autor.

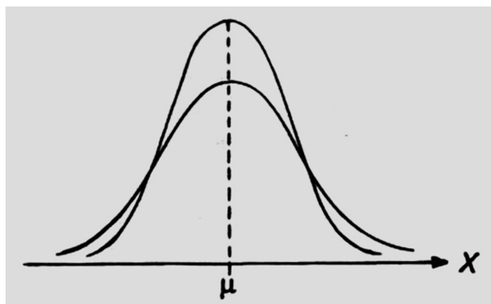


Figura 19.4: Duas distribuições normais de mesma média e com variâncias diferentes.
Fonte: Do autor.

Distribuição normal reduzida

Denomina-se distribuição normal reduzida a distribuição de Z, à distribuição normal de média zero e variância 1.

Uma variável aleatória normal reduzida segue o modelo a seguir:

$$Z = \frac{x - \mu}{\sigma}$$

Onde μ e σ são respectivamente a média e o desvio padrão de uma variável aleatória normal em observação x.

As probabilidades associadas à distribuição normal reduzida são facilmente obtidas em tabelas. Daí o interesse em estudar esse tipo particular de distribuição.

Observe a **Figura 19.5**. A área total sob a curva vale 1. Isso significa que a probabilidade de ocorrer qualquer valor real é 1. A curva é simétrica em torno da média zero. Então, a probabilidade de ocorrer valor maior do que zero é 0,5. Mas qual seria a probabilidade de ocorrer valor entre zero e z = 1,25, por exemplo?

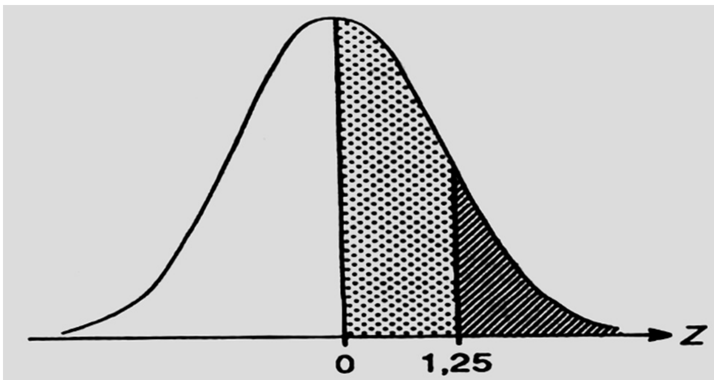


Figura 19.5: Probabilidade de ocorrer valor entre zero e z = 1,25.
Fonte: Do autor.

A probabilidade de ocorrer valor entre zero e z = 1,25 corresponde à área pontilhada na **Figura 19.5**. Essa probabilidade é encontrada na tabela de distribuição normal reduzida dada nesta aula, em anexo. Para mostrar como se usa esse tipo de tabela, reproduziu-se parte dela na **Figura 19.6**:

	0	1	2	3	4	5	6
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279

Figura 19.6: Probabilidade de ocorrer valor entre zero e 1,25 utilizando a tabela de distribuição normal reduzida.
Fonte: Do autor.

Na primeira coluna da tabela apresentada na **Figura 19.6** está o valor 1,2 (para facilitar, este valor foi sombreado). Na primeira linha da tabela apresentada na **Figura 19.6** está o valor 5 (para facilitar, este valor também foi sombreado). O número 1,2 compõe, com o algarismo 5, o número $z = 1,25$. No cruzamento da linha 1,2 com a coluna 5 está o número 0,3944 (também sombreado). Essa é a probabilidade de ocorrer o valor entre zero e $z = 1,25$, que corresponde à área pontilhada na **Figura 19.5**.

Considere outro problema. Qual é a probabilidade de ocorrer valor maior do que $z = 1,25$? Essa probabilidade corresponde à área hachurada na **Figura 19.5**. Como a probabilidade de ocorrer valor maior do que zero é 0,5 e a probabilidade de ocorrer valor entre zero e $z = 1,25$ (área pontilhada) é 0,3944, a probabilidade pedida (área hachurada) é: $0,5 - 0,3944 = 0,1056$ ou 10,56%.

Atividade

Atende aos Objetivos 1, 2, 3 e 4

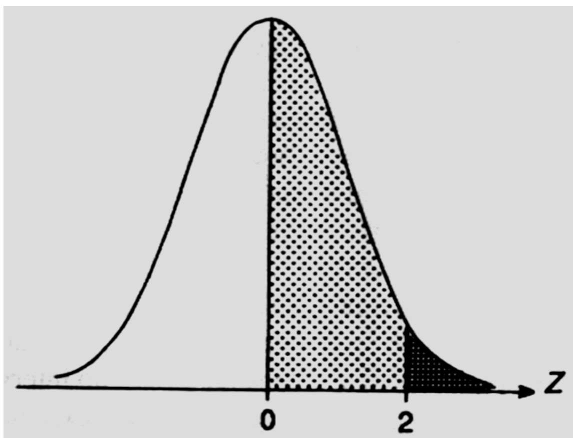
1. Em mulheres a passeio por uma cidade, foi constatado que a quantidade de hemoglobina no sangue é uma variável aleatória com distribuição normal de média $\mu = 6g$ e desvio padrão $\sigma = 1$. Faça o gráfico e calcule a probabilidade de uma mulher que visite a cidade apresentar de 6 a 8g de hemoglobina a cada 100ml de sangue.

Resposta Comentada

Primeiro, vamos calcular:

$$Z = \frac{x - \mu}{\sigma} = \frac{8 - 6}{1} = 2$$

A probabilidade de X assumir valor entre a média 6 e o valor 8 corresponde à probabilidade de Z assumir valor entre a média zero e o valor 2 (área pontilhada no gráfico a seguir). Esta probabilidade constante do anexo 1 é de 0,4772. Então, a probabilidade de uma mulher apresentar de 6g a 8g de hemoglobina por 100ml de sangue é 0,4772 ou 47,72%.

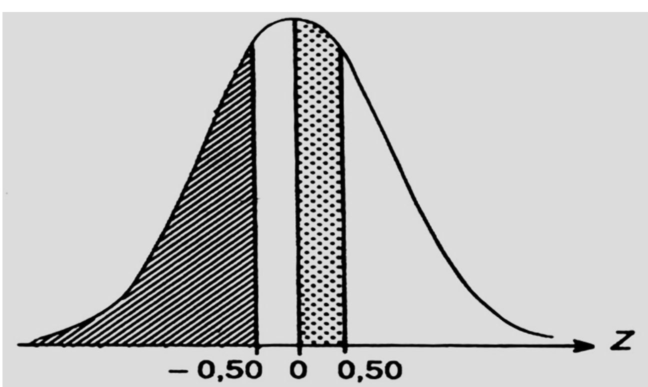


Probabilidade de taxa de hemoglobina entre 6g e 8g.

Atividade

Atende aos Objetivos 1, 2, 3 e 4

2. Qual é a probabilidade de ocorrer valor menor do que z = - 0,50 no problema da Atividade 1? Note que essa probabilidade corresponde à área hachurada na figura a seguir:



Probabilidade de ocorrer valor menor do que z = - 0,5.

Resposta Comentada

Para responder à pergunta, primeiro observe que a área em branco, entre zero e z = - 0,50, é igual à área pontilhada entre zero e z = 0,50. Mas quanto vale essa área? Procure, na primeira coluna da tabela de distribuição normal reduzida, o valor 0,5 e, na primeira linha, o valor zero, para compor o número z = 0,50. No cruzamento entre a linha e a coluna está o valor 0,1915, que é a probabilidade de ocorrer valor entre zero e z = 0,50.

Observe novamente a figura da resposta comentada da Atividade 1. A probabilidade de ocorrer valor menor do que z = - 0,50 é igual à probabilidade de ocorrer valor maior do que z = - 0,50. Como a probabilidade de ocorrer valor maior do que a média zero é 0,5, a probabilidade pedida é dada por: 0,5 - 0,1915 = 0,3085, ou 30,85%.

Atividade

Atende ao Objetivos 1, 2, 3 e 4

3. No problema da Atividade 1, qual é a probabilidade de uma mulher apresentar mais que 18g de hemoglobina por 100ml de sangue?

Resposta Comentada

A figura da resposta comentada da Atividade 1 mostra uma área com dupla hachura, que corresponde à probabilidade procurada. Como para x = 18 corresponde z = 2, e a probabilidade de Z assumir valor entre a média zero e o valor z = 2 é 0,4772, segue-se que a probabilidade de Z assumir valor maior do que 2 é:
0,5 - 0,4772 = 0,0228 ou 2,28%.

Muitas variáveis aleatórias de medidas contínuas apresentam distribuição normal de probabilidades. Aplicando a tabela da curva normal reduzida, podemos calcular probabilidades de quaisquer eventos envolvendo variáveis aleatórias contínuas.

Atividade Final

Atende aos Objetivos 1, 2, 3, 4 , 5 e 6

Os gastos dos hóspedes de um hotel, denominado de X, seguem a distribuição normal de média R$ 1200,00 e desvio padrão R$ 400,00. Calcule as probabilidades:

a) P(1200 < X < 2000)

b) P(X > 2000)

Resposta Comentada

a) Primeiro, é preciso calcular:

$$Z = \frac{x - \mu}{\sigma} = \frac{2000 - 1200}{400} = 2$$

A probabilidade de x assumir valor entre a média 1200 e o valor 2000 corresponde à probabilidade de Z assumir valor entre a média zero e o valor 2. Esta probabilidade, dada na tabela da normal reduzida anexa, é 0, 4772.
b) Como x = 2000 corresponde z = 2, e a probabilidade de z assumir valor entre a média zero e o valor z = 2 é 0, 4772, segue que a probabilidade de Z assumir valor maior do que 2000 é:
0,5 - 0, 4777 = 0,0228 ou 2,28%.

Resumo

As variáveis aleatórias contínuas são as que resultam de processos de medição, como peso e altura de turistas.
A distribuição normal é uma distribuição de probabilidades para variáveis aleatórias contínuas tais como medidas biométricas, tempos gastos em uma tarefa, medidas de produtos industrializados etc.
A distribuição normal tem, entre outras características, de ser simétrica em relação a sua média, e a área total sob a sua curva é igual a 1.
Para o cálculo das probabilidades de variáveis aleatórias normalmente distribuídas, recorremos à tabela da normal reduzida.

Informação sobre a próxima aula

Na próxima aula, iremos apresentar outros problemas envolvendo o cálculo de probabilidades na distribuição normal, inclusive a aproximação normal da binomial.

Anexo

	Último dígito									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4658	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

20 Distribuição normal – cálculo de probabilidades

Meta da aula

Apresentar cálculos de probabilidades na distribuição normal para quaisquer eventos dados.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. calcular probabilidades diversas envolvendo a distribuição normal;

2. aplicar a distribuição normal em situações da vida real e prática.

Pré-requisito

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos básicos de probabilidades.

Introdução

Na aula anterior, aprendemos a calcular probabilidades sob a curva normal quando reduzíamos apenas um dos valores envolvidos no evento de cálculo de probabilidades. Nesta aula, por exemplo, aprenderemos que existem eventos, no cálculo de probabilidades sob a curva normal, que são necessários quando reduzimos os dois valores do evento envolvidos no cálculo da probabilidade.

Probabilidades na distribuição normal

Suponha que a quantidade de colesterol em hóspedes de um *Spa* tenha distribuição normal com média 200mg e desvio padrão 20mg. O gráfico dessa distribuição está apresentado na **Figura 20.1**.

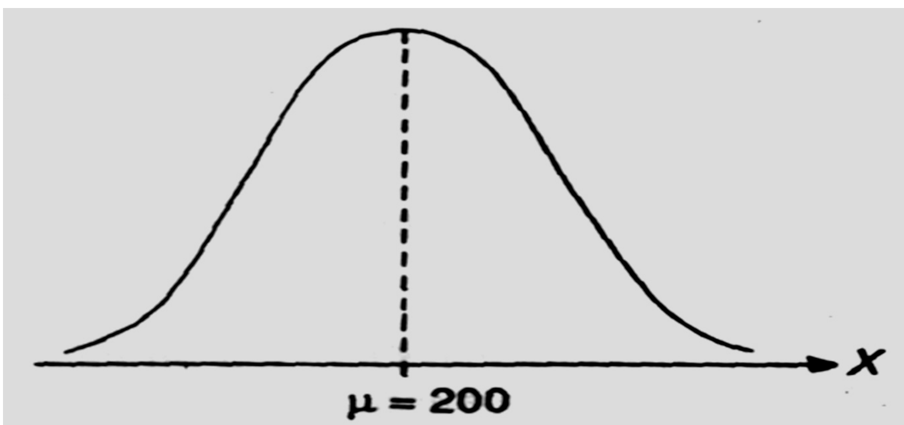


Figura 20.1: Distribuição normal da taxa de colesterol em hóspedes de um *Spa*.
Fonte: Do autor.

Pode existir interesse em calcular a probabilidade de uma pessoa hospedada no *Spa* apresentar entre 200 e 225mg de colesterol. Essa probabilidade corresponde à área pontilhada na **Figura 20.2**.

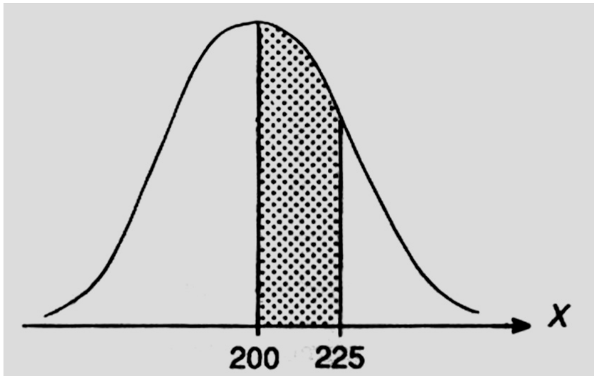


Figura 20.2: Probabilidade da taxa de colesterol entre 200 e 225mg.
Fonte: Do autor.

Para calcular probabilidades associadas à distribuição normal, usa-se um artifício. Sabe-se que, se X tem distribuição normal com média μ e desvio padrão σ, a variável a seguir tem distribuição normal reduzida. Fica, então, fácil obter as probabilidades associadas a qualquer distribuição normal: basta "reduzir" a distribuição e obter as probabilidades na tabela de distribuição normal reduzida, como mostra a Aula 19.

$$Z = \frac{X - \mu}{\sigma}$$

No exemplo em discussão, deseja-se obter a probabilidade de uma pessoa hospedada em um *Spa* apresentar entre 200 e 225mg de colesterol. Como a quantidade de colesterol tem distribuição normal com média μ = 200mg e desvio padrão σ = 20mg, a variável a seguir tem distribuição normal reduzida.

$$Z = \frac{X - 200}{20}$$

Nessa distribuição, a média é zero e, ao valor x = 225mg, corresponde:

$$Z = \frac{225 - 200}{20} = 1,25$$

A área pontilhada na **Figura 20.2** corresponde à área pontilhada na **Figura 20.3**. Então, a probabilidade de X assumir valor entre 200m e 225m é igual à probabilidade de Z assumir valor entre zero e z = 1,25 que, como se viu na Aula 19, é igual a 0,3944 ou 39,44%.

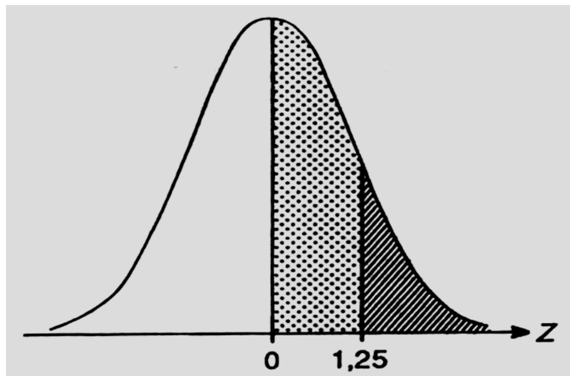


Figura 20.3: Probabilidade de ocorrer valor entre zero e z = 1,25.
Fonte: Do autor.

Considere outro exemplo agora: Qual é a probabilidade de um hóspede desse *Spa* apresentar menos do que 190mg de colesterol? Essa probabilidade corresponde à área hachurada na **Figura 20.4**.

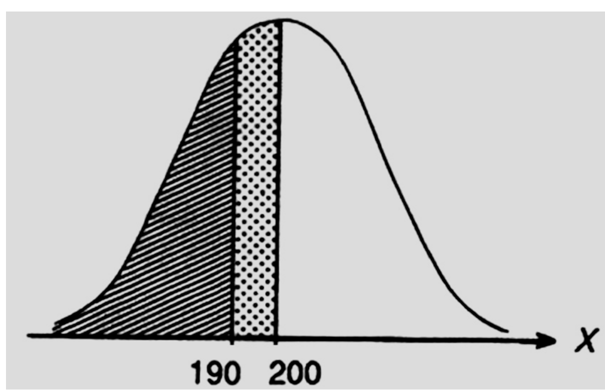


Figura 20.4: Probabilidade de a taxa de colesterol ser menor do que 190mg.
Fonte: Do autor.

Para resolver o problema, é preciso "reduzir" o valor x = 190mg. Obtém-se então:

$$Z = \frac{190 - 200}{20} = -0,50$$

A probabilidade pedida corresponde à probabilidade de Z assumir valor menor do que z = - 0,5 que, como se viu na Aula 19 Atividade 2, é 0,3085 ou 30,85%.

Atividade

Atende aos Objetivos 1 e 2

1. Os gastos dos hóspedes de um hotel, denominado X, seguem a distribuição normal de média R$1.200,00 e desvio padrão R$400,00. Qual a probabilidade de um hóspede apresentar um gasto entre R$1.000,00 e R$2.000,00?

Resposta Comentada

Primeiro, é preciso calcular:

$$Z = \frac{X - \mu}{\sigma} = \frac{1.000 - 1.200}{400} = -0,5$$

Em seguida, é preciso calcular:

$$Z = \frac{X - \mu}{\sigma} = \frac{2.000 - 1.200}{400} = 2,0$$

A probabilidade de x assumir valor entre 1.000 e o valor 2.000 corresponde à probabilidade de Z assumir entre o valor -0,5 e o valor 2. Basta consultar a tabela da normal reduzida entre zero e -0,5, que é igual a 0,1915 e a área entre zero e 2, que é igual a 0,4772 e somar: 0,1915 + 0,4772 = 0,6687 ou 66,87%.

Muitas variáveis aleatórias de medidas contínuas apresentam distribuição normal de probabilidades.

Aplicando a tabela da curva normal reduzida, podemos calcular probabilidades de quaisquer eventos envolvendo variáveis aleatórias contínuas. Mas é fundamental construir a figura da normal em cada situação problema para se verificar qual a área que corresponde à probabilidade do problema em questão.

Atividade Final

Atende aos Objetivos 1 e 2

Suponha que o tempo médio de permanência em uma pousada para turistas em uma cidade seja 50 dias, com desvio padrão igual a 10 dias. Se for razoável admitir que o tempo de permanência tem distribuição aproximadamente normal, qual é a probabilidade de um hóspede, turista na cidade, permanecer na pousada?

a. mais de 30 dias

b. menos de 30 dias

Respostas Comentadas

a.

Chamando de variável aleatória X o tempo de permanência, então a probabilidade pedida é:

$$P(X > 30)$$

Primeiro, vamos calcular:

$$Z = \frac{X - \mu}{\sigma} = \frac{30 - 50}{10} = -2,0$$

É fundamental que o aluno construa a figura da normal nessa situação problema para se verificar qual a área que corresponde à probabilidade do problema em questão. A probabilidade pedida corresponde na figura da distribuição normal reduzida à probabilidade de Z assumir o valor entre zero e – 2,0 mais 0,5. Essa probabilidade é, então, 0,4772 + 0,5 = 0,9772 ou 97,72%.

b.

Chamando de variável aleatória X o tempo de permanência, então a probabilidade pedida é:

$$P(X < 30)$$

Primeiro, vamos calcular:

$$Z = \frac{X - \mu}{\sigma} = \frac{30 - 50}{10} = -2,0$$

É fundamental que o aluno construa a figura da normal nessa situação problema para se verificar qual a área que corresponde à probabilidade do problema em questão. A probabilidade pedida corresponde na figura da distribuição normal reduzida à probabilidade de Z assumir o valor 0,5 menos o valor entre zero e – 2,0, Essa probabilidade é, então, 0,5 – 0,4772 = 0,0228 ou 2,28%.

Resumo

A distribuição normal deve ser usada para calcular probabilidades de eventos associados às variáveis aleatórias que sejam normalmente distribuídas ou aproximadamente distribuídas.
Para obter as probabilidades a qualquer distribuição normal, basta "reduzir" a distribuição e obter as probabilidades na tabela de distribuição normal reduzida.

Informação sobre a próxima aula

Na próxima aula, iremos apresentar o problema de mensurar o grau de correlação entre duas variáveis aleatórias normalmente distribuídas.

21 Correlação simples

Meta da aula

Apresentar os conceitos básicos que envolvem o cálculo da associação de duas variáveis aleatórias.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. diferenciar relação funcional de relação estatística;
2. definir e reconhecer um diagrama de dispersão;
3. definir correlação linear e classificá-la;
4. calcular o coeficiente de correlação linear de Pearson.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos básicos de probabilidades das Aulas 1, 19 e 20.

Introdução

Nas aulas anteriores, nossa preocupação era descrever a distribuição de valores de uma única variável. Com esse objetivo, aprendemos a calcular medidas de tendência central e variabilidade.

Quando, porém, consideramos observações de duas ou mais variáveis, surge um novo problema: as relações que podem existir entre as variáveis estudadas. Nesse caso, as medidas estudadas não são eficientes.

Assim, quando consideramos variáveis como peso e altura de um grupo de turistas, uso do cigarro e incidência de câncer, procuramos verificar se existe alguma relação entre as variáveis de cada um dos pares e qual o grau dessa relação. Para isso, é necessário o conhecimento de novas medidas.

Sendo a relação entre as variáveis de natureza quantitativa, a correlação é o instrumento adequado para descobrir e medir essa relação. Ficaremos restritos às relações entre duas variáveis que é chamada de correlação simples.

Relação funcional e relação estatística

Como sabemos, o perímetro e o lado de um quadrado estão relacionados. A relação que os liga é perfeitamente definida e pode ser expressa por meio de uma sentença matemática:

$$2p = 4\ell$$

Onde 2p é o perímetro e ℓ é o lado do quadrado.

Atribuindo-se, então, um valor qualquer a ℓ, é possível determinar exatamente o valor de 2p.

Consideremos, agora, a relação que existe entre o peso e a estatura de um grupo de turistas hospedados em um hotel. É evidente que essa relação não é do tipo da anterior; ela é bem menos precisa. Assim, pode acontecer que a estaturas diferentes correspondam pesos iguais ou que a estaturas iguais correspondam pesos diferentes.

As relações do tipo perímetro-lado são conhecidas como *relações funcionais* e as do tipo peso-altura, como *relações estatísticas*. As relações funcionais são um caso limite das relações estatísticas pois elas são totalmente precisas.

Quando duas variáveis estão ligadas por uma *relação estatística*, dizemos que existe correlação entre elas.

Diagrama de dispersão

Consideremos uma amostra aleatória, formada por dez dos 98 alunos de uma classe da faculdade A e pelas notas por eles obtidas em Matemática e Estatística.

Tabela 21.1: Notas em Matemática e Estatística

Nos	**Notas**	
	Matemática (X_i)	**Estatística (Y_i)**
1	5,0	6,0
8	8,0	9,0
24	7,0	8,0
38	10,0	10,0
44	6,0	5,0
58	7,0	7,0
59	9,0	8,0
72	3,0	4,0
80	8,0	6,0
92	2,0	2,0

Fonte: Do autor.

Representando, em um sistema coordenado cartesiano ortogonal, os pares ordenados (x_i, y_i), obtemos uma nuvem de pontos que denominamos *diagrama de dispersão*. Esse diagrama nos fornece uma idéia grosseira, porém útil, da correlação existente.

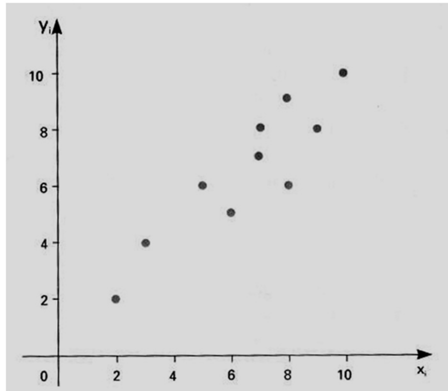


Figura 21.1: Diagrama de dispersão das notas dos alunos da faculdade.
Fonte: Do autor.

Correlação linear

Os pontos obtidos da **Figura 21.1**, vistos em conjunto, formam uma elipse em diagonal. Podemos imaginar que, quanto mais fina for a elipse, mais ela se aproximará de uma reta. Dizemos, então, que a correlação de forma elíptica tem como "imagem" uma reta, sendo, por isso, denominada *correlação linear*.

É possível verificar que a cada correlação está associada como "imagem" uma relação funcional. Por esse motivo, as relações funcionais são chamadas *relações perfeitas*.

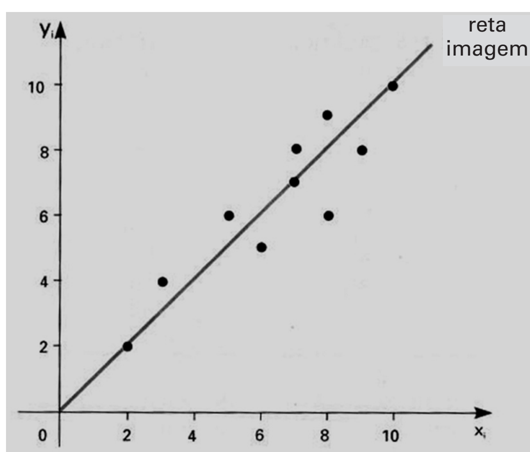


Figura 21.2: Correlação linear.
Fonte: Do autor.

Como a correlação em estudo tem como imagem uma reta ascendente, ela é chamada *correlação linear positiva*.

Assim, uma correlação é:

a) *linear positiva* se os pontos do diagrama têm como "imagem" uma reta ascendente;

b) *linear negativa* se os pontos do diagrama têm como "imagem" uma reta descendente.

Se os pontos apresentam-se dispersos, não oferecendo uma "imagem" definida, concluímos que não há relação alguma entre as variáveis em estudo.

Temos, então:

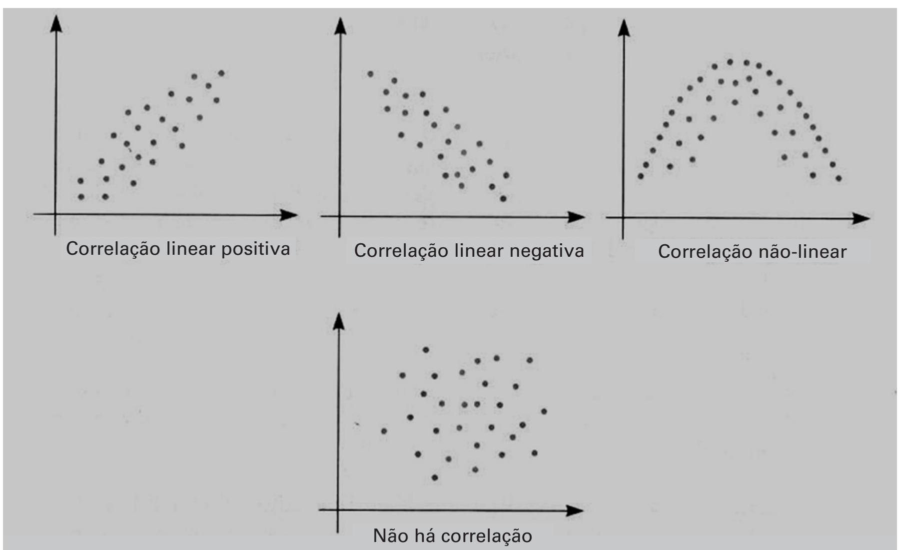


Figura 21.3: Os diferentes tipos de correlação.
Fonte: Do autor.

Coeficiente de correlação linear

O instrumento empregado para a medida da correlação linear é o coeficiente de correlação. Esse coeficiente deve indicar o grau de intensidade da correlação entre duas variáveis e, ainda, o sentido dessa correlação (positiva ou negativa).

Faremos uso do coeficiente de correlação de Pearson, que é dado pela fórmula:

$$r = \frac{n\sum x_i y_i - \sum(x_i)\sum(y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2].[n\sum y_i^2 - (\sum y_i)^2]}}$$

Onde n é o número de observações.

Os valores limites de r são -1 e +1, isto é, o valor de r pertence ao intervalo [-1, +1].

Assim:

a. se a correlação entre duas variáveis é perfeita e positiva, então r = +1;

b. se a correlação é perfeita e negativa, então r = -1;

c. se não há correlação entre as variáveis, então r = 0.

Logicamente:

a. se r = +1, há uma correlação perfeita e positiva entre as variáveis;

b. se r = -1, há uma correlação perfeita e negativa entre as variáveis;

c. se r = 0, ou não há correlação entre as variáveis ou a relação que porventura exista não é linear. Por exemplo, a relação curvilínea.

Para que uma relação possa ser descrita por meio do coeficiente de correlação de Pearson, é imprescindível que ela se aproxime de uma função linear. Uma maneira prática de verificarmos a linearidade da relação é a inspeção do diagrama de dispersão: se a elipse apresenta saliência ou reentrâncias muito acentuadas, provavelmente trata-se de *correlação curvilínea*.

Critério para indicar correlação entre as variáveis como significativa

Para podermos tirar algumas conclusões significativas sobre o comportamento simultâneo das variáveis analisadas, é necessário que:

$$0,6 \leq |r| \leq 1$$

Se $0,3 \leq | r | \leq 0,6$, há uma correlação relativamente fraca entre as variáveis.

Se $0 \leq | r | \leq 0,3$, a correlação é muito fraca entre as variáveis e, praticamente, nada podemos concluir sobre a relação entre as variáveis em estudo.

Exemplo:

Vamos , então, calcular o coeficiente de correlação relativo à **Tabela 21.1.** O modo mais prático para obtermos r é abrir, na tabela, colunas correspondentes aos valores de $x_i y_i$, x^2_i , y^2_i . Assim:

Tabela 21.2: Cálculo do coeficiente de correlação linear

Matemática (x_i)	**Estatística (y_i)**	$x_i y_i$	x^2_i	y^2_i
5	6	30	25	36
8	9	72	64	81
7	8	56	49	64
10	10	100	100	100
6	5	30	36	25
7	7	49	49	49
9	8	72	81	64
3	4	12	9	16
8	6	48	64	36
2	2	4	4	4
$\Sigma = 65$	$\Sigma = 65$	$\Sigma = 473$	$\Sigma = 481$	$\Sigma = 475$

Logo:

$$r = \frac{10 \times 473 - 65 \times 65}{\sqrt{[10 \times 481 - (65)^2] . [10 \times 475 - (65)^2]}}$$

$$r = \frac{4.730 - 4.225}{\sqrt{[4.810 - 4.225] . [4.750 - 4.225]}}$$

$$r = \frac{505}{\sqrt{585 \times 525}}$$

$$r = 0,911$$

Esta correlação é significativa, uma vez que 0,911 está no intervalo $06 \leq | r | \leq 1$.

Atividade

Atende aos Objetivos 3 e 4

1. Um grupo de turistas fez uma avaliação do peso aparente de alguns objetos típicos de uma região visitada. Com o peso real e a média dos pesos aparentes dados pelo grupo, obteve-se tabela:

Peso real	18	30	42	62	73	97	120
Peso aparente	10	23	33	60	91	98	159

Calcule o índice de correlação. Que tipo de correlação a amostra apresenta?

Resposta Comentada

É necessário realizar a tabela de cálculo a seguir:

Cálculo do coeficiente de correlação linear

Peso real (x_i)	**Peso aparente (y_i)**	$x_i y_i$	x^2_i	y^2_i
18	10	180	324	100
30	23	690	900	529
42	33	1.386	1.764	1.089
62	60	3.720	3.844	3.600
73	91	6.643	5.329	8.281
97	98	9.506	9.409	9.604
120	159	19.080	14.400	25.281
= 442	= 474	= 41.205	= 35.970	= 48.484

Logo:

$$r = \frac{7 \times 41.402 - 442 \times 474}{\sqrt{[7 \times 35.970 - (442)^2] . [7 \times 48.484 - (474)^2]}}$$

$$r = \frac{289.814 - 209.508}{\sqrt{56.426 \times 114.712}}$$

$$r = \frac{80.306}{80.453}$$

$$r = 0{,}9982$$

A correlação entre as variáveis é do tipo forte positiva.
Se tivermos duas variáveis aleatórias quantitativas, podemos medir a relação linear entre elas, primeiramente sondando esta relação por meio do diagrama de dispersão e, se for significativa, calcular o seu grau de intensidade por meio do coeficiente de correlação linear de Pearson*.*

Atividade Final

Atende aos Objetivos 3 e 4

Pretendendo-se estudar a relação entre as variáveis "consumo de energia elétrica de agências de turismo(x_i)" e "volume de serviços nas agências de turismo(y_i)", fez-se uma amostragem que inclui vinte agências, computando-se os seguintes valores:

$\Sigma x_i = 11{,}34$, $\Sigma y_i = 20{,}72$, $\Sigma x2i = 12{,}16$, $\Sigma y^2_i = 84{,}96$, $\Sigma x_i y_i = 22{,}13$

Calcule o coeficiente de correlação. Qual o tipo de correlação que as variáveis apresentam?

Resposta Comentada

Podemos aplicar a fórmula do coeficiente de correlação de Pearson diretamente:

$$r = \frac{20 \times 22{,}13 - 11{,}34 \times 20{,}72}{\sqrt{[20 \times 12{,}16 - (11{,}34)^2] \times [20 \times 84{,}96 - (20{,}72)^2]}}$$

$$r = \frac{442{,}60 - 234{,}96}{\sqrt{114{,}60 \times 1.269{,}88}}$$

$$r = \frac{207{,}64}{381{,}48}$$

$$r = 0{,}544$$

A correlação é significativamente positiva.

Resumo

O grau de relação estatística entre duas variáveis aleatórias quantitativas denomina-se correlação.

Podemos ter uma idéia grosseira porém útil do grau de correlação entre duas variáveis utilizando o diagrama de dispersão. A correlação de forma elíptica tem como "imagem" uma reta, denominada correlação linear.

Uma "imagem" ascendente dos pontos no diagrama revela uma correlação linear positiva. Uma "imagem" descendente dos pontos no diagrama revela uma correlação linear negativa.

O instrumento empregado para se medir a correlação linear entre duas variáveis aleatórias quantitativas é o coeficiente de correlação de Pearson(r).

Para que possamos tirar tomar decisões administrativas e científicas com base na relação de duas variáveis, é necessário que $0{,}6 \leq |r| \leq 1$.

Informações sobre a próxima aula

Uma vez caracterizada a relação linear, procuramos descrevê-la por meio de uma função matemática. A regressão é o instrumento adequado para a determinação dos parâmetros dessa correlação. É o que iremos estudar na próxima aula.

22 Regressão simples

Meta da aula

Apresentar os conceitos e a metodologia básica que envolve a aplicação da regressão simples.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. avaliar os objetivos da análise de regressão;
2. explicar o procedimento adotado para estimar os parâmetros da regressão;
3. interpretar o significado da estimação dos parâmetros a e d da reta;
4. avaliar as aplicações da análise da regressão simples;
5. definir o conceito e a utilização da regressão simples.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, é necessário ter internalizado e exercitado os conceitos básicos de probabilidades das Aulas 1, 19, 20 e 21.

Introdução

A análise de regressão simples é uma técnica estatística que é utilizada para relacionar duas variáveis. Aqui, uma variável de interesse, a variável dependente ou de resposta (Y), é relacionada a uma variável independente ou explicativa (X). O objetivo da análise de regressão é construir um modelo de regressão ou uma equação de previsão relacionando a variável dependente com uma variável independente.

Nesta aula, vamos discutir a forma mais simples de análise de regressão: a análise de regressão com duas variáveis e a variável de interesse que se relaciona com apenas uma variável independente. A análise de regressão que envolve mais de uma variável independente é chamada regressão múltipla.

Ajustamento da reta

Sempre que desejamos estudar determinada variável em função de outra, fazemos uma análise de regressão.

Podemos dizer que a análise de regressão tem por objetivo descrever, por meio de um modelo matemático, a relação entre duas variáveis, partindo de n observações das mesmas.

A variável sobre a qual desejamos fazer uma estimativa recebe o nome de *variável dependente* e a outra recebe o nome de *variável independente*.

Assim, supondo X a variável independente e Y a dependente, vamos procurar determinar o ajustamento de uma reta à relação entre essas variáveis, ou seja, vamos obter uma função definida por:

$$Y = aX + b$$

Onde a e b são os parâmetros do modelo a ser estimados.

Sejam duas variáveis X e Y, entre as quais exista uma correlação acentuada, embora não perfeita, como, por exemplo, as que formam a **Tabela 22.1**.

Tabela 22.1: Dados para análise de regressão

X	5	8	7	10	6	7	9	3	8	2
Y	6	9	8	10	5	7	8	4	6	2

O diagrama de regressão é dado por:

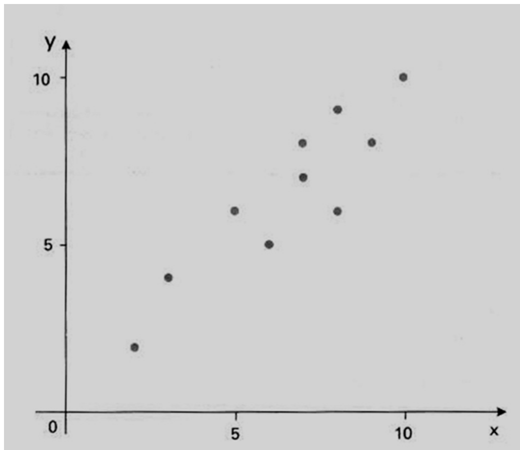


Figura 22.1: Diagrama de dispersão dos dados da **Tabela 22.1**.

Podemos concluir, pela forma do diagrama, que se trata de uma correlação retilínea, de modo a permitir o ajustamento de uma reta, imagem da função definida por:

$$Y = aX + b$$

Vamos, então, calcular os valores dos parâmetros a e b com ajuda das fórmulas:

$$a = \frac{n\sum x_i\, y_i - \sum x_i \sum y_i}{n\sum x^2_i - (\sum x_i)^2}$$

$$b = \overline{Y} - a\overline{X}$$

onde:

n é o número de observações;

$\overline{X}$ é a média dos valores de x_i;

$\overline{Y}$ é a média dos valores de y_i.

Como estamos fazendo uso de uma amostra para obtermos os valores dos parâmetros, o resultado, na realidade, é uma estimativa da verdadeira equação de regressão. Sendo assim, para especificar isso, escrevemos:

$$\hat{Y} = aX + b,$$

onde $\hat{Y}$ é o Y estimado.

Formemos, então, a tabela de valores:

Tabela 22.2: Tabela de valores para o cálculo da regressão

X_i	Y_i	X_iY_i	X^2_i
5	6	30	25
8	9	72	64
7	8	56	49
10	10	100	100
6	5	30	36
7	7	49	49
9	8	72	81
3	4	12	9
8	6	48	64
2	2	4	4
$\Sigma = 65$	$\Sigma = 65$	$\Sigma = 473$	$\Sigma = 481$

Temos assim:

$$a = \frac{10 \times 473 - 6\,5 \times 65}{10 \times 482 - (65)^2} = \frac{505}{585}$$

$$a = 0,8632$$

$$\overline{X} = \frac{65}{10} = 6,5$$

$$\overline{Y} = \frac{65}{10} = 6,5$$

Parte-se então para:

$$b = 6,5 - 0,8632\,\overline{X}\,6,5 = 0,8892$$

onde:

a = 0,86 e b = 0,89.

Logo:

$$\hat{Y} = 0,86X + 0,89.$$

Atividade

Atende aos Objetivos 1, 2 , 3 , 4 e 5

1. Um grupo de turistas fez uma avaliação do peso aparente de alguns objetos típicos de uma região visitada. Com o peso real e a média dos pesos aparentes, dados pelo grupo, obteve-se a seguinte tabela:

Peso real	18	30	42	62	73	97	120
Peso aparente	10	23	33	60	91	98	159

Obtenha a equação da reta.

Resposta Comentada

É necessário realizar a tabela de valores a seguir:

Tabela 22.3: Cálculo da equação de regressão

Peso real (x_i)	**Peso aparente (y_i)**	x_iy_i	x^2_i
18	10	180	324
30	23	690	900
42	33	1.386	1.764
62	60	3.720	3.844
73	91	6.643	5.329
97	98	9.506	9.409
120	159	19.080	14.400
Σ = 442	Σ = 474	Σ = 41.205	Σ = 35.970

Temos assim:

$$a = \frac{7 \times 41.205 - 442 \times 474}{7 \times 35.970 - (442)^2} = \frac{78.927}{56.426}$$

$$a = 1,3988$$

$$X = \frac{442}{7} = 63,1.$$

$$\overline{Y} = \frac{474}{7} = 67,7.$$

Parte-se então para:
$b = 67,7 – 1,3988 \times 63,1 = -20,5643$.
onde:
$a = 1,40$ e $b = -20,56$.
Logo:

$$\hat{Y} = 1,40\ X - 20,56.$$

Se tivermos duas variáveis aleatórias quantitativas, poderemos medir a relação linear entre elas, primeiramente sondando essa relação por meio do diagrama de dispersão e se for significativa, calcular o seu grau de intensidade por meio do coeficiente de correlação linear de Pearson. Uma vez constatada a correlação linear, poderemos descrevê-la por meio de uma equação matemática denominada regressão linear.

Atividade Final

Atende aos Objetivos 1, 2, 3, 4 e 5

Pretendendo-se estudar a relação entre as variáveis "consumo de energia elétrica de agências de turismo (x_i)" e "volume de serviços nas agências de turismo (y_i)", fez-se uma amostragem que inclui vinte agências, computando-se os seguintes valores:

$\Sigma x_i = 11{,}34$, $\Sigma y_i = 20{,}72$, $\Sigma x^2_i = 12{,}16$, $\Sigma y^2_i = 84{,}96$, $\Sigma x_i y_i = 22{,}13$

Obtenha a equação da reta.

Resposta Comentada

Podemos aplicar as fórmulas de estimação dos parâmetros diretamente:

$$a = \frac{20 \times 22{,}13 - 11{,}34 \times 20{,}72}{20 \times 12{,}16 - (11{,}34)^2} = \frac{207{,}6352}{114{,}6044}$$

$$a = 1{,}8118.$$

$$\overline{X} = \frac{11{,}34}{20} = 0{,}567$$

$$\overline{Y} = \frac{20{,}72}{20} = 1{,}038$$

Parte-se então para:

b = 1,038 – 1,8118 x 0,567 = 0,0107

Em que:

a = 1,81 e b = 0,01.

Logo:

$$\hat{Y} = 1,81\,X - 0,01.$$

Resumo

Uma vez caracterizada a relação linear, procuramos descrevê-la por meio de uma função matemática. A regressão é o instrumento adequado para a determinação dos parâmetros dessa correlação.

A variável sobre a qual desejamos fazer uma estimativa recebe o nome de *variável dependente* e a outra recebe o nome de *variável independente*.

A análise de regressão simples é a análise de regressão que é utilizada para relacionar duas variáveis.

Utilizando fórmulas podemos estimar os parâmetros a e b da verdadeira equação de regressão simples.

Informações sobre a próxima aula

Na nossa última aula, discutiremos os conceitos de dados absolutos e dados relativos, essenciais na fase de análise de dados na utilização do método estatístico em pesquisa.

23 Números relativos

Meta da aula

Apresentar os conceitos e os tipos de números relativos.

Objetivos

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

1. conceituar e diferenciar números absolutos e números relativos;
2. identificar a importância e o uso dos dados relativos;
3. conceituar, diferenciar e calcular percentagens, índices, coeficientes e taxas.

Pré-requisitos

Para que você encontre maior facilidade na compreensão desta aula, são necessários os conceitos das Aulas 5, 6 e 7.

Introdução

Os dados dispostos em tabelas e gráficos estatísticos podem ser de duas naturezas: absolutos ou relativos. É importante que o analista de dados saiba o conceito de cada um desses tipos de dados para melhor interpretar os resultados de sua pesquisa, além de ter disponível um poderoso instrumento de interpretação das informações obtidas com o método estatístico, que são os números relativos definidos como percentagens, índices, coeficientes e taxas.

Dados absolutos

Os dados estatísticos resultantes da coleta direta da fonte, sem outra manipulação senão a contagem ou medida, são chamados dados absolutos.

A leitura dos dados absolutos pode ser enfadonha e inexpressiva. Embora esses dados traduzam um resultado exato e fiel, eles não têm a virtude de ressaltar de imediato as suas conclusões numéricas, daí o uso imprescindível que faz a Estatística dos dados relativos.

Dados relativos

Dados relativos são o resultado de comparação por quociente (razões) que se estabelece entre dados absolutos, e têm por finalidade realçar ou facilitar as comparações entre quantidades.

Traduzem-se os dados relativos, em geral, por meio de percentagens, índices, coeficientes e taxas.

Percentagens

Percentagem é a parte proporcional calculada sobre 100 unidades.

Consideremos a série a seguir:

Tabela 23.1: Reservas em hotéis da cidade Y*

Hotéis	Número de reservas
A	19.286
B	1.681
C	234
Total	**21.201**

*Dados fictícios.

Calculemos as percentagens de reservas em cada hotel:

$$\text{Hotel A} \rightarrow \frac{19.286 \times 100}{21.201} = 90,96 = 91,0\%$$

$$\text{Hotel B} \rightarrow \frac{1.681 \times 100}{21.201} = 7,92 = 7,9\%$$

$$\text{Hotel C} \rightarrow \frac{234 \times 100}{21.201} = 1,10 = 1,1\%$$

Com esses dados, podemos formar uma nova coluna na série em estudo.

Tabela 23.2: Reservas em hotéis da cidade Y com percentagens*

Hotéis	Número de reservas	%
A	19.286	91,0
B	1.681	7,9
C	234	1,1
Total	**21.201**	**100,0**

*Dados fictícios.

Os valores dessa nova coluna nos dizem que, de cada 100 reservas dos hotéis considerados da cidade, 91 são para o hotel A, 8, aproximadamente, para o hotel B e 1 no hotel C.

O emprego da percentagem é de grande valia quando é nosso intuito destacar a participação da parte no todo.

Consideremos, agora, a seguinte série.

Tabela 23.3: Reservas em hotéis da cidade Y e W*

Hotéis	Número de reservas cidade Y	Número de reservas cidade W
A	19.286	38.660
B	1.681	3.399
C	234	424
Total	**21.201**	**42.483**

*Dados fictícios.

Qual das cidades tem, comparativamente, maior número de reservas nos hotéis considerados?

Como o número total de reservas é diferente nas duas cidades, não é fácil concluir a respeito usando os dados absolutos. Porém, usando as percentagens, tal tarefa fica bastante facilitada. Assim, acrescentando na **Tabela 23.3** as colunas correspondentes às percentagens, obtemos a seguinte tabela:

Tabela 23.4: Reservas em hotéis da cidade Y e W com percentagens

Hotéis	**Cidade Y**		**Cidade W**	
	Reservas	**%**	**Reservas**	**%**
A	19.286	91,0	38.660	91,0
B	1.681	7,9	3.399	8,0
C	234	1,1	424	1,0
Total	**21.201**	100,0	**42.483**	100,0

A leitura da **Tabela 23.4** permite dizer que os hotéis considerados, comparativamente, contam, praticamente, com o mesmo número de reservas nas duas cidades em que estão instalados.

Do mesmo modo que tomamos 100 para base de comparação, também podemos tomar outro número qualquer, dentre os quais destacamos o número 1. É claro que, supondo o total igual a 1, os dados relativos das parcelas serão todos menores que 1.
Em geral, quando usamos 100 para a base, os dados são arredondados até a primeira casa decimal; e quando tomamos 1 por base, são arredondados até a terceira casa decimal.

Atividade

Atende ao Objetivo 3

1. Complete a tabela a seguir

Pousadas	**Nº de turistas**	**Dados relativos**	
		Por 1	**Por 100**
A	175	0,098	9,8
B	222		
C	202		
D	362		
E	280		
F	540		
Total	1.781	1,000	100,0

Resposta Comentada

Devemos lançar na tabela as percentagens realizadas nas bases 1 e 100:

Pousadas	**Nº de turistas**	**Dados relativos**	
		Por 1	**Por 100**
A	175	0,098	9,8
B	222	0,125	12,5
C	202	0,113	11,3
D	362	0,203	20,3
E	280	0,157	15,7
F	540	0,304	30,4
Total	**1.781**	**1,000**	**100,0**

Índices

Os índices são razões entre duas grandezas tais que uma não inclui a outra.

Exemplos:

$$\text{Densidade demográfica} = \frac{\text{população}}{\text{superfície}}$$

Índices econômicos:

$$\text{População } \textit{per capita} = \frac{\text{Valor total da produção}}{\text{população}}$$

$$\text{Consumo } \textit{per capita} = \frac{\text{consumo do bem ou serviço}}{\text{população}}$$

$$\text{Demanda } \textit{per capita} = \frac{\text{demanda pelo bem ou serviço}}{\text{população}}$$

$$\text{renda } \textit{per capita} = \frac{\text{renda}}{\text{população}}$$

$$\text{receita } \textit{per capita} = \frac{\text{receita}}{\text{população}}$$

Coeficientes

Os coeficientes são razões entre o número de ocorrências e o número total (número de ocorrências e número de não ocorrências).

Exemplos:

$$\text{Coeficiente de natalidade} = \frac{\text{Número de nascimentos}}{\text{População total}}$$

$$\text{Coeficiente de mortalidade} = \frac{\text{Número de óbitos}}{\text{População total}}$$

$$\text{Coeficiente de cancelamento de reservas} = \frac{\text{Número de reservas canceladas}}{\text{Reservas totais}}$$

Taxas

As taxas são os coeficientes multiplicados por uma potência de 10 (10, 100, 1.000 etc.) para tornar o resultado mais inteligível.

Exemplos:

Taxa de mortalidade = coeficiente de mortalidade x 1.000.

Taxa de natalidade = coeficiente de natalidade x 1.000.

Taxa de cancelamento de reservas = coeficiente de cancelamento de reservas x 100.

Atividade

Atende aos Objetivos 2 e 3

2. O hotel A registrou 733.986 reservas no início de um feriadão e 683.816 no fim do feriadão. O hotel B apresentou 436.127 reservas no início do feriadão e 412.457 no fim do feriadão. Qual o hotel que apresentou maior saída de hóspedes?

Reposta Comentada

$$A \rightarrow TSR = \frac{733.986 - 683.816}{733.986} x\, 100 = 0,0683\; x\, 100 = 6,83 = 6,8\%.$$

$$B \rightarrow TSR = \frac{436.127 - 412.457}{436.127} x\, 100 = 0,0542\; x\, 100 = 5,42 = 5,4\%.$$

O hotel que apresentou maior saída de hóspedes foi o A.

Os dados relativos constituem técnicas estatísticas que possibilitam o destaque do que há de mais essencial na informação, o que muitas vezes não acontece de imediato com os dados absolutos. A principal vantagem dos dados na forma relativa é possibilitar comparações de categorias de variáveis, quando o número total em cada categoria for diferente.

Atividade Final

Atende aos Objetivos 2 e 3

Considerando que Minas Gerais, em 1992, apresentou (dados fornecidos pelo IBGE):

população: 15.957.600 habitantes;

superfície: 586.624 km²;

nascimentos: 292.036;

óbitos: 99.282.

Calcule:

a) o índice da densidade demográfica;

b) a taxa de natalidade;

c) a taxa de mortalidade.

Respostas

a) Densidade demográfica $= \dfrac{15.957.600}{586.624} = 27,20$ *habitantes* / km^2.

b) Taxa de natalidade $= \dfrac{292.036}{15.957.600} \times 1.000 = 18,30\%$.

c) Taxa de mortalidade $= \dfrac{99.281}{15.957.600} \times 1.000 = 6,22\%$.

Resumo

Os dados estatísticos resultantes da coleta direta da fonte, sem outra manipulação senão a contagem ou medida, são chamados dados absolutos. Dados relativos são o resultado de comparação por quociente (razões) que se estabelece entre dados absolutos, e têm por finalidade realçar ou facilitar as comparações entre quantidades. Traduzem-se os dados relativos, em geral, por meio de percentagens, índices, coeficientes e taxas.
Percentagem é a parte proporcional calculada sobre 100 unidades. Os índices são razões entre duas grandezas tais que uma não inclui a outra. Os coeficientes são razões entre o número de ocorrências e o número total (número de ocorrências e número de não-ocorrências). As taxas são os coeficientes multiplicados por uma potência de 10 (10, 100, 1.000 etc.) para tornar o resultado mais inteligível. Os dados relativos são importantes na análise de dados, pois permitem de imediato o realce no que há de mais importante no conjunto de informações e permitem comparações entre variáveis de natureza diferente, entre variáveis em unidades de medidas diferentes e entre categorias de variáveis com totais diferentes.

Nota Final

Com esta aula, concluímos o estudo do método estatístico. Esperamos que você possa, assim, estar apto a desenvolver pesquisas quantitativas e realizar análises de dados na sua área do saber com excelência.

Desejo sucesso em sua área de formação e que você possa exercê-la de forma prazerosa e plena.

Referências

Aula 13

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D. A.; BERENSON, M. L.; STEPHAN, D. *Estatística*: teoria e aplicações usando microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5. ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2. ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 14

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. São Paulo: Atual, 1981. 321 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. São Paulo: Atlas, 2002.

TOLEDO. G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 15

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5. ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2 ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 16

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5 ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2 ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 17

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2 ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 18

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5. ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2. ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 19

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. 2. ed. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 20

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5. ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 21

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. Estatística básica. São Paulo: Atlas, 1983. 459 p.

Aula 22

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

Aula 23

GATTI, B. H.; FERRES N. L. *Estatística básica para ciências humanas*. 3. ed. São Paulo: Alfa-ômega, 1978. 163 p.

LEVINE, D.A.; BERENSON, M.L.; STEPHAN, D. *Estatística*: teoria e aplicações usando Microsoft EXCEL em português. Rio de Janeiro: LCT, 1998. 811 p.

MORETTIN, P. A.; BUSSAB, W.O. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2003. 526 p.

NAZARETH, H. *Curso básico de estatística*. São Paulo: Ática, 1996. 160 p.

TIBONI, C. G. *Estatística básica para o curso de turismo*. São Paulo: Atlas, 2003.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. São Paulo: Atlas, 1983. 459 p.

ISBN 978-85-7648-624-4

9 788576 486244

UENF
Universidade Estadual
do Norte Fluminense

UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO
UERJ

Universidade Federal Fluminense
uff

Universidade Federal Rural do Rio de Janeiro

UNIRIO

FAPERJ
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

GOVERNO DO
Rio de Janeiro
SECRETARIA DE
CIÊNCIA E TECNOLOGIA

UAB
UNIVERSIDADE
ABERTA DO BRASIL

Ministério
da Educação

BRASIL
UM PAÍS DE TODOS
GOVERNO FEDERAL