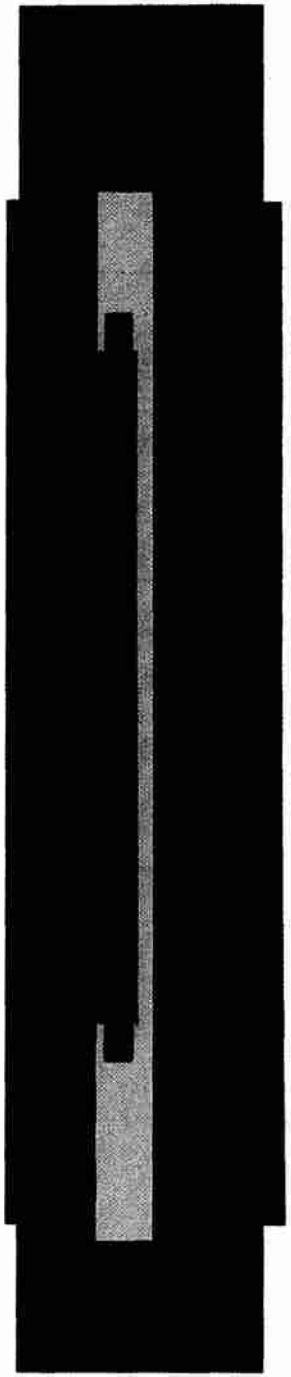
edusp



Probabilidade: Um Curso Introdutório

Carlos A. B. Dantas

1

Noções Básicas de Probabilidade

1.1 EXPERIMENTOS ALEATÓRIOS

Hoje em dia grande quantidade de jogos são oferecidos, entre os quais citamos, por exemplo: a loteria federal, a sena e a loteria esportiva. É natural que se pense nas chances de ganhar um prêmio antes de decidir em qual deles jogar.

Um torcedor de futebol procura avaliar as chances de vitória de seu clube antes de cada jogo de que ele participa. A loteria esportiva foi criada em função do interesse do brasileiro pelo futebol e de sua paixão por jogos. Na loteria esportiva em cada rodada são escolhidos treze jogos e uma aposta consiste da escolha em cada jogo de um dos possíveis resultados, ou seja, vitória de um dos dois clubes ou o empate.

Muitas vezes ao acordar nos perguntamos: será que vai chover? De um modo ou de outro atribuímos um valor à chance de chover e então decidimos o tipo de roupa que usaremos e se levaremos ou não um guarda-chuva conosco.

Pode-se facilmente imaginar uma série de outras situações em que nos deparamos com a incerteza quanto à ocorrência de uma das possíveis alternativas na situação que se está vivenciando. Por exemplo, ao chegar a uma bifurcação em que haja duas opções de trajeto para se dirigir ao local desejado, procura-se avaliar as condições de trânsito nos dois caminhos para decidir-se por um deles.

Um analista de sistemas atribui chances aos possíveis números de usuários que estarão ligados a uma rede durante um certo período. Um engenheiro industrial avalia as chances de um determinado processo encontrar-se em equilíbrio, ou atribui chances para as possíveis proporções de peças defeituosas por ele produzidas. Um médico defronta-se com a incerteza em relação ao efeito provocado pela administração de um novo remédio a um determinado paciente.

Há uma grande classe de experimentos que, ao serem repetidos nas mesmas condições, produzem resultados diferentes. Ou, em outros termos, experimentos que, quando realizados, não apresentam resultados previsíveis de antemão.

DEFINIÇÃO 1.1.1 *Experimentos que ao serem repetidos nas mesmas condições não produzem o mesmo resultado são denominados experimentos aleatórios.*

Damos a seguir alguns exemplos de experimentos aleatórios:

EXEMPLO 1.1.1 Quando retiramos um lote de peças num processo de produção, observamos que o número de peças defeituosas varia de lote para lote.

EXEMPLO 1.1.2 O número de chamadas telefônicas que chegam a uma central em um determinado intervalo de tempo não pode ser determinado de antemão.

EXEMPLO 1.1.3 Se escolhermos uma lâmpada do processo de fabricação e observarmos o seu tempo de duração, verificaremos que esse tempo varia de lâmpada para lâmpada.

EXEMPLO 1.1.4 Se lançarmos um dado sobre uma superfície plana e observarmos o número que aparece na face superior, não poderemos determinar *a priori* qual será esse número.

EXEMPLO 1.1.5 Se selecionarmos um casal de um conjunto de casais e observarmos o sexo do primogênito, não poderemos determiná-lo *a priori*, e o mesmo variará de casal para casal.

DEFINIÇÃO 1.1.2 *Os experimentos que ao serem repetidos nas mesmas condições conduzem ao mesmo resultado são denominados determinísticos.*

Eis alguns exemplos de experimentos determinísticos: Se deixarmos uma pedra cair de uma certa altura, podemos determinar sua posição e velocidade para qualquer instante de tempo posterior à queda. Se aquecermos a água a 100 graus centígrados, ela entrará em ebulição.

Nosso objetivo será construir um modelo matemático para representar experimentos aleatórios. Isto será feito em duas etapas: na primeira descreveremos para cada experimento aleatório o conjunto de seus resultados possíveis, e na segunda procuraremos atribuir pesos a cada resultado que reflitam a sua maior ou menor chance de ocorrer quando o experimento é realizado.

1.2 ESPAÇO AMOSTRAL E EVENTOS

DEFINIÇÃO 1.2.1 *Denominaremos espaço amostral associado a um experimento o conjunto de seus resultados possíveis.*

O espaço amostral será representado por um conjunto S, cujos elementos serão denominados eventos simples ou pontos amostrais. Sempre que o experimento for realizado, suporemos que ocorrerá um e apenas um evento simples.

EXEMPLO 1.2.1 No exemplo 1.1.4, que corresponde ao lançamento de um dado, o espaço amostral é o conjunto: $S = \{1, 2, 3, 4, 5, 6\}$.

EXEMPLO 1.2.2 Uma moeda é lançada duas vezes sobre uma superfície plana. Em cada um dos dois lançamentos pode ocorrer cara (C) ou coroa ($\overline{C}$). O espaço amostral é o conjunto:

$$S = \{C\,C, C\,\overline{C}, \overline{C}\,C, \overline{C}\,\overline{C}\}.$$

EXEMPLO 1.2.3 Três peças são retiradas de uma linha de produção. Cada peça é classificada em boa (B) ou defeituosa (D). O espaço amostral associado a esse experimento é:

$$S = \{BBB, BBD, BDB, BDD, DBB, DBD, DDB, DDD\}.$$

Nos exemplos 1.2.1, 1.2.2 e 1.2.3 o espaço amostral é finito. Apresentaremos a seguir exemplos de experimentos aleatórios cujos espaços amostrais não são finitos.

EXEMPLO 1.2.4 Uma moeda é lançada sucessivamente até que apareça cara pela primeira vez. Se ocorrer cara no primeiro lançamento o experimento termina. Se ocorrer coroa no primeiro lançamento, faz-se um segundo lançamento e se então ocorrer cara o experimento termina. Se não ocorrer cara nos dois primeiros lançamentos, faz-se um terceiro lançamento e caso não ocorra cara, faz-se um quarto lançamento e assim por diante até que ocorra a primeira cara, quando o experimento termina. O espaço amostral é o conjunto:

$$S = \{C, \overline{C}\,C, \overline{C}\,\overline{C}\,C, \ldots\}.$$

Note que os pontos desse espaço amostral podem ser postos em correspondência biunívoca com o conjunto dos números naturais e portanto ele é infinito, porém enumerável.

EXEMPLO 1.2.5 Considere o exemplo 1.1.2 em que observamos o número de chamadas telefônicas que chegam a uma central durante um determinado intervalo de tempo. O espaço amostral é o conjunto: $S = \{0, 1, 2, 3, \ldots\}$.

EXEMPLO 1.2.6 Considere a situação do exemplo 1.1.3 em que observamos o tempo de vida de uma lâmpada. O espaço amostral é o conjunto dos números reais não negativos. Ou seja:

$$S = \{x : x\,real\,, x \geq 0\}.$$

EXEMPLO 1.2.7 A umidade do ar pode ser registrada por meio de um higrômetro. Um higrômetro pode ser acoplado a um dispositivo que possui um ponteiro que desliza sobre papel milimetrado e registra em cada instante a umidade do ar. Se as leituras são feitas no intervalo de tempo $[0, T]$ então o resultado é uma curva que a cada $t \in [0, T]$ associa $x(t)$ que designa a umidade do ar no instante t. É razoável supor-se que $x(t)$ é uma função contínua de t, no intervalo $[0, T]$. O espaço amostral nesse caso é o conjunto:

$$S = \{x : x \text{ é uma função contínua em } [0, T]\}.$$

Nos exemplos 1.2.4 e 1.2.5 o espaço amostral é infinito porém enumerável, isto é, pode ser posto em correspondência biunívoca com o conjunto dos naturais. Nos exemplos 1.2.6 e 1.2.7 o espaço amostral é infinito não enumerável.

Seja S o espaço amostral associado a um experimento aleatório.

DEFINIÇÃO 1.2.2 *Denominaremos de evento a todo resultado ou subconjunto de resultados de um experimento.*

Os eventos serão representados por subconjuntos do espaço amostral. Os eventos representados por um conjunto unitário, isto é, contendo somente um ponto do espaço amostral são denominados eventos simples. Diremos que o evento A ocorre quando o resultado do experimento é um evento simples pertencente a A.

EXEMPLO 1.2.8 No exemplo 1.2.3 consideremos o evento A: duas peças são boas. Tem-se :

$$A = \{BBD, BDB, DBB\}.$$

Então A ocorre se ocorrer um dos três eventos simples BBD, BDB, ou DBB.

EXEMPLO 1.2.9 No exemplo 1.2.4 consideramos o evento A: a primeira cara ocorre em um lançamento que é um múltiplo de 3. Temos então:

$$A = \{\overline{C}\,\overline{C}C, \overline{C}\,\overline{C}\,\overline{C}\,\overline{C}\,\overline{C}C...\}.$$

Os eventos simples de A têm $3n-1$ coroas que precedem a ocorrência da primeira cara na posição $3n$, para $n = 1, 2, \ldots$

EXEMPLO 1.2.10 No exemplo 1.2.2 considere o evento B: o número de caras é igual ao número de coroas,

$$B = \{C\overline{C}, \overline{C}C\}.$$

1.3 OPERAÇÕES ENTRE EVENTOS

DEFINIÇÃO 1.3.1 *A reunião de dois eventos A e B, denotada $A \cup B$, é o evento que ocorre se pelo menos um deles ocorre.*

DEFINIÇÃO 1.3.2 *A interseção de dois eventos A e B, denotada $A \cap B$, é o evento que ocorre se ambos ocorrem.*

DEFINIÇÃO 1.3.3 *O complementar do evento A, denotado A^c, é o evento que ocorre quando A não ocorre.*

Como os eventos são subconjuntos do espaço amostral, podemos representar a reunião, a interseção de dois eventos e o complementar de um evento pelos diagramas utilizados para representar subconjuntos de um dado conjunto.

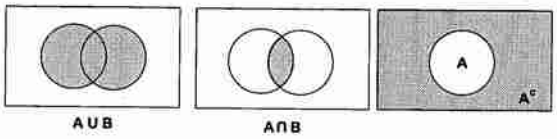


Figura 1.1: União, interseção e complementar de eventos.

EXEMPLO 1.3.1 Uma urna contém bolas numeradas de um a quinze. Uma bola é retirada da urna e seu número anotado. Sejam A e B os

seguintes eventos: A: o número da bola retirada é par, B: o número da bola retirada é múltiplo de 3. Determinemos os eventos $A \cup B$, $A \cap B$ e A^c.

O espaço amostral S associado a esse experimento é o conjunto:

$$S = \{1, 2, \ldots, 15\}.$$

Para A, B, $A \cup B$, $A \cap B$ e A^c temos:

$$A = \{2, 4, 6, 8, 10, 12, 14\} \quad B = \{3, 6, 9, 12, 15\}$$

$$A \cup B = \{2, 3, 4, 6, 8, 9, 10, 12, 14, 15\}$$

$$A \cap B = \{6, 12\} \quad A^c = \{1, 3, 5, 7, 9, 11, 13, 15\}.$$

Dizemos que o evento A implica o evento B, que denotamos $A \subset B$, se para todo $\omega \in A$ tivermos $\omega \in B$. Isto corresponde à situação em que a ocorrência de A garante inevitavelmente a ocorrência de B.

Os eventos A e B são iguais se $A \subset B$ e $B \subset A$.

Os eventos A e B são ditos mutuamente exclusivos, se eles não podem ocorrer simultaneamente. Isto equivale a $A \cap B = \emptyset$.

Apresentamos no próximo lema algumas propriedades dessas operações entre eventos.

LEMA 1.3.1 *Sejam A, B e C eventos do espaço amostral S, temos:*

a) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
b) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
c) $(A \cup B)^c = A^c \cap B^c$
d) $(A \cap B)^c = A^c \cup B^c$.

Vamos demonstrar *a* e *d* e deixar *b* e *c* a cargo do leitor.

Para demonstrar a igualdade em *a* precisamos mostrar que todo elemento pertencente ao lado esquerdo pertence ao lado direito e vice-versa.

Se $\omega \in (A \cup B) \cap C$ então $\omega \in (A \cup B)$ e $\omega \in C$. Daí decorre que ($\omega \in A$ ou $\omega \in B$) e $\omega \in C$ e portanto ($\omega \in A$ e $\omega \in C$) ou ($\omega \in B$ e $\omega \in C$), ou seja, ($\omega \in (A \cap C)$) ou ($\omega \in (B \cap C)$), que

implica $\omega \in (A \cap C) \cup (B \cap C)$. Podemos percorrer estas implicações de trás para frente e verificar que são verdadeiras, donde decorre a igualdade dos conjuntos.

d) Seja $\omega \in (A \cap B)^c$, então $\omega \notin (A \cap B)$, que implica que $\omega \notin A$ ou $\omega \notin B$, que por sua vez implica que $\omega \in A^c$ ou $\omega \in B^c$, isto é, $\omega \in (A^c \cup B^c)$. Partindo de $\omega \in (A^c \cup B^c)$ e fazendo o percurso inverso nós obtemos a igualdade.

Vamos concluir esta seção definindo as operações de uma reunião e de uma interseção enumerável de eventos.

DEFINIÇÃO 1.3.4 *O evento $\cup_{i=1}^{\infty} A_i$ é o evento que ocorre quando pelo menos um dos eventos A_i, para $i = 1, 2, \ldots$ ocorre.*

DEFINIÇÃO 1.3.5 *O evento $\cap_{i=1}^{\infty} A_i$ é o evento que ocorre quando todos os eventos A_i, $i = 1, 2, \ldots$ ocorrerem.*

1.4 DEFINIÇÕES: CLÁSSICA, FREQÜENTISTA E SUBJETIVA DE PROBABILIDADE

As primeiras tentativas de se atribuir probabilidades a eventos aleatórios surgem na Idade Média. Os jogos de dados já eram praticados desde vários milênios antes da era cristã, mas não há menção sobre cálculos associados a chances de ocorrência de resultados dos lançamentos. É na Idade Média, com Galileu, que encontramos a primeira vez o conceito de eventos "igualmente possíveis".

A definição que denominaremos de clássica baseia-se no conceito primitivo de eventos igualmente possíveis.

Consideremos um experimento com número *finito* de eventos simples. Vamos supor que podemos, por alguma razão, uma razão de simetria, por exemplo, atribuir a mesma chance de ocorrência a cada um dos eventos simples desse experimento. Nessas condições adotaremos a seguinte definição de probabilidade.

DEFINIÇÃO 1.4.1 *Consideremos um espaço amostral S com N eventos simples, que suporemos igualmente possíveis. Seja A um evento de S composto de m eventos simples. A probabilidade de A, que denotaremos $P(A)$, é definida por:*

$$P(A) = \frac{m}{N}. \tag{1.1}$$

Observemos que assim definida, a probabilidade é uma função definida na classe dos eventos ou, o que é equivalente, na classe dos subconjuntos do espaço amostral e satisfaz as propriedades estabelecidas no seguinte lema:

LEMA 1.4.1 *Seja S um espaço amostral finito satisfazendo as condições da definição 1.4.1. A probabilidade definida por (1.1) satisfaz:*

i) $P(A) \geq 0$, *para todo* $A \subset S$;

ii) Se A e B são eventos mutuamente exclusivos, então:
$$P(A \cup B) = P(A) + P(B)$$

iii) $P(S) = 1$.

DEMONSTRAÇÃO i) Como $N > 0$ e $m \geq 0$ segue que $P(A) \geq 0$.

Suponha que A tem m_1 eventos simples e que B tem m_2 eventos simples. Como A e B são mutuamente exclusivos segue-se que eles não têm eventos simples comuns, logo o número de eventos simples de $A \cup B$ é $m_1 + m_2$. Usando a definição obtemos ii).

Como o número de eventos simples de S é N, segue da definição que $P(S) = 1$.

EXEMPLO 1.4.1 No experimento que consiste em lançar-se um dado, supondo-se o mesmo balanceado, pode-se atribuir probabilidade $\frac{1}{6}$ a cada um dos eventos simples 1, 2, 3, 4, 5, e 6. O evento "o número obtido quando se lança o dado é par" tem probabilidade 0,5.

Nas situações em que a definição clássica se aplica, para calcular a probabilidade de um evento A precisamos contar o número de eventos simples do espaço amostral e de A. Para facilitar essa tarefa exporemos na próxima seção alguns métodos de contagem.

1.4.1 Definição Freqüentista de Probabilidade

Ao concluirmos que um evento é aleatório, desejamos poder atribuir ao mesmo um número que reflita suas chances de ocorrência quando o experimento é realizado. Vimos acima que em determinadas circunstâncias podemos atribuir a mesma chance a todos os eventos simples associados ao experimento. Quando o número de eventos simples do espaço amostral não for finito, esta possibilidade fica afastada.

Uma outra maneira de determinar a probabilidade de um evento consiste em repetir-se o experimento aleatório, digamos n vezes, e anotar quantas vezes o evento A associado a esse experimento ocorre. Seja $n(A)$ o número de vezes em que evento A ocorreu nas n repetições do experimento. A razão

$$f_{n,A} = \frac{n(A)}{n} \tag{1.2}$$

é denominada freqüência relativa de A nas n repetições do experimento.

Repetindo-se o experimento um grande número de vezes, nas mesmas condições, e de modo que as repetições sucessivas não dependam dos resultados anteriores, observa-se que a freqüência relativa de ocorrências do evento A tende a uma constante p.

A estabilidade da freqüência relativa, para um grande número de observações, foi inicialmente notada em dados demográficos e em resultados de lançamentos de dados.

Buffon, no século XVIII, realizou 4 040 lançamentos de uma moeda e observou a ocorrência de 2 048 caras. A freqüência relativa observada foi 0,5064. Karl Pearson fez 24 000 lançamentos de uma moeda, tendo obtido freqüência relativa de 0,5005 para caras.

Os dados seguintes referem-se ao número de nascimentos durante um ano classificados quanto ao sexo.

Meses	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.
Masc.	3743	3550	4017	4173	4117	3944
Fem.	3537	3407	3866	3711	3775	3665
Total	7280	6957	7883	7884	7892	7609

Meses	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Masc.	3964	3797	3712	3512	3392	3761
Fem.	3621	3596	3491	3391	3160	3371
Total	7585	7393	7203	6903	6552	7132

Para os dados do ano todo, a freqüência relativa de masculino foi $\frac{45682}{88273} = 0,5175$.

Seja S o espaço amostral associado a um experimento aleatório. Considerando-se n repetições desse experimento nas mesmas condições, observemos que a freqüência relativa está definida na classe dos eventos de S e suas propriedades são dadas no seguinte lema:

LEMA 1.4.2 *A freqüência relativa $f_{n,A}$ definida na classe dos eventos do espaço amostral S satisfaz as seguintes condições:*

a) Para todo evento A, $0 \leq f_{n,A} \leq 1$;

b) Se A e B são dois eventos de S mutuamente exclusivos, temos: $f_{n,A\cup B} = f_{n,A} + f_{n,B}$;

c) $f_{n,S} = 1$.

DEMONSTRAÇÃO A parte a decorre do fato que $n(A) \geq 0$. Como os eventos A e B são mutuamente exclusivos, toda vez que um deles ocorre o outro não ocorre e portanto o número de ocorrências de $A \cup B$ é igual a soma do número de ocorrências de A com o número de ocorrências de B, isto é: $n(A \cup B) = n(A) + n(B)$. Dividindo-se por n obtemos b. Como em toda realização do experimento algum ponto de S ocorre, segue-se que c é verdadeira.

Houve tentativas de se definir a probabilidade como limite da freqüência relativa, já que se observava que, como foi mencionado, a freqüência relativa $f_{n,A}$ se aproxima de uma constante quando n tende a infinito. Estes esforços não foram bem-sucedidos. Voltaremos a tratar deste assunto quando, no capítulo de teoremas limites, apresentarmos a lei dos grandes números. Por hora nos contentaremos em mencionar que as propriedades que a definição clássica satisfaz (lema 1.4.1) são também satisfeitas pela freqüência relativa (lema 1.4.2) e servem de base intuitiva para a definição axiomática que apresentaremos a seguir, após apresentarmos a definição subjetiva.

1.4.2 Definição Subjetiva de Probabilidade

A fundamentação freqüentista da probabilidade baseia-se na hipótese de que existe uma realidade física e que as probabilidades descrevem aspectos dessa realidade de modo análogo ao que as leis da mecânica fazem no caso de um modelo determinístico. A probabilidade de um evento associado a um experimento independe, portanto, do observador, sendo obtida como o valor do qual se aproxima a freqüência relativa de ocorrências desse evento em um grande número de repetições do experimento. Há no entanto situações em que a repetição do experimento não pode ser realizada e outras em que não pode ser realizada em idênticas condições. Eis alguns exemplos dessas situações:

a) Um paciente é submetido a um novo tipo de cirurgia e desejamos saber se ele ficará bom.

b) Desejamos saber se haverá um tremor de terra no Rio Grande do Norte no próximo ano.

c) Desejamos saber quem vencerá o próximo jogo entre São Paulo e Palmeiras.

No primeiro exemplo não se pode falar em repetição do experimento, pois, trata-se de uma nova técnica cirúrgica que estará sendo empregada. No segundo, temos notícia de raras ocorrências de tremores de terra no Rio Grande do Norte. No caso do jogo entre São Paulo e Palmeiras, sabemos que há estatísticas de um grande número de jogos entre São Paulo e Palmeiras, mas que as condições entre um jogo e outro variam bastante. Uma corrente de probabilistas considera a probabilidade de um evento como sendo a medida da crença que o observador possui na ocorrência do evento. Desse modo, a probabilidade será em geral diferente para distintas pessoas em decorrência das diferentes opiniões que elas têm sobre a ocorrência do evento. Em uma outra descrição equivalente, a probabilidade de um evento é o valor que cada observador estaria inclinado a apostar na realização do evento.

O leitor que toma contato com esta matéria pela primeira vez pode estranhar essa colocação e preocupar-se com a consistência da teoria que se constrói com essa formulação. Para lidar com essa

questão são estabelecidas regras de comportamento racional para o observador. Desta forma, para que ele seja racional, suas opiniões precisam necessariamente obedecer certas regras, que são justamente os axiomas de probabilidade, apresentados logo a seguir.

Quanto à efetiva aferição das probabilidades subjetivas, utiliza-se em geral um padrão, ou seja, uma unidade de incerteza. Um especialista em sismologia, por exemplo, compara sua "opinião" sobre a ocorrência de um terremoto com a ocorrência de bola branca na retirada de uma bola de uma urna contendo m bolas brancas e n bolas pretas. Sua probabilidade para a ocorrência do terremoto seria então $\frac{m}{n+m}$. A probabilidade assim entendida é denominada subjetiva ou pessoal. Sugerimos ao leitor interessado consultar a bibliografia indicada no final do livro.

1.4.3 Definição Axiomática de Probabilidade

A probabilidade será definida numa classe de eventos do espaço amostral que satisfaz certas propriedades. Todas as operações que definimos entre os eventos conduzem a novos eventos que pertencem a essa classe. Para indicações de uma abordagem rigorosa sugerimos ao leitor consultar a bibliografia.

DEFINIÇÃO 1.4.2 *Probabilidade é uma função definida numa classe $\mathcal{F}$ de eventos de S que satisfaz as seguintes condições:*

a) $P(A) \geq 0$ *para todo* $A \in \mathcal{F}$;

b) Se $(A_n)_{n\geq 1}$ *é uma seqüência de eventos de* $\mathcal{F}$, *que são mutuamente exclusivos, então:*

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n); \tag{1.3}$$

c) $P(S) = 1$.

Observe que a propriedade (ii) do lema 1.4.1 foi substituída pela condição b desta definição.

No que segue vamos considerar espaços amostrais enumeráveis ou que são intervalos ou reunião de intervalos da reta. Em algumas

situações consideraremos subconjuntos do R^n que são generalizações de intervalos.

Se o espaço amostral S é enumerável, podemos definir a probabilidade na classe de todos os subconjuntos de S que é denominado também conjunto das partes de S e denotado por $\mathcal{P}(S)$. Representemos nesse caso o espaço amostral S da seguinte forma $S = \{\omega_1, \omega_2, \ldots\}$. Associemos a cada ω_n, $n = 1, 2, \ldots$, um número $P(\omega_n)$, tal que $P(\omega_n) \geq 0$ e $\sum_{n=1}^{\infty} P(\omega_n) = 1$.

Denominaremos $P(\omega_n)$ de probabilidade do evento simples ω_n, $n = 1, 2, \ldots$.

DEFINIÇÃO 1.4.3 *Seja S um espaço amostral enumerável e seja A um subconjunto de S. A probabilidade de A é definida da seguinte maneira:*

$$P(A) = \sum_{n:\omega_n \in A} P(\omega_n). \tag{1.4}$$

Podemos verificar que a probabilidade definida dessa maneira satisfaz os axiomas da definição 1.4.2. Sugerimos ao leitor que verifique esses axiomas quando o espaço amostral é finito.

1.5 MÉTODOS DE CONTAGEM

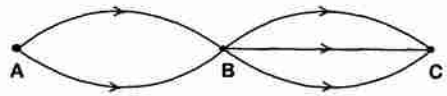
A definição clássica atribuiu a um evento A, composto de M eventos simples, probabilidade $\frac{M}{N}$, onde N é o número de eventos simples do espaço amostral. Para calcularmos a probabilidade de um evento qualquer precisamos portanto contar o número de eventos simples desse evento. Veremos nesta seção alguns métodos de contagem que nos auxiliarão nessa tarefa.

Um procedimento muito elementar de contagem tem sido apresentado sob o título de princípio fundamental da contagem. Nós seguiremos essa tradição e iniciaremos pela apresentação desse princípio.

1.5.1 Princípio Fundamental da Contagem

Suponhamos que uma tarefa pode ser executada em duas etapas. Se a primeira etapa pode ser realizada de n maneiras e a segunda etapa de m maneiras então a tarefa completa pode ser executada de $n.m$ maneiras.

EXEMPLO 1.5.1 Desejamos ir da cidade A à cidade C. Os caminhos de A a C passam pela cidade B. Se há dois caminhos que ligam A a B e três caminhos que ligam B a C, de quantas maneiras podemos ir de A a C ? O número de caminhos que ligam A a C é seis. Se designarmos por 1 e 2 os caminhos que ligam A a B e por 3, 4 e 5 os caminhos que ligam B a C, então os seis caminhos que ligam A a C são: 13, 14, 15, 23, 24, 25.



DEFINIÇÃO 1.5.1 *Uma amostra de tamanho n de um conjunto C que tem N elementos é um subconjunto de n elementos retirados de C.*

As amostras podem ser retiradas de um conjunto de duas maneiras: com reposição ou sem reposição. Nas amostras retiradas com reposição cada elemento selecionado é reposto no conjunto antes da próxima retirada. No caso de amostras sem reposição, como o nome diz, os elementos não são repostos após cada retirada.

Os elementos da amostra poderão ainda ser ordenados ou não.

DEFINIÇÃO 1.5.2 *Uma amostra é dita ordenada se os seus elementos forem ordenados, isto é, se duas amostras com os mesmos elementos, porém em ordens distintas, forem consideradas diferentes.*

EXEMPLO 1.5.2 Considere uma classe com vinte estudantes. O conselho de classe é formado por três estudantes: um presidente, um secretário e um tesoureiro. Ao escolhermos uma amostra de três estudantes para formarem o conselho, deveremos considerar as amostras ordenadas, pois ainda que duas amostras sejam formadas pelas mesmas pessoas, se elas executam tarefas distintas, devem ser consideradas como diferentes.

As amostras não ordenadas sem reposição, de tamanho n de um conjunto com N elementos, são denominadas na maioria dos textos elementares de probabilidade ou de combinatória de combinações de N elementos tomados n a n. Quando não for estabelecida nenhuma qualificação, estaremos admitindo que os elementos são todos distintos e que a amostra é não ordenada. As amostras ordenadas sem reposição são denominadas arranjos. Utilizaremos tanto um nome como outro.

Vamos agora determinar o número de amostras de cada tipo.

LEMA 1.5.1 *O número de amostras ordenadas sem reposição de tamanho n, de um conjunto com N elementos, que será denotado por $(N)_n$, é dado por:*

$$(N)_n = N(N-1)...(N-n+1). \tag{1.5}$$

DEMONSTRAÇÃO As amostras são retiradas sem reposição, portanto o primeiro elemento da amostra pode ser retirado de N maneiras, o segundo de $(N-1)$ maneiras, e assim por diante até o n-ésimo que pode ser retirado de $(N-(n-1))$ maneiras. Pelo princípio fundamental da contagem, o número de maneiras de retirar uma amostra de tamanho n é dado pelo produto desses números.

EXEMPLO 1.5.3 No exemplo 1.5.2 o número de maneiras que o conselho de classe pode ser formado é igual ao número de amostras ordenadas sem reposição de tamanho 3 de um conjunto com 20 elementos. Pelo lema temos:

$$(20)_3 = 20.19.18 = 6\,840\,.$$

EXEMPLO 1.5.4 Considere o conjunto das quatro primeiras letras do alfabeto $\{a, b, c, d\}$. O número de amostras ordenadas sem reposição de tamanho 3 é igual a $(4)_3 = 4.3.2 = 24$. Para referência futura nós

aproveitamos para listar essas 24 amostras:

abc	*abd*	*acd*	*bcd*
acb	*adb*	*adc*	*bdc*
bac	*bad*	*cad*	*cbd*
bca	*bda*	*cda*	*cdb*
cab	*dab*	*dac*	*dbc*
cba	*dba*	*dca*	*dcb*.

LEMA 1.5.2 *O número de amostras ordenadas com reposição de tamanho n, de um conjunto com N elementos é igual a N^n.*

DEMONSTRAÇÃO De fato, como após cada retirada o elemento retirado é reposto, então em cada uma das n retiradas temos N escolhas possíveis. Pelo princípio fundamental da contagem o número dessas amostras é N^n.

EXEMPLO 1.5.5 No exemplo 1.5.4 determinamos o número de amostras ordenadas sem reposição, de tamanho 3 do conjunto $\{a, b, c, d\}$. Vamos agora determinar o número de amostras ordenadas com reposição.

O número de amostras de tamanho 3 retiradas com reposição é pelo lema 1.5.2 igual a 4^3. A título de ilustração vamos obter esse número exibindo o conjunto dessas amostras.

As amostras ordenadas de tamanho 3, com reposição, do conjunto $\{a, b, c, d\}$ incluem:

(i) as amostras sem reposição, que coincidem com as amostras com reposição em que não há repetição. O número dessas é 24 e estão listadas no exemplo 1.5.4.

(ii) as amostras onde há pelo menos uma repetição. Vamos listar as amostras em que o elemento "a" aparece repetido pelo menos uma vez

aaa	*aab*	*aac*	*aad*
aba	*aca*	*ada*	
baa	*caa*	*daa*	

O leitor poderá completar o conjunto dessas amostras notando que falta listar as amostras em que aparecem repetidos os elementos

b, c ou d. Nós obtivemos 24 amostras ordenadas sem reposição e dez amostras ordenadas em que aparece repetido o elemento a. O leitor obterá trinta amostras em que aparecem repetidos os elementos b, c, e d. Assim o número de amostras ordenadas com reposição é 64, que coincide com o valor dado pelo lema.

EXEMPLO 1.5.6 Suponha que a data de nascimento de qualquer pessoa pode ser considerada igualmente distribuída entre os 365 dias de um ano. Se em uma sala existem n pessoas, qual é a probabilidade de que todas tenham nascido em dias diferentes?

Denotemos por A esse evento. O número de conjuntos de n dias em que nasceram as n pessoas é igual ao número de amostras ordenadas com reposição de tamanho n de um conjunto com 365 elementos, que é igual a 365^n. Datas distintas de nascimento das n pessoas correspondem a amostras ordenadas sem reposição de tamanho n de um conjunto com 365 elementos, cujo número é $(365)_n$. Assim,

$$P(A) = \frac{(365)_n}{365^n} = (1 - \frac{1}{365})(1 - \frac{2}{365}) \cdots (1 - \frac{n-1}{365})$$

Pode-se mostrar que para $n = 23, P(A) < \frac{1}{2}$, e portanto a probabilidade do complementar de A, que é o evento "pelo menos duas pessoas fazem aniversário no mesmo dia", é maior que $\frac{1}{2}$. Observemos que o problema que resolvemos é um caso particular do seguinte: Se de um conjunto de N elementos retirarmos amostras ordenadas com reposição de tamanho n, qual é a probabilidade de que a amostra não tenha nenhum elemento repetido? De fato, as amostras em que não há elemento repetido são as amostras ordenadas sem reposição. Se p denota essa probabilidade, então:

$$p = \frac{(N)_n}{N^n} . \tag{1.6}$$

EXEMPLO 1.5.7 Um icosaedro regular tem suas faces numeradas de 1 a 20. O icosaedro é lançado quatro vezes. Qual é a probabilidade de que não apareça nenhuma face repetida?

O espaço amostral pode ser considerado como o conjunto das amostras ordenadas com reposição de tamanho 4 de um conjunto com 20 elementos. Não ocorrer repetição equivale a retirar amostras ordenadas sem reposição. A razão do número dessas amostras nos dá a probabilidade p procurada.

$$p = \frac{(20)_4}{20^4} = 0,0363.$$

DEFINIÇÃO 1.5.3 *Uma amostra ordenada sem reposição de tamanho n de um conjunto com n elementos será denominada uma permutação dos n elementos.*

LEMA 1.5.3 *O número de permutações de n elementos, denotado P_n, é dado por:*

$$P_n = n! \tag{1.7}$$

onde $n! = 1.2 \ldots n$.

DEMONSTRAÇÃO Basta substituir N por n na expressão de $(N)_n$ dada pela fórmula 1.5.

EXEMPLO 1.5.8 Considere o conjunto dos inteiros de 1 a 3. O número de permutações desse conjunto é $P_3 = 6$ e as permutações são as seguintes:

123, 132, 213, 231, 312, e 321.

DEFINIÇÃO 1.5.4 *Uma amostra é dita não ordenada se os seus elementos não forem ordenados, assim uma amostra não ordenada de tamanho n coincide com um subconjunto de tamanho n.*

Uma amostra não ordenada, de tamanho n, sem reposição, de um conjunto com N elementos será também, como mencionamos, denominada uma combinação de N elementos tomados n a n. O número dessas amostras será denotado $C_{N,n}$.

LEMA 1.5.4 *O número de amostras não ordenadas sem reposição de tamanho n, de um conjunto com N elementos, é dado por:*

$$C_{N,n} = \frac{(N)_n}{P_n}. \tag{1.8}$$

DEMONSTRAÇÃO Vamos designar o conjunto de N elementos por $\{a_1, a_2, \ldots, a_N\}$. Uma amostra não ordenada sem reposição de tamanho n é um subconjunto desse conjunto com n elementos. Consideremos, por exemplo, a amostra de tamanho n composta pelos elementos: $a_1, a_2, \ldots, a_n$. Esta amostra pode gerar $n!$ amostras ordenadas sem reposição. Como isto é válido para qualquer amostra não ordenada e o número dessas é $C_{N,n}$ temos:

$$(N)_n = C_{N,n}.P_n$$

$$C_{N,n} = \frac{(N)_n}{P_n} = \frac{N(N-1)\ldots(N-n+1)}{n!}.$$

Multiplicando-se o numerador e o denominador por $(N-n)!$ podemos reescrever $C_{N,n}$ da seguinte forma:

$$C_{N,n} = \frac{N!}{n!(N-n)!} \tag{1.9}$$

que é o coeficiente binomial $\binom{N}{n}$.

EXEMPLO 1.5.9 Seis times participam de um torneio de basquete. Cada uma das equipes enfrenta todas as demais. Quantos jogos serão realizados?

Para determinar o número de jogos, precisamos calcular o número de amostras não ordenadas de tamanho 2 de um conjunto com 6 elementos. Pela fórmula (1.8) temos:

$$C_{6,2} = \frac{6.5}{2} = 15.$$

Como mencionamos acima, $C_{N,n}$ coincide com o coeficiente binomial $\binom{N}{n}$. Este nome deriva do fato que $\binom{N}{n}$ é o coeficiente de $a^n b^{N-n}$ na expansão do binômio $(a+b)^N$. De fato, temos:

$$(a+b)^N = \sum_{n=0}^{N} \binom{N}{n} a^n b^{N-n}. \tag{1.10}$$

Deixamos a cargo do leitor provar essa igualdade e as seguintes propriedades dos coeficientes binomiais:

$$\binom{N}{n} = \binom{N}{N-n} \; ; \; \binom{N}{n} = \binom{N-1}{n-1} + \binom{N-1}{n}.$$

EXEMPLO 1.5.10 Uma comissão formada por três estudantes deve ser escolhida em uma classe de vinte estudantes para organizar os jogos interclasses. De quantas maneiras essa comissão pode ser escolhida?

Como a comissão deve ter três membros distintos, as amostras devem ser selecionadas sem reposição, e como a ordem da escolha dos participantes é irrelevante, trata-se de amostras não ordenadas. Pela fórmula (1.8) temos:

$$C_{20,3} = \frac{(20)_3}{3!} = \frac{20.19.18}{6} = 1\,140.$$

Quando selecionamos uma amostra não ordenada de tamanho n de um conjunto com N elementos, nós particionamos esse conjunto em dois subconjuntos; um com n elementos e o outro com $N-n$ elementos. Vimos que isto pode ser feito de $\binom{N}{n}$ maneiras. É natural que procuremos resolver o seguinte problema: Dado um conjunto com N elementos, de quantas maneiras podemos particioná-lo em k subconjuntos, de modo que o primeiro tenha n_1 elementos, o segundo tenha n_2 elementos, ... e o k-ésimo tenha n_k elementos, com $n_1 + n_2 + \cdots + n_k = N$.

Em vez de dar a resposta no caso geral, vamos considerar o exemplo que corresponde à distribuição de cartas no jogo do bridge. Nesse jogo participam quatro jogadores e as 52 cartas do baralho são distribuídas entre eles. Trata-se portanto de particionar um conjunto com 52 elementos em quatro subconjuntos, cada um com 13 elementos. Vamos então determinar de quantas maneiras isto pode ser feito.

Podemos selecionar 13 cartas a serem dadas ao primeiro jogador de tantas maneiras quantas são as amostras não ordenadas sem reposição (combinações) de tamanho 13 de um conjunto com 52 elementos. Esse número é igual a $\binom{52}{13}$. Restam no baralho 39 cartas. Podemos escolher 13 cartas para o segundo jogador de $\binom{39}{13}$ maneiras. Das 26 restantes, podemos escolher 13 para o terceiro jogador de $\binom{26}{13}$ maneiras. As 13 restantes irão para o quarto jogador. Assim o número de maneiras de particionar as 52 cartas do baralho em

quatro subconjuntos com 13 cartas em cada um deles é:

$$\binom{52}{13}\binom{39}{13}\binom{26}{13}.$$

O mesmo raciocínio nos fornecerá a prova do lema que enunciamos a seguir:

LEMA 1.5.5 *O número de partições de um conjunto de N elementos em k subconjuntos, com $n_1, n_2, \ldots, n_k$ elementos, respectivamente, é igual a:*

$$\frac{N!}{n_1!n_2!...n_k!}. \tag{1.11}$$

DEMONSTRAÇÃO A linha de raciocínio é a mesma do exemplo. Nós selecionamos inicialmente do conjunto com N elementos um subconjunto de tamanho n_1. Do conjunto remanescente com $N - n_1$ elementos selecionamos um subconjunto com n_2 elementos. Dos $N-(n_1+n_2)$ retiramos n_3 e assim sucessivamente até que na última etapa restam n_k elementos e o processo termina. A primeira retirada pode ser feita de $\binom{N}{n_1}$ maneiras, a segunda de $\binom{N-n_1}{n_2}$, e assim por diante, sendo que a última pode ser feita de $\binom{N-(n_1+n_2+...+n_{k-2})}{n_{k-1}}$ maneiras. Pelo princípio fundamental de contagem, o número de maneiras de retirar $n_1, n_2, \ldots, n_{k-1}$, restando n_k elementos para o último subconjunto, é igual ao produto:

$$\binom{N}{n_1}\binom{N-n_1}{n_2}\cdots\binom{N-(n_1+n_2+\cdots+n_{k-2})}{n_{k-1}}.$$

Substituindo nessa fórmula cada coeficiente binomial pela sua expressão em termos dos fatoriais obtemos:

$$\frac{N!}{n_1!(N-n_1)!}\cdot\frac{(N-n_1)!}{n_2!(N-(n_1+n_2))!}\cdots$$

$$\cdots\frac{(N-(n_1+n_2+\cdots+n_{k-2}))!}{n_{k-1}!(N-(n_1+n_2+\cdots+n_{k-1}))!}$$

$$=\frac{N!}{n_1!n_2!..n_k!}.$$

EXEMPLO 1.5.11 No jogo de pôquer com quatro participantes é comum usar-se 32 cartas. As cartas pertencem a um de quatro naipes, a saber: paus, espadas, ouros e copas. As denominações das cartas são: sete, oito, nove, dez, valete, dama, rei e ás. Numa primeira etapa são dadas cinco cartas a cada jogador. Consideremos as cartas dadas a um jogador na primeira etapa. Qual é a probabilidade de que ele receba um par de ases?

Vamos considerar o espaço amostral como o conjunto das amostras não ordenadas sem reposição, de tamanho 5, de um conjunto com 32 elementos. Pelo lema 1.5.4 o número de pontos do espaço amostral é $\binom{32}{5}$.

No pôquer, dizer que o jogador recebe um par de ases quer dizer que ele só tem um par e três outras cartas distintas. Vamos designar por A o evento "o jogador recebe um par de ases". Para calcular o número de pontos de A, observemos que podemos selecionar os dois ases de $\binom{4}{2}$ maneiras; podemos selecionar as outras três cartas que não devem ser ases e ser distintas de $\binom{7}{3}\binom{4}{1}^3$ maneiras. Esse produto corresponde a selecionar três denominações do conjunto {sete, oito, nove, dez, valete, dama, rei} e de cada uma das quatro cartas das denominações selecionadas, escolher uma carta. Assim:

$$P(A)=\frac{\binom{4}{2}\binom{7}{3}\binom{4}{1}^3}{\binom{32}{5}}=0,11.$$

Observemos que alternativamente nós podemos adotar para espaço amostral associado à distribuição de cartas na primeira etapa o conjunto das partições de 32 cartas em cinco subconjuntos, quatro com cinco cartas cada um, correspondendo às cartas entregues aos quatro jogadores, e um quinto subconjunto com 12 cartas que são as que permanecem sem serem distribuídas. Vamos determinar o número de pontos do espaço amostral e calcular a probabilidade de que um jogador tenha um par de ases. O número dessas partições pelo lema 1.5.5 é igual a: $\frac{(32)!}{(12)!(5!)^4}$. Consideremos agora as partições em que um jogador especificado recebe um par de ases. Uma mão com um par de ases e outras três cartas diferentes do ás e diferentes

entre si pode ser escolhida de $\binom{4}{2}\binom{7}{3}\binom{4}{1}^3$ maneiras. Para cada mão desse jogador com o par de ases o número de maneiras que os três outros jogadores podem receber suas cartas é igual ao número de partições de um conjunto de 27 cartas em quatro subconjuntos, um deles com doze cartas e os outros três com cinco cartas cada um. O número dessas partições é igual a $\frac{27!}{12!(5!)^3}$. Para a probabilidade de obtermos um par de ases, obtemos:

$$\frac{\binom{4}{2}\binom{7}{3}\binom{4}{1}^3 (27)!}{\frac{(12)!(5!)^3}{(32)!}} = \frac{\binom{4}{2}\binom{7}{3}\binom{4}{1}^3}{\binom{32}{5}}$$

que é o mesmo valor que obtivemos acima.

1.6 PROPRIEDADES DA PROBABILIDADE

Vamos apresentar nesta seção algumas das propriedades da probabilidade que decorrem diretamente ou quase diretamente dos axiomas da definição 1.4.2.

LEMA 1.6.1 *Denotemos por ϕ o evento impossível. Temos $P(\phi) = 0$.*

DEMONSTRAÇÃO Seja A um evento de S de probabilidade positiva; seja ϕ o evento impossível, podemos exprimir o evento A da seguinte maneira: $A = A \bigcup_{i=1}^{\infty} \phi_i$, onde para todo $i \geq 1, \phi_i = \phi$.

Então pelo axioma b da definição 1.4.2 segue que:

$$P(A) = P(A) + \sum_{i=1}^{\infty} P(\phi).$$

Subtraindo $P(A)$ de ambos os membros, segue-se que a igualdade acima só faz sentido se $P(\phi) = 0$.

LEMA 1.6.2 *Se os eventos $A_1, A_2, \ldots, A_n$ são mutuamente exclusivos, então:*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i). \tag{1.12}$$

DEMONSTRAÇÃO Basta considerarmos a seqüência $A_1, A_2, \ldots, A_n$, $A_k = \phi, k \geq n+1$ e aplicar b da definição 1.4.2. Como pelo lema anterior $P(\phi) = 0$ o resultado segue.

LEMA 1.6.3 *Se A^c é o complementar do evento A, então:*

$$P(A^c) = 1 - P(A). \tag{1.13}$$

DEMONSTRAÇÃO A e A^c são eventos mutuamente exclusivos cuja reunião é S. Daí decorre que $P(A) + P(A^c) = 1$. Subtraindo-se $P(A)$ tem-se o resultado do lema.

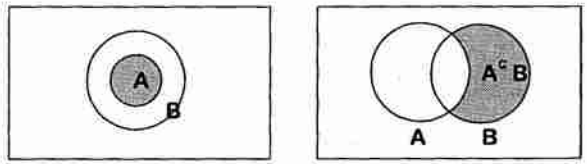
LEMA 1.6.4 *Sejam A e B dois eventos do espaço amostral S, tais que $A \subset B$, tem-se:*

$$P(A) \leq P(B). \tag{1.14}$$

DEMONSTRAÇÃO Como $A \subset B$, segue-se que $B = A \cup A^c B$. Da aditividade da probabilidade (lema 1.6.2) segue que:

$$P(B) = P(A) + P(A^c B)$$

Como $P(A^c B) \geq 0$ concluímos que (1.14) é verdadeira.



LEMA 1.6.5 *Sejam A e B dois eventos quaisquer do espaço amostral S, tem-se:*

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{1.15}$$

DEMONSTRAÇÃO $A \cup B$ pode ser escrito como uma reunião de dois eventos mutuamente exclusivos: $A \cup B = A \cup A^c B$. Do lema 1.6.2 segue que:

$$P(A \cup B) = P(A) + P(A^c B). \tag{1.16}$$

Notando que $B = A^c B \cup AB$, vem:

$$P(B) = P(A^c B) + P(AB).$$

Tirando-se o valor de $P(A^c B)$ e substituindo na expressão (1.16), nós obtemos:

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

EXEMPLO 1.6.1 No exemplo 1.3.1 consideramos um urna contendo bolas numeradas de um a quinze. Suponha que uma bola é retirada da urna ao acaso, isto é, de modo que cada bola tem probabilidade $\frac{1}{15}$ de ser retirada. Calcule as probabilidades dos eventos A e B definidos no exemplo 1.3.1, isto é, A: o número da bola retirada é par; B: o número da bola retirada é múltiplo de 3, e dos eventos AB, $A \cup B$ e $(A \cup B)^c$.

Como os eventos A, B e AB têm respectivamente 7, 5 e 2 pontos e o espaço amostral 15, temos: $P(A) = \frac{7}{15}$, $P(B) = \frac{5}{15}$, $P(AB) = \frac{2}{15}$.

Podemos obter $P(A \cup B)$ diretamente, contando o número de pontos de $A \cup B$ ou utilizando a expressão (1.15). Como $A \cup B$ tem dez pontos, tem-se: $P(A \cup B) = \frac{10}{15}$. Substituindo os valores de $P(A)$, $P(B)$ e de $P(AB)$ em (1.15), obtém-se esse mesmo valor. Finalmente $P((A \cup B)^c) = 1 - P(A \cup B) = \frac{5}{15}$.

No exemplo 1.5.11, calculamos a probabilidade do jogador receber um par de ases. Fixada qualquer denominação, a probabilidade de o jogador receber um par dessa denominação é a mesma que a de receber um par de ases.

EXEMPLO 1.6.2 Nas condições do Exemplo 1.5.11, calcular a probabilidade de o jogador receber um par de ases ou um par de reis.

Vamos designar por A o evento "receber um par de ases" e por B o evento "receber um par de reis". O evento "receber um par de ases ou um par de reis" é o evento $A \cup B$. Notemos que pela definição de "um par" no jogo de pôquer o jogador somente pode receber um par de uma certa carta e três outras cartas distintas entre si e daquelas do par. Assim os eventos "um par de ases" e "um par de reis" são mutuamente exclusivos e portanto

$$P(AB) = 0.$$

Pela expressão (1.15) temos então:

$$P(A \cup B) = \frac{\binom{4}{2}\binom{7}{3}\binom{4}{1}^3}{\binom{32}{5}} + \frac{\binom{4}{2}\binom{7}{3}\binom{4}{1}^3}{\binom{32}{5}}.$$

LEMA 1.6.6 *Sejam $A_1, A_2, \ldots, A_n$ eventos do espaço amostral S, temos:*

$$P(\bigcup_{k=1}^{n} A_k) = 1 - P(\bigcap_{k=1}^{n} A_k^c), \tag{1.17}$$

$$P(\bigcup_{k=1}^{n} A_k) = P(A_1) + P(A_1^c A_2) + \cdots + P(A_1^c A_2^c \cdots A_{n-1}^c A_n), \tag{1.18}$$

$$P(\bigcup_{k=1}^{n} A_k) \leq P(A_1) + P(A_2) + \cdots + P(A_n). \tag{1.19}$$

DEMONSTRAÇÃO A relação:

$$(\bigcup_{k=1}^{n} A_k)^c = \bigcap_{k=1}^{n} A_k^c$$

é uma extensão da propriedade c do lema 1.3.1. Temos:

$$P(\bigcup_{k=1}^{n} A_k)^c = P(\bigcap_{k=1}^{n} A_k^c).$$

Pelo lema 1.6.3 segue que:

$$1 - P(\bigcup_{k=1}^{n} A_k) = P(\bigcap_{k=1}^{n} A_k^c),$$

donde segue a expressão (1.17).

Mostremos que

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_1^c A_2 \cup \cdots \cup A_1^c A_2^c \cdots A_{n-1}^c A_n \qquad (1.20)$$

e que os eventos do lado direito da igualdade são mutuamente exclusivos.

Considere $x \in (\bigcup_{k=1}^{n} A_k)$. Então $x \in A_k$, para algum $k, 1 \leq k \leq n$. Tome o menor k tal que $x \in A_k$, então $x \in A_1^c A_2^c .. A_{k-1}^c A_k$. Logo x pertence ao conjunto do lado direito de (1.20).

Tomemos dois eventos quaisquer do lado direito de (1.20) e verifiquemos que são mutuamente exclusivos. Sejam i e j inteiros tais que: $1 \leq i < j \leq n$. Consideremos $A_1^c A_2^c \cdots A_{i-1}^c A_i$ e $A_1^c A_2^c \cdots A_{j-1}^c A_j$. Note que o primeiro desses conjuntos está contido em A_i e o segundo está contido em A_i^c e portanto eles são mutuamente exclusivos. Tomemos agora x pertencente ao lado direito em (1.20). Então para algum k, $1 \leq k \leq n$, $x \in A_1^c A_2^c \cdots A_{k-1}^c A_k$. Logo $x \in A_k$ e portanto $x \in \bigcup_{i=1}^{n} A_i$.

Para provar (1.18) basta calcular a probabilidade de ambos os lados da igualdade (1.20). Como os eventos da união no lado direito são mutuamente exclusivos, o resultado decorre do lema 1.6.2.

A desigualdade (1.19) segue diretamente da fórmula (1.18), pois como foi visto acima $A_1^c A_2^c \cdots A_{i-1}^c A_i \subset A_i$, e portanto

$$P(A_1^c A_2^c \cdots A_{i-1}^c A_i) \leq P(A_i) \text{ , para todo } i = 2, 3, \cdots, n.$$

Somando-se obtemos o resultado desejado.

EXEMPLO 1.6.3 Uma moeda balanceada é lançada dez vezes. Vamos calcular a probabilidade que ocorra cara em pelo menos um dos lançamentos de números 2, 4, 6 ou 8.

Vamos designar que o evento cara ocorra no lançamento $2k$ por A_{2k}, para $k = 1, 2, 3, 4$. O evento de interesse é: $\bigcup_{k=1}^{4} A_{2k}$. Vamos utilizar a expressão (1.16) para calcular a probabilidade desse evento. Notemos que o evento $\bigcap_{k=1}^{4} A_{2k}^c$ corresponde à ocorrência de coroa nos lançamentos de números 2, 4, 6 e 8. O número de pontos do espaço amostral que satisfazem esta condição é 2^6, pois, nós fixamos coroa nesses quatro lançamentos e nos outros seis temos duas possibilidades em cada um deles. Deste modo:

$$P(\bigcup_{k=1}^{4} A_{2k}) = 1 - P(\bigcap_{k=1}^{4} A_{2k}^c) = 1 - \frac{2^6}{2^{10}} = 0,0625.$$

Antes de encerrarmos esta seção vamos enunciar a generalização do lema 1.6.5 para uma reunião de um número inteiro positivo de eventos do espaço amostral.

LEMA 1.6.7 *Sejam $A_1, A_2, \ldots, A_n$ eventos de um espaço amostral onde está definida uma probabilidade P, temos:*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n} \sum_{j>i} P(A_i A_j) +$$

$$+ \sum_{i=1}^{n} \sum_{j>i} \sum_{k>j} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n) \qquad (1.21)$$

Não daremos aqui a prova desse lema.

1.7 PROBABILIDADE CONDICIONAL

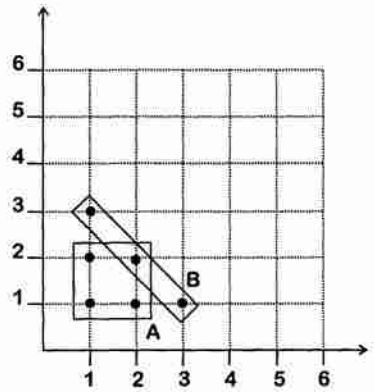
Os conceitos de probabilidade condicional e de independência de eventos são conceitos típicos da Teoria das Probabilidades e que servem para distingui-la de outros ramos da Matemática. Vamos introduzir o conceito de probabilidade condicional considerando uma situação especial em que o espaço amostral tem eventos equiprováveis.

Vamos considerar o experimento que consiste em lançar um dado duas vezes em uma superfície plana e observar o número de pontos na face superior do dado em cada um dos lançamentos. Vamos supor que não se presencie os lançamentos do dado, mas se receba a seguinte informação: "em cada um dos lançamentos, o número de pontos observados é menor ou igual a dois". Vamos denotar por A esse evento. Nestas condições, pergunta-se: qual é a probabilidade de que a soma dos pontos nos dois lançamentos seja igual a quatro? Ou seja, designando por B o evento "soma dos pontos nos dois lançamentos igual a quatro", queremos saber qual é a probabilidade

de ocorrer o evento B, sabendo-se que o evento A ocorreu? Para o espaço amostral associado aos dois lançamentos e para os eventos A e B temos:

$$S = \{(i,j) : i,j \text{ são inteiros } 1 \leq i \leq 6, 1 \leq j \leq 6\}$$

$$B = \{(1,3),(2,2),(3,1)\}\,, \quad A = \{(1,1),(1,2),(2,1),(2,2)\}.$$



Dizer que o evento A ocorreu é equivalente a dizer que pode não se levar em conta qualquer ponto do espaço amostral que não pertença a A, ou seja, pode considerar-se o evento A como novo espaço amostral para o experimento. Desta maneira a probabilidade de B ocorrer dado A é igual a $\frac{1}{4}$, pois dos quatro pontos de A apenas o ponto $(2,2) \in B$ e os quatro pontos são equiprováveis.

Para espaços amostrais com eventos equiprováveis pode-se adotar este procedimento como definição de probabilidade condicional do evento B dado o evento A. Ele serve, no entanto, de motivação para a seguinte definição:

DEFINIÇÃO 1.7.1 *Sejam A e B dois eventos de um espaço amostral e supondo que $P(A) > 0$, a probabilidade condicional de B dado A é definida por:*

$$P(B \mid A) = \frac{P(AB)}{P(A)}\,. \tag{1.22}$$

Retornemos ao exemplo acima. Temos $AB = \{(2,2)\}$ e portanto $P(AB) = \frac{1}{36}$, $P(A) = \frac{4}{36}$. Aplicando-se a fórmula (1.22) obtemos o valor $\frac{1}{4}$, já encontrado acima.

Da fórmula (1.22) que define a probabilidade condicional do evento B dado o evento A, obtemos a seguinte expressão:

$$P(AB) = P(A)P(B \mid A)\,. \tag{1.23}$$

Esta expressão e sua generalização para uma interseção de n eventos permitem construir probabilidades em espaços amostrais que representam experimentos realizados em seqüência, em que a ocorrência de um evento na k-ésima etapa depende das ocorrências nas $k-1$ etapas anteriores. Vejamos inicialmente um exemplo para $n = 2$.

EXEMPLO 1.7.1 Considere uma urna com três bolas brancas e sete bolas vermelhas. Duas bolas são retiradas da urna, uma após a outra sem reposição. Determinar o espaço amostral e as probabilidades associadas a cada ponto amostral.

O espaço amostral é o conjunto $\{B_1B_2, B_1V_2, V_1B_2, V_1V_2\}$

O evento B_1B_2 é o evento que corresponde a ocorrer branca na primeira retirada e branca na segunda. Para os outros pontos do espaço amostral a interpretação é análoga.

Utilizando (1.23) temos:

$$P(B_1B_2) = P(B_1)P(B_2 \mid B_1) = \frac{3}{10}\frac{2}{9} = \frac{2}{30}$$

$$P(B_1V_2) = P(B_1)P(V_2 \mid B_1) = \frac{3}{10}\frac{7}{9} = \frac{7}{30}$$

$$P(V_1B_2) = P(V_1)P(B_2 \mid V_1) = \frac{7}{10}\frac{3}{9} = \frac{7}{30}$$

$$P(V_1V_2) = P(V_1)P(V_2 \mid V_1) = \frac{7}{10}\frac{6}{9} = \frac{14}{30}\,.$$

A fórmula (1.23) pode ser generalizada de modo a exprimir a probabilidade da interseção de n eventos $A_1, A_2, \ldots, A_n$ por meio das probabilidades condicionais sucessivas.

LEMA 1.7.1 *Sejam $A_1, A_2, \ldots, A_n$ eventos do espaço amostral S, onde está definida a probabilidade P, tem-se:*

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\cdots P(A_n \mid A_1A_2\cdots A_{n-1}). \quad (1.24)$$

DEMONSTRAÇÃO Vamos demonstrar por indução. Para $n = 2$ esta fórmula se reduz à fórmula (1.23). Suponha que (1.24) vale para $n-1$ eventos, isto é,

$$P(A_1A_2\cdots A_{n-1}) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\cdots P(A_{n-1} \mid A_1A_2\cdots A_{n-2}). \quad (1.25)$$

Aplicando-se a fórmula (1.22) aos eventos $A_1A_2\cdots A_{n-1}$ e A_n temos:

$$P(A_1A_2\cdots A_{n-1}A_n) = P(A_1A_2\cdots A_{n-1})P(A_n \mid A_1A_2\cdots A_{n-1}).$$

Substituindo nesta igualdade a expressão de $P(A_1A_2\cdots A_{n-1})$ dada por (1.25), obtemos (1.24).

EXEMPLO 1.7.2 Vamos retomar o exemplo 1.7.1 e calcular a probabilidade de obter o seguinte resultado: $B_1B_2V_3V_4B_5$ em cinco retiradas de bolas da urna, sem reposição. O índice representa o número da retirada. Pela fórmula (1.24) temos:

$$\begin{aligned} P(B_1B_2V_3V_4B_5) &= P(B_1)P(B_2 \mid B_1)P(V_3 \mid B_1B_2). \\ &\quad .P(V_4 \mid B_1B_2V_3)P(B_5 \mid B_1B_2V_3V_4) \\ &= \frac{3}{10}\frac{2}{9}\frac{7}{8}\frac{6}{7}\frac{1}{6} = \frac{1}{120}. \end{aligned}$$

As probabilidades condicionais sucessivas foram calculadas levando-se em conta as mudanças na composição da urna após cada retirada. Assim após a primeira retirada, em que saiu branca, a urna fica com duas brancas e sete vermelhas; após a segunda em que também saiu branca, fica com uma branca e sete vermelhas; após a terceira em que saiu vermelha, fica com uma branca e seis vermelhas e assim por diante.

1.7.1 Fórmula das Probabilidades Totais e Fórmula de Bayes

LEMA 1.7.2 *Seja $B_1, B_2, \ldots, B_n$ uma partição do espaço amostral S, isto é, esses eventos são mutuamente exclusivos e sua reunião é S; seja A um evento e P uma probabilidade definida nos eventos de S, temos:*

$$P(A) = \sum_{k=1}^{n} P(A \mid B_k)P(B_k). \quad (1.26)$$

DEMONSTRAÇÃO Como $S = \bigcup_{k=1}^{n} B_k$

$$A = AS = A(\bigcup_{k=1}^{n} B_k) = \bigcup_{k=1}^{n} AB_k$$

Calculando-se a probabilidade de A obtemos:

$$P(A) = \sum_{k=1}^{n} P(AB_k) = \sum_{k=1}^{n} P(A \mid B_k)P(B_k).$$

Esta última igualdade foi obtida exprimindo $P(AB_k)$ pela fórmula (1.23).

A fórmula (1.26) é conhecida como a fórmula das probabilidades totais. Note que ela permite calcular a probabilidade de um evento A quando se conhece as probabilidades de um conjunto de eventos disjuntos cuja reunião é o espaço amostral (uma partição de S) e as probabilidades condicionais de A dado cada um deles.

EXEMPLO 1.7.3 São dadas três urnas com as seguintes composições: a urna um tem três bolas brancas e cinco vermelhas, a urna dois tem quatro bolas brancas e duas vermelhas e a urna três tem uma bola branca e três vermelhas. Escolhe-se uma das três urnas de acordo com as seguintes probabilidades: urna um com probabilidade $\frac{2}{6}$, urna

dois com probabilidade $\frac{3}{6}$ e urna três com probabilidade $\frac{1}{6}$. Uma bola é retirada da urna selecionada. Calculemos a probabilidade da bola escolhida ser branca.

Designemos por A o evento que corresponde a retirar uma bola branca da urna selecionada, e por B_i, para $1 \leq i \leq 3$, o evento "a urna i é selecionada". Temos então $A = A(B_1 \cup B_2 \cup B_3)$.

Utilizando a fórmula (1.26) para $n = 3$ vem:

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3).$$

Os dados do exemplo nos fornecem:

$$P(B_1) = \frac{2}{6},\ P(B_2) = \frac{3}{6},\ P(B_3) = \frac{1}{6};$$

$$P(A \mid B_1) = \frac{3}{8},\ P(A \mid B_2) = \frac{4}{6},\ P(A \mid B_3) = \frac{1}{4}.$$

Substituindo-se esses valores na expressão acima obtemos

$$P(A) = \frac{1}{2}\,.$$

A probabilidade condicional, definida na classe dos eventos do espaço amostral, satisfaz as propriedades estabelecidas no seguinte lema.

LEMA 1.7.3 *Seja A um evento tal que $P(A) > 0$. A probabilidade condicional satisfaz:*

i) Para todo evento B $P(B \mid A) \geq 0$;

ii) Se $B_1, B_2, \ldots, B_n$ são eventos mutuamente exclusivos então:

$$P(\bigcup_{k=1}^{n} B_k \mid A) = \sum_{k=1}^{n} P(B_k \mid A); \tag{1.27}$$

iii) Se S denota o espaço amostral $P(S \mid S) = 1$.

DEMONSTRAÇÃO A parte i decorre imediatamente da definição de probabilidade condicional e do fato da probabilidade de um evento ser sempre não negativa.

A parte ii decorre da definição de probabilidade condicional e da aditividade da probabilidade, expressa pela fórmula (1.12). De fato, sejam $B_1, B_2, \ldots, B_n$ eventos mutuamente exclusivos

$$\begin{aligned} P(\bigcup_{k=1}^{n} B_k \mid A) &= \frac{P((\bigcup_{k=1}^{n} B_k) \cap A)}{P(A)} = \frac{P(\bigcup_{k=1}^{n}(B_k \cap A))}{P(A)} = \\ &= \frac{\sum_{k=1}^{n} P(B_k \cap A)}{P(A)} = \frac{\sum_{k=1}^{n} P(B_k \mid A)P(A)}{P(A)}\,. \end{aligned}$$

A demonstração de iii é imediata, pois, $P(S \mid S) = \frac{P(S)}{P(S)} = 1$.

Vamos agora deduzir a fórmula de Bayes. Como veremos, sua dedução é bem simples, porém permite interpretação bastante profunda e que é responsável pelo desenvolvimento de uma linha de fundamentos da estatística que hoje em dia é denominada Bayesiana.

LEMA 1.7.4 *Seja B um evento e A_1 e A_2 uma partição do espaço amostral S, isto é $A_1 \cap A_2 = \phi$ e $A_1 \cup A_2 = S$. Seja P uma probabilidade definida nos eventos de S. Temos para $i = 1, 2$:*

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)}\,. \tag{1.28}$$

DEMONSTRAÇÃO Pela definição de probabilidade condicional temos:

$$P(A_i \mid B) = \frac{P(A_i B)}{P(B)},$$

mas $P(A_i B) = P(A_i)P(B \mid A_i)$ e expressando $P(B)$ pela fórmula das probabilidades totais (1.26) nós obtemos a fórmula (1.28).

Vamos fazer algumas observações sobre a fórmula de Bayes:

a) Ela permanece válida quando se considera uma partição finita do espaço amostral S, ou seja, se $A_1, A_2, \ldots, A_n$ é uma partição de S, e B é um evento de S, então, para todo $i = 1, 2, \ldots, n$, tem-se:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{k=1}^{n} P(B \mid A_k)P(A_k)}\,. \tag{1.29}$$

b) Como $A_1, A_2, \ldots, A_n$ é uma partição do espaço amostral S, isto é, $\cup_{i=1}^{n} A_i = S$ e $A_i \cap A_j = \phi$, segue-se que $P(A_1) + P(A_2) + \cdots + P(A_n) = 1$. É fácil verificar que $\sum_{i=1}^{n} P(A_i \mid B) = 1$. De fato isto decorre das propriedades *ii* e *iii* do lema 1.7.3.

EXEMPLO 1.7.4 Uma companhia monta rádios cujas peças são produzidas em três de suas fábricas denominadas A_1, A_2 e A_3. Elas produzem, respectivamente, 15%, 35% e 50% do total. As probabilidades das fábricas A_1, A_2 e A_3 produzirem peças defeituosas são 0,01; 0,05 e 0,02, respectivamente. Uma peça é escolhida ao acaso do conjunto das peças produzidas. Essa peça é testada e verifica-se que é defeituosa. Qual é a probabilidade que tenha sido produzida pela fábrica A_i, para $i = 1, 2, 3$?

Designemos por B o evento "a peça é boa" e por D o evento "a peça é defeituosa". Sabemos que: $P(A_1) = 0,15$, $P(A_2) = 0,35$ e que $P(A_3) = 0,5$, pois a peça é escolhida ao acaso do conjunto das peças. Sabemos ainda dos dados do problema que $P(D \mid A_1) = 0,01$, $P(D \mid A_2) = 0,05$ e $P(D \mid A_3) = 0,02$. Desejamos calcular $P(A_i \mid D)$, isto é, a probabilidade condicional de a peça ter sido produzida na fabricada A_i, para $i = 1, 2, 3$, sabendo-se que ela é defeituosa.

Pela fórmula de Bayes temos, para $i = 1, 2, 3$:

$$P(A_i \mid D) = \frac{P(D \mid A_i)P(A_i)}{P(D \mid A_1)P(A_1) + P(D \mid A_2)P(A_2) + P(D \mid A_3)P(A_3)}$$

$$P(D \mid A_1)P(A_1) + P(D \mid A_2)P(A_2) + P(D \mid A_3)P(A_3) =$$

$$= (0,01)(0,15) + (0,05)(0,35) + (0,02)(0,5) = 0,0290.$$

Para $i = 1$, o numerador é $P(D \mid A_1)P(A_1) = 0,0015$, para $i = 2$ é $P(D \mid A_2)P(A_2) = 0,0175$ e para $i = 3$ é $P(D \mid A_3)P(A_3) = 0,01$. Substituindo-se, obtemos:

$$P(A_1 \mid D) = 0,052,\ P(A_2 \mid D) = 0,603,\ P(A_3 \mid D) = 0,345.$$

O uso da fórmula de Bayes nos fornece a seguinte interpretação: como as fábricas A_1, A_2 e A_3 são responsáveis, respectivamente, por 15%, 35% e 50% da produção, se retirarmos uma peça ao acaso da linha de produção, as probabilidades de que essa peça venha das fábricas A_1, A_2 e A_3 são respectivamente iguais a 0,15, 0,35 e 0,5. Por outro lado se retiramos a peça ao acaso e deconhecemos a procedência da peça, isto é, de que fábrica ela veio, e somos informados de que ela é defeituosa, então levando em conta esta informação proveniente do experimento, as probabilidades que a peça tenha vindo de A_1, A_2 ou de A_3 passam a valer, respectivamente, 0,052, 0,603 e 0,345.

1.8 INDEPENDÊNCIA DE EVENTOS

Vamos introduzir a noção de independência para dois eventos e posteriormente estender a definição para um número qualquer de eventos.

DEFINIÇÃO 1.8.1 *Sejam A e B dois eventos e suponha que $P(A) > 0$. O evento B é dito independente do evento A se:*

$$P(B \mid A) = P(B). \tag{1.30}$$

Esta definição corresponde à noção intuitiva da independência do evento B em relação ao evento A, pois diz que a probabilidade de B não se altera com a informação de que o evento A ocorreu.

Se o evento B é independente do evento A decorre da fórmula (1.23) que:

$$P(AB) = P(A)P(B). \tag{1.31}$$

Se o evento B é independente do evento A então nós esperamos que A também seja independente de B. De fato isto ocorre, como é verificado a seguir:

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

A primeira igualdade é a definição de probabilidade condicional, e a segunda decorre da fórmula (1.31).

Se a fórmula (1.31) é válida, então essas igualdades mostram que A é independente de B. Logo A e B são independentes se e somente se a fórmula (1.31) valer. Então podemos adotar essa fórmula

como a definição de independência. Note que esta fórmula, apesar de não ser tão intuitiva como a expressão da definição 1.8.1, é simétrica em relação aos dois eventos, e além disso ela é adequada para a generalização da definição de independência para mais que dois eventos.

EXEMPLO 1.8.1 Vamos considerar uma urna, como no exemplo 1.7.1, que contém três bolas brancas e sete bolas vermelhas. Retiram-se duas bolas da urna, pórem, após a primeira retirada, a bola é reposta na urna.

Vamos supor que as dez bolas são numeradas de 1 a 10, as brancas recebem números de 1 a 3 e as vermelhas de 4 a 10. Podemos considerar como espaço amostral associado a esse experimento o conjunto:

$$S = \{(i,j) : i,j \text{ inteiros }, 1 \leq i \leq 10, 1 \leq j \leq 10\}.$$

Se partirmos da hipótese de que qualquer das 10 bolas tem a mesma chance de ser escolhida, como as duas retiradas são feitas com reposição, podemos associar aos pontos do espaço amostral probabilidade $\frac{1}{100}$. Obtemos para as probabilidades dos eventos, B_1B_2, B_1V_2, V_1B_2, V_1V_2, onde o índice indica a retirada, os seguintes valores:

$$\begin{aligned} P(B_1B_2) &= \frac{9}{100} \\ P(B_1V_2) &= \frac{21}{100} \\ P(V_1B_2) &= \frac{21}{100} \\ P(V_1V_2) &= \frac{49}{100}. \end{aligned}$$

Vamos justificar a primeira igualdade, pois a justificação para as demais é a mesma. Pode-se tirar branca na primeira vez de três maneiras e na segunda retirada também de três maneiras, pois a primeira bola é reposta na urna antes da retirada da segunda. Assim o evento B_1B_2 pode ocorrer de nove maneiras e o espaço amostral

tem 100 pontos. Para os eventos B_1 e B_2 temos:

$$\begin{aligned} P(B_1) &= \frac{30}{100} = \frac{3}{10} \\ P(B_2) &= \frac{30}{100} = \frac{3}{10}. \end{aligned}$$

Observe que tanto B_1 quanto B_2 têm 30 pontos amostrais,

$$\begin{aligned} B_1 &= \{(i,j) : 1 \leq i \leq 3, 1 \leq j \leq 10\} \\ B_2 &= \{(i,j) : 1 \leq i \leq 10, 1 \leq j \leq 3\}. \end{aligned}$$

De modo análogo obtemos $P(V_1)$ e $P(V_2)$. Vemos assim que:

$$\begin{aligned} P(B_1B_2) &= P(B_1)P(B_2) \\ P(B_1V_2) &= P(B_1)P(V_2) \\ P(V_1B_2) &= P(V_1)P(B_2) \\ P(V_1V_2) &= P(V_1)P(V_2). \end{aligned}$$

Esse exemplo nos mostra formalmente que os eventos B_1 e B_2; B_1 e V_2; V_1 e B_2 e V_1 e V_2 são independentes. A seguir partiremos da independência e usaremos a fórmula do produto (1.31) para definir a probabilidade em espaços amostrais associados a experimentos repetidos.

Vamos encontrar com freqüência situações em que um experimento é repetido um número arbitrário de vezes. Se consideramos um evento associado a cada repetição, precisamos generalizar a definição de independência para mais de dois eventos.

Antes de prosseguir, daremos dois exemplos que auxiliarão a determinar as condições que devem ser impostas para definir a independência de três ou mais eventos.

EXEMPLO 1.8.2 Duas moedas são lançadas. Consideremos os seguintes eventos: A_1 – cara no primeiro lançamento, A_2 – cara no segundo lançamento e A_3 – nos dois lançamentos ocorre a mesma face.

O espaço amostral é o conjunto $S = \{CC, C\overline{C}, \overline{C}C, \overline{C}\,\overline{C}\}$, onde C designa cara e $\overline{C}$ coroa.

Suporemos que a moeda é balanceada e portanto que os pontos do espaço amostral têm probabilidade $\frac{1}{4}$.

$$A_1 = \{CC, C\overline{C}\}, \quad A_2 = \{CC, \overline{C}C\}, \quad A_3 = \{CC, \overline{C}\,\overline{C}\},$$
$$A_1A_2 = \{CC\}, \quad A_1A_3 = \{CC\}, \quad A_2A_3 = \{CC\},$$
$$A_1A_2A_3 = \{CC\}.$$

Calculando-se as probabilidades desses eventos, obtemos:

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2},$$

$$P(A_1A_2) = P(A_2A_3) = P(A_1A_3) = \frac{1}{4},$$

$$P(A_1A_2A_3) = \frac{1}{4}.$$

Daí decorre que vale a igualdade (1.31) para os eventos tomados dois a dois, mas

$$P(A_1A_2A_3) \neq P(A_1)P(A_2)P(A_3).$$

EXEMPLO 1.8.3 Lança-se uma moeda três vezes. O espaço amostral é o conjunto das seqüências de caras e coroas de comprimento 3. Consideremos os eventos:

$$A = \{CCC, CC\overline{C}, C\overline{C}C, C\overline{C}\,\overline{C}\}$$
$$B = \{C\overline{C}C, C\overline{C}\,\overline{C}, \overline{C}CC, \overline{C}C\overline{C}\}$$
$$C = \{C\overline{C}\,\overline{C}, \overline{C}CC, \overline{C}C\overline{C}, \overline{C}\,\overline{C}C\}.$$

Decorre daí que: $P(A) = P(B) = P(C) = \frac{1}{2}$

Temos:

$$AB = \{C\overline{C}C, C\overline{C}\,\overline{C}\}$$
$$BC = \{C\overline{C}\,\overline{C}, \overline{C}CC, \overline{C}C\overline{C}\}$$
$$AC = \{C\overline{C}\,\overline{C}\}$$
$$ABC = \{C\overline{C}\,\overline{C}\}.$$

Calculando-se as probabilidades desses eventos, obtemos o seguinte:

$$P(AB) = \frac{1}{4} = P(A)P(B)$$

$$P(AC) = \frac{1}{8} \neq P(A)P(C)$$

$$P(BC) = \frac{3}{8} \neq P(B)P(C)$$

$$P(ABC) = \frac{1}{8} = P(A)P(B)P(C)$$

As igualdades acima mostram que os eventos A e B são independentes, que os eventos B e C e os eventos A e C não são independentes e que A, B e C satisfazem a forma produto.

Sugerimos ao leitor que considere no exemplo 1.8.3 os eventos A_1 – cara no primeiro lançamento, A_2 – cara no segundo lançamento, e A_3 – cara no terceiro lançamento e verifique que, para esses eventos, todas as quatro expressões acima se transformam em igualdades, isto é:

$$P(A_1A_2) = P(A_1)P(A_2)$$

$$P(A_1A_3) = P(A_1)P(A_3)$$

$$P(A_2A_3) = P(A_2)P(A_3)$$

$$P(A_1A_2A_3) = P(A_1)P(A_2)P(A_3). \qquad (1.32)$$

Esta análise mostra que, para definirmos independência de três eventos, devemos impor que essas quatro igualdades valham. Esses exemplos nos sugerem que para definir a independência para n eventos, sendo n um inteiro positivo, devemos exigir a validade da forma produto para todo subconjunto de k dos n eventos, para $2 \leq k \leq n$.

DEFINIÇÃO 1.8.2 *Os eventos $A_1, A_2, \ldots, A_n$ são independentes se:*

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}) \qquad (1.33)$$

para todo $k = 2, 3, \ldots n$ e todo $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$ tal que $i_1 < i_2 < \cdots < i_k$.

O número de equações que devem ser satisfeitas nesta definição é $2^n - n - 1$.

De fato, para todo k, $2 \leq k \leq n$, há uma equação para cada subconjunto de tamanho k dos n eventos, e como o número desses subconjuntos é $\binom{n}{k}$, há para cada k esse número de equações. O total de equações é, portanto:

$$\sum_{k=2}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{0} - \binom{n}{1} = 2^n - n - 1.$$

Para obter esta última igualdade, utilizamos a fórmula (1.10) pondo $a = b = 1$.

1.9 EXERCÍCIOS

1. Defina o espaço amostral para cada um dos seguintes experimentos aleatórios: (a) lançam-se dois dados e anota-se a configuração obtida; (b) conta-se o número de peças defeituosas, no intervalo de uma hora, de uma linha de produção; (c) investigam-se famílias com quatro crianças e anota-se a configuração obtida, segundo o sexo; (d) numa entrevista telefônica com dez assinantes, pergunta-se se o proprietário tem ou não máquina de secar roupa; (e) de um fichário com seis nomes, sendo três de mulheres e três de homens, seleciona-se ficha após ficha até que o último nome de mulher seja selecionado.

2. Uma moeda é lançada três vezes. Descreva o espaço amostral. Considere os eventos A_i: cara no i-ésimo lançamento, para $i = 1, 2, 3$. Determine os seguintes eventos: (a) $A_1^c \cap A_2$; (b) $A_1^c \cup A_2$; (c) $(A_1^c \cap A_2^c)^c$; (d) $A_1 \cap (A_2 \cup A_3)$.

3. Suponha que o espaço amostral é o intervalo $[0, 1]$ dos reais. Considere os eventos $A = [x : 1/4 \leq x \leq 5/8]$ e $B = [x : 1/2 \leq x \leq 7/8]$. Determine os eventos: (a) A^c; (b) $A \cap B^c$; (c) $(A \cup B)^c$; (d) $A^c \cup B$.

4. Quais das seguintes relações são verdadeiras:
(a) $(A \cup B) \cap (A \cup C) = A \cup (A \cap C)$; (b) $A \cup B = (A \cap B^c) \cup B$; (c) $A^c \cap B = A \cup B$; (d) $(A \cup B)^c \cap C = A^c \cap B^c \cap C^c$.

5. Sejam A, B e C três eventos de um espaço amostral. Determine expressões em função de A, B e C para os eventos: (a) somente A ocorre; (b) todos os três eventos ocorrem; (c) pelo menos dois eventos ocorrem; (d) exatamente dois eventos ocorrem; (e) não mais do que dois eventos ocorrem; (f) A e B ocorrem, mas C não ocorre; (g) pelo menos um dos eventos ocorre; (h) exatamente um dos eventos ocorre; (i) nenhum dos eventos ocorre.

6. Prove e interprete as seguintes identidades:
(a) $\binom{n}{r} = \binom{n}{n-r}$; (b) $\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}$.

7. Prove que: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

8. Mostre que:
(a) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$; (b) $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$;
(c) $\sum_{k=r}^{n} \binom{k}{r} = \binom{n+1}{r+1}$; (d) $\sum_{k=0}^{n} \binom{m}{r-k}\binom{n}{k} = \binom{n+m}{r}$, $n \leq r \leq m$;
(e) $\sum_{k=0}^{n} \binom{n}{k} = \binom{2n}{n}$; (f) $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$;
(g) $\sum_{k=0}^{n} k(k-1)\binom{n}{k} = n(n-1)2^{n-2}$;
(h) $\binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{n-1} = \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{n}$, com n par.

9. A, B e C são três eventos de um mesmo espaço amostral, tais que: $P(B) = 0,5$, $P(C) = 0,3$, $P(B \mid C) = 0,4$ e $P[A \mid (B \cap C)] = 0,5$. Calcule $P(A \cap B \cap C)$.

10. Prove que se A e B são dois eventos de um espaço amostral S e P é uma probabilidade definida nos eventos de S, então: $P((A \cap B^c) \cup (B \cap A^c)) = P(A) + P(B) - 2P(A \cap B)$.

11. Sejam A e B dois eventos de um mesmo espaço de probabilidades. Sabendo-se que $P(A) = 0,7$ e $P(B) = 0,6$, determine o valor máximo e mínimo de $P(A \cap B)$.

12. Sejam A, B e C três eventos independentes dois a dois tal que $A \cap B \cap C = \emptyset$. Dado que $P(A) = P(B) = P(C) = x$, determine o maior valor possível de x.

13. Sejam A e B eventos de um mesmo espaço amostral. Se A e B são independentes, prove que A e B^c; A^c e B; A^c e B também são independentes.

14. Sejam A e B dois eventos de um espaço amostral. Suponha que $P(A) = 0,4$ e $P(A \cup B) = 0,7$ e $P(B) = p$. (a) Para que valor de p os eventos A e B são mutuamente exclusivos? (b) Para que valor de p os eventos A e B são independentes?

15. Prove que: (a) Se $P(A^c) = \alpha$ e $P(B^c) = \beta$, então $P(A \cap B) \geq 1 - \alpha - \beta$; (b) Se $P(A \mid B) \geq P(A)$, então $P(B \mid A) \geq P(B)$.

16. Mostre que: $P(E^c \cap F^c) = 1 - P(E) - P(F) + P(E \cap F)$.

17. Uma urna contém duas bolas brancas e duas pretas. As bolas são retiradas ao acaso, sucessivamente e sem reposição. (a) Qual é a probabilidade de sair uma bola preta na primeira retirada? (b) Qual é a probabilidade de que a primeira bola preta apareça somente na quarta retirada? (c) Qual é a probabilidade de que a segunda bola preta apareça logo na segunda retirada? (d) Qual é a probabilidade de que a segunda bola preta apareça somente na quarta retirada?

18. Um dado é viciado de modo que um número par é duas vezes mais provável que um número ímpar. Encontre a probabilidade de que: (a) um número par ocorra; (b) um número primo ocorra; (c) um número primo par ocorra em um lançamento.

19. Um número é escolhido, ao acaso, entre os números inteiros de 1 a 20. Considere os eventos: A : o número escolhido é múltiplo de 3; B: o número escolhido é par.
Descreva os eventos $A \cap B$, $A \cup B$, $A \cap B^c$ e calcule suas probabilidades.

20. No exercício 1 considere os pontos amostrais em a) e c) equiprováveis. Em a) calcule a probabilidade da soma dos pontos dos dois dados ser igual a seis. Em c) calcule a probabilidade de uma família ter três filhos do sexo masculino.

21. Em média, 5% dos produtos vendidos por uma loja são devolvidos. Qual é a probabilidade de que, nas quatro próximas unidades vendidas deste produto, duas sejam devolvidas?

22. Um comitê é formado por quatro homens e duas mulheres. Dois membros do comitê são selecionados sucessivamente, ao acaso e sem reposição. Calcule a probabilidade de cada um dos resultados: HH, HM, MH e MM.

23. Uma cidade tem 30.000 habitantes e três jornais: A, B e C. Uma pesquisa de opinião revela que 12.000 lêem A, 8.000 lêem B; 7.000 lêem A e B; 6.000 lêem C; 4.500 lêem A e C; 1.000 lêem B e C e 500 lêem A, B e C. Selecionamos ao acaso um habitante dessa cidade. Qual a probabilidade de que ele leia: (a) pelo menos um jornal? (b) somente um jornal?

24. É dada a distribuição de 300 estudantes segundo o sexo e a área de concentração:

	Biologia	Exatas	Humanas
Masculino	52	40	58
Feminino	38	32	80

Um estudante é sorteado ao acaso.

(a) Qual é a probabilidade de que ele seja do sexo feminino e da área de humanas? (b) Qual é a probabilidade de que ele seja do sexo masculino e não seja da área de biológicas? (c) Dado que foi sorteado um estudante da área de humanas, qual é a probabilidade de que ele seja do sexo feminino?

25. Um restaurante popular apresenta dois tipos de refeições: salada completa e um prato à base de carne. 20% dos fregueses do sexo masculino preferem salada, e 30% das mulheres preferem carne. 75% dos fregueses são homens. Considere os seguintes eventos:

H : o freguês é homem, M : o freguês é mulher,
A : o freguês prefere salada, B : o freguês prefere carne.

Calcule:

(a) $P(H)$, $P(A \mid H)$ e $P(B \mid H)$;

(b) $P(A \cup H)$ e $P(A \cap H)$;

(c) $P(M \mid A)$.

26. Um indivíduo tem n chaves, das quais somente uma abre uma porta. Ele seleciona, a cada tentativa, uma chave ao acaso sem reposição e tenta abrir a porta. Qual é a probabilidade de que ele abra a porta na k-ésima tentativa ($k = 1, 2, \ldots, n$)?

27. O jogo da loto consiste em selecionar-se cinco dezenas do conjunto de cem dezenas de 00 a 99. Qual a probabilidade de se acertar a quina (5 dezenas) se marcar-se 10 dezenas no volante?

28. Duas cartas são retiradas simultaneamente de um baralho. Qual é a probabilidade de que: (a) ambas sejam de espadas; (b) uma seja de espadas e a outra de copas.

29. Suponha que existam dez livros que devem ser colocados em uma estante, sendo quatro de matemática, três de química, dois de física e um dicionário. Se quisermos que os livros de mesmo assunto fiquem juntos, de quantas maneiras isso será possível?

30. Em um jornal existem dez jornalistas. Se quisermos colocar três jornalistas trabalhando na sede do jornal, cinco em reportagem externa e dois de reserva, de quantas maneiras isso poderá ser feito?

31. Ache a probabilidade de que uma mão de pôquer seja um: (a) *royal flush* (dez, valete, dama, rei e ás do mesmo naipe); (b) quatro de um mesmo tipo (quatro cartas do mesmo valor); (c) trinca e par (um par e uma trinca de cartas do mesmo valor); (d) seguida (cinco cartas em seqüência, não importando o naipe); (e) três cartas do mesmo tipo (três valores iguais mais duas cartas diferentes); (f) dois pares (dois pares do mesmo valor mais uma outra carta); (g) um par (um par de valores iguais mais três cartas diferentes).

32. Uma moeda é lançada até se obter a primeira cara. Determine: (a) a probabilidade de que isto ocorra em um lançamento de número par; (b) a probabilidade de que isto ocorra em um lançamentos de número ímpar.

33. Um dispositivo eletrônico é formado por três partes. Cada parte tem probabilidade de 0,9 de funcionar bem e 0,1 de falhar. O funcionamento de cada parte não depende das demais. O dispositivo falha se duas ou mais falham. Calcule a probabilidade de falha do dispositivo.

34. Três máquinas A, B e C produzem 50%, 30% e 20%, respectivamente, do total de peças de uma fábrica. As porcentagens de produção defeituosa destas máquinas são 3%, 4% e 5%. Se uma peça é selecionada aleatoriamente, ache a probabilidade de ela ser defeituosa. Se a peça selecionada é defeituosa, encontre a probabilidade de ter sido produzida pela máquina C.

35. A probabilidade de que a porta de uma casa esteja trancada à chave é 3/5. Um chaveiro possui 25 chaves das quais três abrem essa porta. Qual a probabilidade de que um indivíduo entre na casa, se ele puder escolher, ao acaso, somente uma chave do chaveiro?

36. Numa urna onde existiam oito bolas brancas e seis azuis, foi perdida uma bola de cor desconhecida. Uma bola foi retirada da urna. Qual é a probabilidade de a bola perdida ser branca, dado que a bola retirada é branca?

37. A probabilidade de que um aluno saiba a resposta de uma questão de um exame de múltipla escolha é p. Há m respostas possíveis para cada questão, das quais apenas uma é correta. Se o aluno não sabe a resposta para uma dada questão, ele escolhe ao acaso uma das m respostas possíveis. (a) Qual é a probabilidade de o aluno responder corretamente uma questão?

(b) Se o aluno respondeu corretamente a questão, qual é a probabilidade de que ele tenha "chutado" a resposta?

38. De quantas maneiras diferentes r bolas distintas podem ser distribuídas, ao acaso, em n urnas numeradas de 1 a n? Qual é a probabilidade de que pelo menos uma urna tenha duas bolas? Qual é a probabilidade de cada urna conter no máximo uma bola?

39. Seleciona-se ao acaso uma amostra casual de tamanho r de uma população de tamanho n, sem reposição. Qual é a probabilidade de que um elemento fixado seja incluído na amostra? Se a amostragem for com reposição, qual é a probabilidade de que um elemento fixado seja incluído pelo menos uma vez na amostra?

40. Distribuindo-se r bolas distintas em n urnas numeradas de 1 a n, qual é a probabilidade de que a urna 1 seja ocupada por r_1 bolas, a urna 2 seja ocupada por r_2 bolas, ... , a urna n seja ocupada por r_n bolas, de modo que $r_1 + r_2 + \cdots + r_n = r$, $r_j \geq 0, j = 1, 2, \ldots, n$?

41. Refaça o exercício 40, supondo bolas idênticas ou indistinguíveis.

42. De uma urna com n bolas numeradas de $1, 2, \ldots, n$, seleciona-se ao acaso e sem reposição todas as bolas, uma de cada vez. Dizemos que ocorre um pareamento na j-ésima seleção $(1 \leq j \leq n)$, se nessa seleção for selecionada a bola de número j. Determine a probabilidade de que (a) ocorra pelo menos um pareamento; (b) não ocorra pareamento algum; (c) ocorram exatamente r pareamentos, $r = 0, 1, \ldots, n$.

43. Suponha que r bolas distintas sejam distribuídas aleatoriamente em n urnas. Determine a probabilidade de que uma urna específica contenha exatamente k bolas, $k = 0, 1, \ldots, r$.

2

Variáveis Aleatórias e Distribuições de Probabilidade

2.1 VARIÁVEIS ALEATÓRIAS

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, também variarão os valores numéricos que lhes são associados. Em alguns experimentos os resultados já são descritos numericamente; no exemplo 1.1.2 observamos o número de chamadas telefônicas que chegam a uma central, que é um inteiro não negativo; no exemplo 1.1.3 observamos o tempo de duração de uma lâmpada, que é um número real não negativo, e no exemplo 1.1.4 o número de pontos que aparecem na face superior de um dado, que é um inteiro de 1 a 6. No entanto, em um grande número de situações, os pontos do espaço amostral não são expressos numericamente. No exemplo 1.1.5 observa-se o sexo do primogênito representado pela letra M para masculino e pela letra F para feminino. Neste caso o espaço amostral é o conjunto $\{M, F\}$. No exemplo 1.2.2 observamos cara (C) ou co-

roa ($\overline{C}$) em cada um dos dois lançamentos de uma moeda; o espaço amostral nesse caso é o conjunto $\{CC, C\overline{C}, \overline{C}C, \overline{C}\,\overline{C}\}$. Podemos estar interessados no número de caras que ocorre nos dois lançamentos, e portanto é natural definirmos uma função que associa a cada ponto desse espaço amostral o seu número de caras. No exemplo 1.2.3 em que retiramos três peças de uma linha de produção e verificamos se cada uma delas é boa (B) ou defeituosa (D) o espaço amostral é o conjunto $\{BBB, BBD, BDB, BDD, DBB, DBD, DDB, DDD\}$. Podemos estar interessados no número de peças defeituosas e não na ordem em que estas ocorrem nos experimentos. É natural que definamos no espaço amostral uma função que associa a cada ponto do espaço amostral o número de peças defeituosas.

EXEMPLO 2.1.1 Consideremos o exemplo 1.2.2, em que uma moeda é lançada duas vezes. Seja X a função definida no espaço amostral que é igual ao número de caras nos dois lançamentos. Temos:

Espaço Amostral	Valores de X
CC	2
$C\overline{C}$	1
$\overline{C}C$	1
$\overline{C}\,\overline{C}$	0

A correspondência entre os pontos do espaço amostral e os valores da variável X também pode ser expressa da seguinte maneira: $X(CC) = 2$, $X(C\overline{C}) = 1, X(\overline{C}C) = 1, X(\overline{C}\,\overline{C}) = 0$.

EXEMPLO 2.1.2 No exemplo 1.2.3 consideremos o número de peças defeituosas, que denominaremos de Y, entre as três peças retiradas da linha de produção. Assim, associando-se a cada ponto do espaço amostral o número de peças defeituosas, tem-se:

Espaço Amostral	Valores de Y
BBB	0
BBD	1
BDB	1
BDD	2
DBB	1
DBD	2
DDB	2
DDD	3

Em probabilidade a função que associa a cada ponto do espaço amostral um número real é denominada uma variável aleatória. Para referência futura, registramos isto na seguinte definição.

DEFINIÇÃO 2.1.1 *Uma variável aleatória é uma função definida num espaço amostral, que assume valores reais.*

Apresentamos a seguir alguns exemplos de variáveis aleatórias. Deixamos a cargo do leitor especificar o espaço amostral em cada caso.

EXEMPLO 2.1.3 a) O número de peças defeituosas entre n peças retiradas de uma linha de produção.

b) O número de partículas radioativas desintegradas em um dado intervalo de tempo.

c) O número de veículos que passam por um posto de pedágio durante uma hora.

d) A duração de um componente de um circuito.

e) O tempo de vida até fadiga de um cabo de aço.

f) O nível de água em uma represa num dado instante.

g) O número de primogênitos do sexo masculino em dez famílias.

Nos exemplos 2.1.1, 2.1.2 e 2.1.3 a e g os valores assumidos pelas variáveis correspondentes pertencem a um conjunto finito de inteiros não negativos. Nos exemplos 2.1.3 b e c o conjunto de valores que as variáveis assumem é o conjunto do inteiros não negativos, que

é enumerável infinito. Nos exemplos 2.1.3 *d*, *e* e *f* os valores das variáveis aleatórias podem ser qualquer real não negativo.

DEFINIÇÃO 2.1.2 *As variáveis aleatórias que assumem valores em um conjunto enumerável serão denominadas discretas e aquelas que assumem valores num intervalo da reta real serão denominadas contínuas.*

Desenvolveremos o estudo das variáveis aleatórias discretas e contínuas separadamente, pois as ferramentas utilizadas para tratalas são diferentes. Em cursos mais avançados é possível dar um tratamento unificado. Nas notas do final do livro daremos indicações bibliográficas de textos onde o leitor poderá encontrar esse tratamento.

2.2 DISTRIBUIÇÃO DE PROBABILIDADE DE VARIÁVEIS ALEATÓRIAS DISCRETAS

Nesta seção nos restringiremos a considerar variáveis aleatórias discretas. Para conhecermos uma variável aleatória discreta, além de conhecermos os seus valores, é preciso que conheçamos também as probabilidades associadas aos mesmos.

DEFINIÇÃO 2.2.1 *A distribuição de probabilidade de uma variável aleatória discreta X, definida em um espaço amostral S, é uma tabela que associa a cada valor de X sua probabilidade.*

EXEMPLO 2.2.1 Vamos determinar a distribuição de probabilidade da variável aleatória X do exemplo 2.1.1. Para cada valor de X determinamos os pontos amostrais em que X é igual a esse valor, em outras palavras determinamos a imagem inversa de cada valor de X. Na tabela a seguir, na primeira coluna estão os valores de X, na segunda as imagens inversas desses valores e na terceira suas probabilidades.

Valores de X	Pontos amostrais	Probabilidades
0	$\overline{C}\,\overline{C}$	1/4
1	$\overline{C}\,C, C\,\overline{C}$	1/2
2	$C\,C$	1/4.

Os valores das probabilidades, na tabela acima, são obtidos da seguinte maneira:

$$P[X=0] = P(\overline{C}\,\overline{C}) = \frac{1}{4}$$

$$P[X=1] = P(\overline{C}\,C) + P(C\,\overline{C}) = \frac{1}{2}$$

$$P[X=2] = P(C\,C) = \frac{1}{4}.$$

EXEMPLO 2.2.2 Para determinarmos a distribuição de probabilidade da variável aleatória Y, que dá o número de peças defeituosas no exemplo 2.1.2, precisamos conhecer a probabilidade definida no espaço amostral. Suponhamos que a fração de peças boas que é produzida é 0,9 e que a fração de defeituosas é portanto 0,1. Suponhamos ainda que as retiradas podem ser consideradas independentes. Na tabela a seguir indicamos na coluna 1, os valores de Y, na coluna 2 os pontos amostrais onde a variável aleatória Y assume os valores correspondentes da coluna 1, e na coluna 3 temos as probabilidades associadas a esses valores.

Valores de Y	Pontos amostrais	Probabilidades
0	BBB	$(0,9)^3$
1	BBD, BDB, DBB	$3\,(0,9)^2(0,1)$
2	BDD, DBD, DDB	$3\,(0,9)(0,1)^2$
3	DDD	$(0,1)^3$

Ilustramos como se calcula $P[Y=2]$. As probabilidades dos outros valores da tabela são calculadas de modo análogo. Temos Y=2 se ocorrer um dos três pontos amostrais indicados na terceira linha da tabela, isto é, BDD, DBD ou DDB. Este conjunto de pontos é a imagem inversa do valor 2 pela variável aleatória Y, assim:

$$P[Y=2] = P(BDD) + P(DBD) + P(DDB) = 3(0,9)(0,1)^2$$

Nesta etapa utilizamos a hipótese de independência, para escrever $P(BDD) = P(B)P(D)^2$, $P(DBD) = P(B)P(D)^2$ e $P(DDB) = P(B)P(D)^2$. Ver a definição 1.8.2

EXEMPLO 2.2.3 Vamos considerar o experimento que consiste em lançar-se um dado sucessivamente sobre uma superfície plana. Vamos observar o número do lançamento em que ocorre 6 pela primeira vez, que denotaremos por X, que é uma variável aleatória cujos valores são os inteiros positivos. Vamos determinar a distribuição de probabilidade de X.

Temos para todo $n \geq 1$:

$$P(X = n) = (\frac{5}{6})^{(n-1)} \frac{1}{6}.$$

De fato, como os lançamentos são independentes, a probabilidade de que não ocorra 6 nos (n-1) primeiros lançamentos e ocorra 6 no n-ésimo lançamento do dado é a dada pela expressão acima.

2.3 DENSIDADES DE PROBABILIDADE

As variáveis aleatórias contínuas, como o tempo de duração de uma chamada telefônica, o tempo de vida de uma lâmpada ou a altura da água em uma represa num dado instante assumem valores na reta ou em intervalos da reta. Não podemos esperar que possamos atribuir probabilidades aos valores de uma variável contínua da mesma maneira que o fizemos para as variáveis discretas, pois a soma de uma quantidade não enumerável de números positivos não poderia ser igual a um.

Assim, em vez de se atribuir, como no caso discreto, probabilidades aos valores da variável, pode-se atribuir probabilidades a intervalos de valores da variável contínua por meio de uma função, que guarda muita analogia com a função que descreve a densidade de massa. Relembremos como é definida a densidade de massa.

Suponhamos que se tenha uma barra de metal em que a massa esteja distribuída de forma irregular. A densidade de massa no ponto x pode ser obtida da seguinte maneira. Considere um intervalo $(x, x+\Delta x)$ e determine a massa da seção da barra determinada por esse

intervalo. Dividindo-se a massa dessa seção pelo seu volume, tem-se a densidade da seção. Se a largura Δx dessa seção tende a zero, então as densidades dessas seções se aproximam da densidade da barra no ponto x. Lembremos ainda que se tivermos a densidade e quisermos determinar a massa total em uma seção da barra que vai do ponto a ao ponto b, integramos a densidade de a até b.



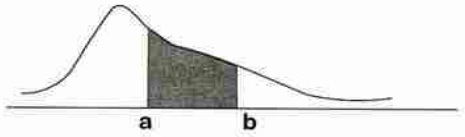
A densidade de probabilidade é definida de modo análogo. É uma função não negativa tal que sua integral num dado intervalo é igual à probabilidade da variável pertencer ao intervalo. Impõe-se ainda a condição de que a integral estendida à reta toda seja igual a um, pois ao ser realizado o experimento algum evento ocorre. Reunimos estas afirmações na seguinte definição.

DEFINIÇÃO 2.3.1 *A densidade de probabilidade de uma variável aleatória contínua é uma função $f(x) \geq 0$, tal que:*

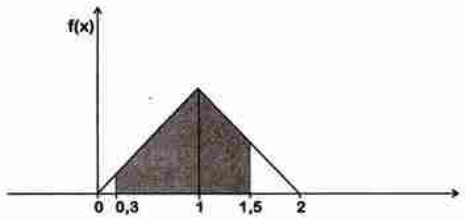
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

A probabilidade de uma variável aleatória X pertencer a um intervalo (a, b], é dada por:

$$P[a < X \leq b] = \int_a^b f(x)dx \qquad (2.1)$$



EXEMPLO 2.3.1 Seja $f(x) = x$ para $0 \leq x \leq 1$, $f(x) = 2 - x$ para $1 \leq x \leq 2$ e zero no complementar.



Observe que a integral de $f(x)$ no intervalo [0, 2] onde a densidade positiva é igual a 1, que é a área do triângulo na figura acima. Vamos calcular usando (2.1) as probabilidades: $P[0 \leq X \leq 0,8]$ e $P[0,3 \leq X \leq 1,5]$.

$$P[0 \leq X \leq 0,8] = \int_0^{0,8} x dx = \frac{x^2}{2}|_0^{0,8} = \frac{(0,8)^2}{2} = 0,32$$

$$P[0,3 \leq X \leq 1,5] = \int_{0,3}^{1} x dx + \int_1^{1,5} (2-x) dx =$$

$$= \frac{x^2}{2}|_{0,3}^{1} + 2x|_1^{1,5} - \frac{x^2}{2}|_1^{1,5} = 0,83$$

2.4 FUNÇÃO DE DISTRIBUIÇÃO DE UMA VARIÁVEL ALEATÓRIA

Voltamos a considerar nesta seção variáveis aleatórias gerais. O conceito de função de distribuição, também conhecido por distribuição cumulativa de probabilidade, que introduziremos aplica-se tanto a variáveis aleatórias discretas quanto a variáveis aleatórias contínuas. A função de distribuição nos dá uma outra maneira de descrever como as probabilidades são associadas aos valores ou aos intervalos de valores de uma variável aleatória.

• DEFINIÇÃO 2.4.1 *A função de distribuição de uma variável aleatória X é uma função que a cada número real x associa o valor:*

$$F(x) = P[X \leq x]\,. \tag{2.2}$$

A notação $[X \leq x]$ é usada para designar o conjunto $\{\omega \in S : X(\omega) \leq x\}$, isto é, denota a imagem inversa do intervalo $(-\infty, x]$ pela variável aleatória X.

• LEMA 2.4.1 *A função de distribuição de uma variável aleatória X satisfaz as seguintes condições:*

a) $0 \leq F(x) \leq 1$.

b)F(x) é não decrescente e contínua à direita

c) $\lim_{x \to -\infty} F(x) = 0$ *e* $\lim_{x \to \infty} F(x) = 1$.

DEMONSTRAÇÃO Iremos dar um esboço da demonstração, já que não dispomos dos instrumentos para dar uma demonstração rigorosa.

a) Como $F(x)$ representa uma probabilidade segue-se que $0 \leq F(x) \leq 1$.

b) Se $x_1 \leq x_2$ então $\{\omega : X(\omega) \leq x_1\} \subset \{\omega : X(\omega) \leq x_2\}$ e portanto pelo lema 1.6.4 segue que $P\{\omega : X(\omega) \leq x_1\} \leq P\{\omega : X(\omega) \leq x_2\}$, o que prova que $F(x_1) \leq F(x_2)$.

A demonstração da continuidade à direita de $F(x)$ depende de uma propriedade das probabilidades que enunciamos a seguir, mas cuja demonstração não apresentaremos.

Se uma seqüência de eventos A_n é decrescente e se aproxima de um evento A, então a seqüência das probabilidades $P(A_n)$ também é decrescente e tem como limite $P(A)$.

Considere uma seqüência de números reais x_n, que seja decrescente e tenha como limite x. A seqüência de eventos $[X \leq x_n]$ é decrescente e se aproxima de $[X \leq x]$. Usando a propriedade que acabamos de mencionar, fica demonstrada a continuidade à direita.

c) Para uma seqüência de eventos A_n que cresce para um evento A vale uma propriedade análoga, isto é, $P(A_n)$ converge para $P(A)$. Se considerarmos uma seqüência de números reais x_n que tende a infinito, segue-se que a seqüência de eventos $[X \leq x_n]$ tende ao evento $[X < \infty]$ e portanto $P[X \leq x_n]$ converge para $P[X < \infty]$. Mas esse evento é o espaço amostral todo, que tem probabilidade um. Assim fica demonstrada a primeira afirmação de c. Para demonstrar a segunda, basta considerar uma seqüência x_n que tende a -∞. A seqüência de eventos $[X \leq x_n]$ tende ao ϕ e portanto

$F(x_n) = P[X \leq x_n]$ tende a zero.

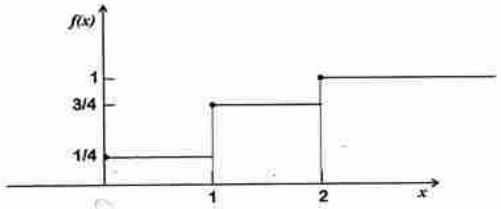
EXEMPLO 2.4.1 Vamos determinar a função de distribuição da variável aleatória X do exemplo 2.1.1, que é igual ao número de caras em dois lançamentos de uma moeda.

Temos $F(x) = 0$ para todo x tal que $-\infty < x < 0$, pois $[X \leq x] = \phi$ e pelo lema 1.6.1, $P[X \leq x] = 0$.

Para x tal que $0 \leq x < 1$, $[X \leq x] = \{\overline{C}\,\overline{C}\}$ e portanto $F(x) = P(\overline{C}\,\overline{C}) = \frac{1}{4}$.

Para x tal que $1 \leq x < 2$, $[X \leq x] = \{\overline{C}\,\overline{C}, \overline{C}\,C, C\,\overline{C}\}$ e portanto $F(x) = P(\overline{C}\,\overline{C}) + P(\overline{C}\,C) + P(C\,\overline{C}) = \frac{3}{4}$.

Para x tal que, $x \geq 2$, $[X \leq x] = \{\overline{C}\,\overline{C}, \overline{C}\,C, C\,\overline{C}\,, C\,C\}$ e portanto $F(x) = 1$.



Conforme demonstrado e ilustrado na figura, o valor de $F(x)$ é zero para todo $x < 0$. O limite à esquerda no ponto zero é portanto zero, mas o valor de $F(x)$ no ponto 0 é $\frac{1}{4}$, logo o ponto 0 é um ponto de descontinuidade de $F(x)$. Observe ainda que para $0 \leq x < 1$, $F(x) = \frac{1}{4}$ e que portanto o limite à direita no ponto 0, isto é, por valores de x maiores que zero, é igual a $\frac{1}{4}$, que coincide com o valor de $F(0)$, ou seja $F(x)$ é contínua à direita no ponto zero, como estabelece o lema 2.4.1. Além disso, vemos que $P(X = 0) = F(0) - \lim_{x \uparrow 0} F(x)$. Para os outros pontos de descontinuidade, vale uma expressão análoga como se pode verificar.

Vamos agora mostrar qual é a relação entre a função de distribuição e a distribuição de probabilidade de uma variável aleatória discreta.

Seja X uma variável aleatória discreta cuja distribuição de probabilidade associa aos valores $x_1, x_2, \ldots, x_n$ as respectivas probabilidades $P[X = x_1]$, $P[X = x_2]$, $\ldots$, $P[X = x_n]$. Para calcular o valor de $F(x)$ usando a definição 2.4.1, precisamos calcular a probabilidade do evento:

$$[X \leq x] = \bigcup_{i:x_i \leq x} [X = x_i].$$

Como os valores de X são todos distintos, os eventos do lado direito são mutuamente exclusivos, e calculando-se a probabilidade de ambos os membros dessa igualdade tem-se:

$$F(x) = \sum_{i:x_i \leq x} P[X = x_i]. \tag{2.3}$$

A expressão (2.3) nos mostra que se for dada a distribuição de probabilidade de uma variável aleatória discreta, fica determinada sua função de distribuição.

Para todo ponto de descontinuidade da função de distribuição, que são os valores x_i, $i = 1, 2, \ldots$ onde a variável discreta assume um valor com probabilidade positiva, temos:

$$P[X = x_i] = F(x_i) - \lim_{x \uparrow x_i} F(x). \tag{2.4}$$

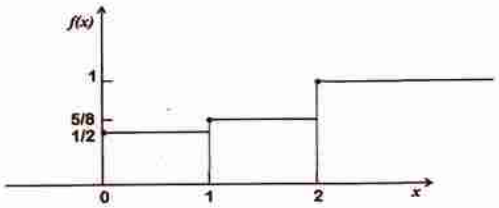
Observe que a expressão (2.4) nos dá a distribuição de probabilidade de X em termos da função de distribuição. As expressões (2.3) e (2.4) nos dão maneiras equivalentes de atribuir probabilidades aos valores da variável aleatória X, pois dada uma delas obtemos a outra.

EXEMPLO 2.4.2 A função de distribuição de uma variável aleatória X é dada por: $F(x) = 0$ para $x < 0$, $F(x) = 1/2$ para $0 \leq x < 1$, $F(x) = 5/8$ para $1 \leq x < 2$ e $F(x) = 1$ para $x \geq 2$. Determinar a distribuição de probabilidade de X.

Observemos que os saltos da função de distribuição de X ocorrem nos pontos 0, 1 e 2. Desta maneira temos:

$$P[X = 0] = F(0) - \lim_{x \uparrow 0} F(x) = 0,5$$

$$P[X = 1] = F(1) - \lim_{x \uparrow 1} F(x) = 0,625 - 0,5 = 0,125$$



$$P[X = 2] = F(2) - \lim_{x \uparrow 2} F(x) = 1 - 0,625 = 0,375.$$

Para uma variável aleatória contínua com densidade de probabilidade $f(x)$ podemos obter a função de distribuição $F(x)$ integrando-se a densidade de probabilidade,

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(y)dy. \tag{2.5}$$

Se a densidade $f(x)$ for contínua no seu campo de definição, então decorre do teorema fundamental do cálculo que:

$$F'(x) = f(x). \tag{2.6}$$

Sejam a e b dois números reais, com $a < b$. Vamos obter a probabilidade da variável aleatória pertencer ao intervalo $(a, b]$ em termos da função de distribuição F.

Partindo-se da igualdade $(-\infty, b] = (-\infty, a] \cup (a, b]$ e tomando-se a imagem inversa dos dois lados da igualdade pela variável aleatória X nós obtemos:

$$\{\omega : X(\omega) \leq b\} = \{\omega : X(\omega) \leq a\} \cup \{\omega : a < X(\omega) \leq b\}.$$

Como os dois eventos do lado direito são mutuamente exclusivos, calculando-se a probabilidade tem-se:

$$P\{\omega : X(\omega) \leq b\} = P\{\omega : X(\omega) \leq a\} + P\{\omega : a < X(\omega) \leq b\}.$$

Daí decorre que: $F(b) = F(a) + P[a < X \leq b]$ e portanto

$$P[a < X \leq b] = F(b) - F(a). \tag{2.7}$$

EXEMPLO 2.4.3 Consideremos a variável aleatória X com densidade de probabilidade dada no exemplo 2.3.1. Vamos determinar sua função de distribuição e calcular $P[0,7 \leq X \leq 1,8]$.

$F(x) = 0$ para $x \leq 0$, pois a densidade de probabilidade de X é nula para $x \leq 0$.

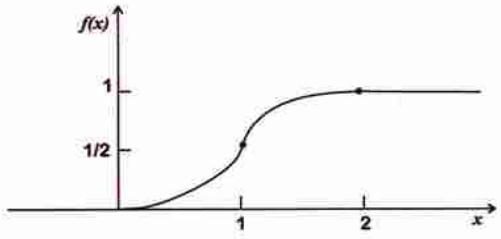
Para $0 \leq x < 1$, $F(x) = \int_{-\infty}^{x} ydy = \int_{0}^{x} ydy = \frac{x^2}{2}$

Para $1 \leq x < 2$,

$F(x) = \int_0^1 ydy + \int_1^x (2 - y)dy = \frac{1}{2} + (2x - 2) - (\frac{x^2}{2} - \frac{1}{2}) =$
$= -\frac{x^2}{2} + 2x - 1$

Para $x \geq 2$, $F(x) = 1$

Pela fórmula (2.7) temos $P[0,7 \leq X \leq 1,8] = F(1,8) - F(0,7) = 0,735$.



Observe que se X é uma variável contínua então a probabilidade de X assumir um dado valor é zero, e portanto quando consideramos as probabilidades de X tomar valores em intervalos não faz diferença se os extremos estão incluídos ou não.

EXEMPLO 2.4.4 Considere uma variável aleatória cuja densidade de probabilidade é dada por: $f(x) = 0$ para $x \leq 0$ e por $f(x) = 2e^{-2x}$ para $x > 0$. Determinar a função de distribuição de X e calcular $P[0,5 < X < 2]$.

Para $x \leq 0$, $F(x) = 0$.

Para $x > 0$,

$$F(x) = \int_0^x 2e^{-2t}dt = -e^{-2t}|_0^x = 1 - e^{-2x}.$$

Para calcular $P[0,5 < X < 2]$ podemos nos utilizar da expressão (2.7), obtendo:

$$P[0,5 < X < 2] = (1 - e^{-4}) - (1 - e^{-1}) = e^{-1} - e^{-4}.$$

2.5 ESPERANÇA MATEMÁTICA DE VARIÁVEIS ALEATÓRIAS DISCRETAS

Nosso referencial é um espaço amostral S, cujos pontos amostrais designaremos por ω e uma probabilidade P definida em S.

DEFINIÇÃO 2.5.1 *A esperança matemática de uma variável aleatória discreta X, definida em um espaço amostral S no qual está definida uma probabilidade P é dada por:*

$$E(X) = \sum_{\omega \in S} X(\omega)P(\omega) \tag{2.8}$$

desde que $\sum_{\omega \in S} |X(\omega)|P(\omega) < \infty$.

O próximo lema mostra que a esperança pode ser expressa por meio da distribuição de probabilidade de X.

LEMA 2.5.1 *A esperança matemática de uma variável aleatória discreta X que assume os valores x_i, com respectivas probabilidades $P[X = x_i]$, para $i = 1, 2, \ldots$ é dada por:*

$$E(X) = \sum_{i=1}^{\infty} x_i P[X = x_i]. \tag{2.9}$$

DEMONSTRAÇÃO Vamos partir da fórmula (2.8) e obter a fórmula (2.9) e mostrar que o caminho inverso ao percorrido é válido também. Vamos denotar os pontos do espaço amostral por ω_j, para $j = 1, 2, \ldots$ Como a série $\sum_{j=1}^{\infty} X(\omega_j)P(\omega_j)$ é absolutamente convergente, a ordem de suas parcelas pode ser modificada obtendo-se a mesma soma.

$$\begin{aligned} E(X) &= \sum_{j=1}^{\infty} X(\omega_j)P(\omega_j) = \sum_{i=1}^{\infty}[\sum_{j:X(\omega_j)=x_i} X(\omega_j)P(\omega_j)] = \\ &= \sum_{i=1}^{\infty}[\sum_{j:X(\omega_j)=x_i} x_i P(\omega_j)] = \sum_{i=1}^{\infty}[x_i \sum_{j:X(\omega_j)=x_i} P(\omega_j)] = \\ &= \sum_{i=1}^{\infty} x_i P[X = x_i]. \end{aligned}$$

A primeira igualdade acima é a definição de esperança. A segunda é obtida agrupando-se na soma as parcelas em que X assume cada um de seus valores x_1, x_2, ... A terceira igualdade é obtida substituindo-se $X(\omega_j)$ por x_i no conjunto onde $X(\omega_j)$ é igual a x_i. A quarta igualdade é obtida pondo em evidência x_i que é constante em relação à soma. Finalmente a última igualdade resulta de $P[X = x_i] = \sum_{j:X(\omega_j)=x_i} P(\omega_j)$.

Observe que se pode partir do último termo da direita nas igualdades acima e percorrer as igualdades no sentido contrário, justificando cada igualdade como fizemos. Isto prova o lema.

Notemos que se X e Y sao variáveis aleatórias definidas em um espaço amostral Ω e se c é um número real, então cX e $X + Y$ definidas em Ω da seguinte maneira também são variáveis aleatórias.

Para cada $\omega \in \Omega$,

$(X + Y)(\omega) = X(\omega) + Y(\omega)$ e $cX(\omega) = cX(\omega)$.

A esperança matemática, que muitas vezes referiremos simplesmente por esperança, tem a propriedade da linearidade que é verificada no próximo lema.

LEMA 2.5.2 *Se as esperanças das variáveis aleatórias X e Y existem, então existe a esperança de $X + Y$ e se c é uma constante tem-se:*

$$E(X + Y) = E(X) + E(Y) \tag{2.10}$$

$$E(cX) = cE(X). \tag{2.11}$$

DEMONSTRAÇÃO Mostremos que $E(X + Y)$ existe.

$$\sum_{\omega \in S} |X(\omega) + Y(\omega)|P(\omega) \leq \sum_{\omega \in S}[|X(\omega)| + |Y(\omega)|]P(\omega) =$$

$$= \sum_{\omega \in S} |X(\omega)|P(\omega) + \sum_{\omega \in S} |Y(\omega)|P(\omega) < \infty;$$

$$E(X + Y) = \sum_{\omega \in S} (X + Y)(\omega)P(\omega) = \sum_{\omega \in S} (X(\omega) + Y(\omega))P(\omega) =$$

$$= \sum_{\omega \in S} X(\omega)P(\omega) + \sum_{\omega \in S} Y(\omega)P(\omega) = E(X) + E(Y).$$

A verificação da fórmula (2.11) é imediata e fica a cargo do leitor.

Este resultado pode ser facilmente estendido para a soma de um número inteiro n de variáveis aleatórias, isto é, se existem as esperanças de X_1, X_2, ... , X_n então existe a esperança de $X_1 + X_2 + \cdots + X_n$ e tem-se:

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n). \quad (2.12)$$

A demonstração fica a cargo do leitor como exercício.

EXEMPLO 2.5.1 Uma loteria vende 100 bilhetes. O preço de cada bilhete é R$ 1,20 e o bilhete sorteado paga um prêmio de R$ 100,00. Você compra um bilhete. Qual é a esperança de seu ganho?

Designemos por X a variável aleatória que é igual ao seu ganho. X assume o valor - R$ 1,20 com probabilidade 0,99 e R$ 98,80 com probabilidade 0,01.Temos para a esperança de X:

$$E(X) = -1,20 \times 0,99 + 98,80 \times 0,01 = -0,20.$$

Vamos dar duas interpretações para o conceito de esperança de uma variável aleatória, uma interpretação física e uma interpretação baseada na definição freqüentista de probabilidade.

2.5.1 *Interpretação Física*

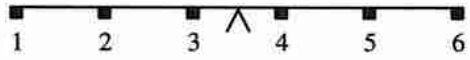
Consideremos uma variável aleatória X que assume os valores x_1, x_2, ... , x_n com probabilidades p_1, p_2, ... , p_n respectivamente. Vamos representar sua distribuição de probabilidade por meio de uma distribuição de massas em um eixo horizontal. Em cada um dos valores x_i penduramos um peso com massa M_i, para $i = 1, 2, \ldots, n$, e juntamos esses pontos por uma haste rígida de massa desprezível. Designemos a massa total por M, isto é, $M = M_1 + M_2 + \cdots + M_n$. As massas M_i para $i = 1, 2, \ldots, n$ são escolhidas de modo que $p_i = \frac{M_i}{M}$, para $i = 1, 2, \ldots, n$.

Vamos mostrar que o valor da esperança é igual à abscissa do centro de gravidade dessa distribuição de massas. Se x_G denota a abscissa do centro de gravidade, tem-se:

$$x_G.M = \sum_{i=1}^{n} x_i.M_i$$

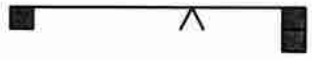
$$x_G = \sum_{i=1}^{n} x_i \frac{M_i}{M} = \sum_{i=1}^{n} x_i.p_i = E(X).$$

Consideremos, por exemplo, a distribuição de probabilidade da variável aleatória que se obtém observando-se o número de pontos que aparece na face superior quando se lança um dado balanceado. Denotemos por X essa variável aleatória. Sua distribuição de probabilidade associa a cada um dos inteiros de 1 a 6 probabilidade $\frac{1}{6}$. Usando-se a representação descrita tem-se:



O centro de gravidade dessa distribuição de massas é o ponto 3,5, que é o valor da esperança de X, e o ponto em torno do qual a distribuição de massas se equilibra.

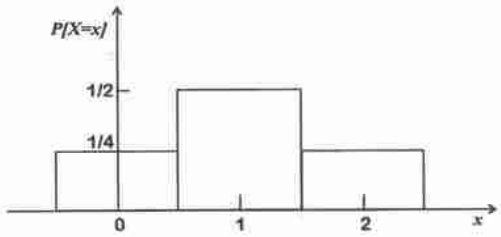
Consideremos agora uma urna com duas bolas vermelhas e quatro brancas. Retira-se uma bola da urna. Seja Y a variável aleatória que assume o valor zero se a bola é vermelha e um se a bola é branca. A distribuição de probabilidade de Y é dada por: $P[Y = 0] = \frac{1}{3}$ e $P[Y = 1] = \frac{2}{3}$. Representando-se esta distribuição como descrito acima, tem-se:



O ponto em torno do qual a distribuição se equilibra quando se coloca um fulcro é o centro de gravidade, ou seja, o ponto $\frac{2}{3}$, que é o valor da esperança.

Em vez de considerarmos uma distribuição de massas em uma barra horizontal, vamos considerar a massa unitária espalhada de

maneira uniforme no histograma, da distribuição de probabilidade da variável aleatória Y. Vamos considerar a distribuição da variável aleatória do exemplo 2.2.1 e construir o histograma dessa distribuição. Se colocarmos um fulcro na abscissa que corresponde à esperança, como este ponto coincide com o centro de gravidade da quantidade de massa espalhada no histograma, vemos que este é o ponto onde o histograma se equilibra.



2.5.2 *Interpretação Baseada na Formulação Freqüentista de Probabilidade*

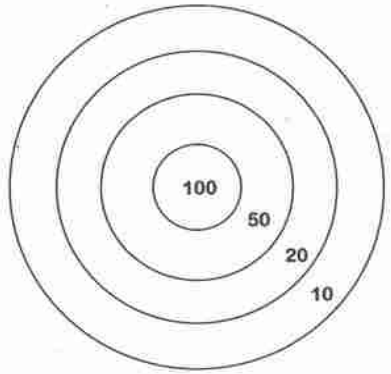
Consideremos um alvo de forma circular dividido em quatro regiões adjacentes como na figura. A região 1 é o círculo interno, a região 2 é a coroa circular adjacente, a região 3 é a próxima coroa circular e a quarta região é a última coroa circular. O número de pontos que o atirador faz ao acertar um tiro nessas regiões é o seguinte: 100 se acertar na região 1; 50 pontos se acertar na região 2; 20 pontos se acertar na região 3; 10 se acertar na região 4 e zero se não acertar em nenhuma das 4 regiões.

Vamos supor que o atirador faz uma seqüência de n disparos nesse alvo. Se ele acerta n_1 vezes na região 1, n_2 vezes na região 2, n_3 vezes na região 3, n_4 na região 4 e não acerta em nenhuma das regiões n_0 vezes, então o número total de pontos que ele faz é igual a:

$$0n_0 + 100n_1 + 50n_2 + 20n_3 + 10n_4. \tag{2.13}$$

O número médio de pontos é igual a:

$$0\frac{n_0}{n} + 100\frac{n_1}{n} + 50\frac{n_2}{n} + 20\frac{n_3}{n} + 10\frac{n_4}{n}. \tag{2.14}$$



Se o número de tiros que é dado ao alvo é grande, então as frações $\frac{n_i}{n}$, para i=0, 1, 2, 3, e 4, são aproximadamente iguais às probabilidades do atirador não acertar no alvo ou acertar nas regiões 1, 2, 3 ou 4 respectivamente, e portanto a esperança é aproximadamente igual ao número médio de pontos feitos pelo atirador.

Vemos que podemos interpretar a esperança matemática de uma variável aleatória discreta como a média dos valores observados dessa variável em uma seqüência grande de repetições do experimento em que, em cada realização do mesmo, observa-se um valor da variável.

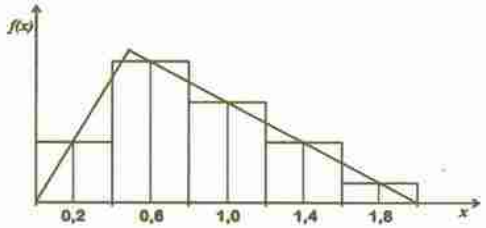
2.6 ESPERANÇA MATEMÁTICA DE VARIÁVEIS ALEATÓRIAS CONTÍNUAS

Para introduzir o conceito de esperança para variáveis aleatórias contínuas, vamos considerar uma variável aleatória X com densidade de probabilidade

$$f(x) = \begin{cases} 2x, & \text{para } 0 \leq x \leq 0,5, \\ -\frac{2}{3}x + \frac{4}{3}, & \text{para } 0,5 \leq x \leq 2 \\ 0, & \text{no complementar.} \end{cases}$$

Vamos fazer o gráfico dessa densidade e espalhar na figura delimitada pelo mesmo e pelo eixo das abscissas uma massa unitária,

como fizemos no caso de um histograma de uma variável aleatória discreta.



Vamos construir variáveis aleatórias discretas que aproximam esta variável contínua X. O sentido da aproximação ficará claro à medida que avançarmos. Vamos dividir o conjunto dos valores de X, ou seja, o intervalo [0,2], em cinco partes iguais, de comprimento 0,4. Denominemos X_5 a variável cujos valores são os pontos médios desses intervalos, isto é, que assume os valores: 0,2; 0,6; 1,0; 1,4; 1,8. As probabilidades associadas a esses valores serão iguais às áreas dos retângulos cujas bases são iguais ao comprimento dos intervalos e cujas alturas traçadas nos pontos médios dos intervalos são iguais ao valor das densidades nesses pontos. No ponto 0,2 o valor da altura é $f(0,2) = 2(0,2) = 0,4$. Nos demais pontos a expressão da densidade é $f(x) = -\frac{2}{3}x + \frac{4}{3}$. Assim substituindo os valores 0,6; 1,0; 1,4; 1,8 nessa espressão, obtemos para essas alturas 0,93; 0,66; 0,4; 0,13. Como as probabilidades desses valores são iguais às areas dos retângulos, temos, para a distribuição de probabilidade de X_5,

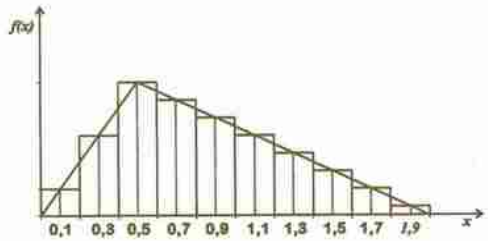
Valores de X_5	Probabilidades
0,2	0,160
0,6	0,372
1.0	0,264
1,4	0,160
1,8	0,052

Calculemos a esperança da variável X_5. Pela fórmula (2.9) temos:

$$\begin{aligned} E(X_5) &= (0,2)(0,16) + (0,6)(0,372) + \\ &\quad +(1,0)(0,264) + (1,4)(0,16) + (1,8)(0,052) \\ &= 0,836 \end{aligned}$$

Agora, se dividirmos o intervalo [0, 2] em 10 partes iguais, em vez de 5, e procedermos da mesma maneira, vamos obter uma variável discreta X_{10} que terá 10 valores, 0,1; 0,3; 0,5; 0,7; 0,9; 1,1; 1,3; 1,5; 1,7 e 1,9. As probabilidades associadas a esses valores serão iguais ao produto de 0,2 que é a base dos retângulos pelas alturas que são iguais ao valor da densidade nesses pontos. Apresentamos na tabela a seguir na primeira coluna os valores de X, na segunda os valores da densidade de X nos pontos médios dos intervalos e na terceira as probabilidades desses valores.

x	$f(x)$	$P[X = x]$	x	$f(x)$	$P[X = x]$
0,1	0,200	0,040	1,1	0,600	0,120
0,3	0,600	0,120	1,3	0,466	0,093
0,5	1,000	0,200	1,5	0,333	0,066
0,7	0,867	0,173	1,7	0,200	0,040
0,9	0,733	0,146	1,9	0,066	0,013



Utilizando-se novamente a fórmula (2.9) obtém-se:

$$\begin{aligned} E(X_{10}) &= (0,1)(0,04) + (0,3)(0,12) + (0,5)(0,2) + \\ &\quad +(0,7)(0,173) + (0,9)(0,146) + (1,1)(0,12) + \\ &\quad +(1,3)(0,093) + (1,5)(0,066) + (1,7)(0,04) + \\ &\quad +(1,9)(0,013) = 0,8357. \end{aligned}$$

Podemos construir uma seqüência de aproximações, tomando um número de divisões cada vez maior, de modo que os comprimentos dos intervalos tendam a zero. Analisando esse exemplo pode-se perceber que, quando o número de divisões é grande, o histograma da variável discreta é bastante próximo da densidade de probabilidade da variável original X e as esperanças dessas variáveis discretas, que são somas do tipo $\sum_i x_i P[X = x_i]$, aproximam-se da integral definida de $xf(x)$, onde $f(x)$ é a densidade de probabilidade de X. Observe que esse é o esquema de aproximações que se encontra no cálculo para definir a integral definida de uma função contínua.

Desta maneira, definimos a esperança da variável contínua X da seguinte maneira:

DEFINIÇÃO 2.6.1 *A esperança de uma variável aleatória X, com densidade de probabilidade $f(x)$ é dada por*

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx. \tag{2.15}$$

A esperança de uma variável aleatória é usualmente referida como uma medida de posição da distribuição dessa variável. O sentido em que é empregado o termo medida de posição não é muito preciso. Vimos um dos sentidos possíveis, ou seja, que o valor da esperança é igual à abscissa do centro de gravidade da distribuição de massas colocadas em cada um dos valores da variável e numericamente iguais às probabilidades desses valores. Observemos que a esperança está sempre compreendida entre os valores extremos da variável.

EXEMPLO 2.6.1 Calcular a esperança da variável aleatória X cuja densidade de probabilidade é dada por:

$$f(x) = \begin{cases} 2x, & \text{para } 0 \leq x \leq 0,5 \\ -\frac{2}{3}x + \frac{4}{3}, & \text{para } 0,5 \leq x \leq 2 \\ 0, & \text{no complementar.} \end{cases}$$

Pela fórmula (2.15) tem-se:

$$\begin{aligned} EX &= \int_0^{0,5} x.2xdx + \int_{0,5}^{2} x\left(-\frac{2}{3}x + \frac{4}{3}\right)dx = \\ &= 2\frac{x^3}{3}\Big|_0^{0,5} + \left(-\frac{2}{3}\right)\frac{x^3}{3}\Big|_{0,5}^{2} + \frac{4}{3}\frac{x^2}{2}\Big|_{0,5}^{2} = \\ &= \frac{1}{12} - \frac{63}{36} + \frac{2}{3}\left(4 - \frac{1}{4}\right) = \frac{30}{36} = 0,8333. \end{aligned}$$

2.7 FUNÇÕES DE VARIÁVEIS ALEATÓRIAS

Do mesmo modo que a soma de variáveis aleatórias é uma nova variável aleatória, as operações elementares de produto e quociente também produzem novas variáveis aleatórias. Elas são definidas da seguinte maneira:

$$(X.Y)(\omega) = X(\omega).Y(\omega)$$

$$(\frac{X}{Y})(\omega) = \frac{X(\omega)}{Y(\omega)}.$$

Para o quociente, exigimos que o denominador seja diferente de zero, salvo casos em que convencionamos que nessa situação o quociente é infinito.

As operações de determinar-se para cada par de valores de X e de Y qual o menor valor e qual o maior valor deles, que denotaremos por $min\{X, Y\}$, $max\{X, Y\}$, também são variáveis aleatórias definidas em Ω. Formalmente essas variáveis são definidas para cada $\omega \in \Omega$ da seguinte maneira:

$$min(X, Y)(\omega) = min\{X(\omega), Y(\omega)\}$$

$$max(X, Y)(\omega) = max\{X(\omega), Y(\omega)\}.$$

Notemos que as definições de soma, produto, min e max se estendem para um número finito de variáveis aleatórias.

EXEMPLO 2.7.1 Em uma vila de cinco residências considere o consumo mensal de energia elétrica. Não é difícil de convencermo-nos que esses consumos mensais, que denotaremos por X_1, X_2, X_3, X_4, X_5, são variáveis aleatórias. Então o consumo total das cinco residências, o menor consumo entre elas e o maior consumo entre elas também são variáveis aleatórias. Se num dado mês os consumos das cinco casas foram: 120, 176, 149, 263 e 190 quilowatts-hora, então para esse mês temos: $X_1 + X_2 + X_3 + X_4 + X_5 = 898$, $min\{X_1, X_2, X_3, X_4, X_5\} = 120$, $max\{X_1, X_2, X_3, X_4, X_5\} = 263$.

Mais adiante, veremos como determinar a distribuição de probabilidade de funções de variáveis aleatórias. Isto é importante, pois é freqüente que estejamos interessados em uma função de uma variável aleatória quando se realiza um experimento aleatório. Considere, por exemplo, um seguro de acidentes de automóvel. O segurado paga um prêmio para obter uma cobertura para acidentes com seu carro e em geral para danos causados a terceiros envolvidos no acidente. O valor pago pela Cia. de Seguro vai depender da extensão do dano incorrido no acidente. Esta é uma das situações mais usuais em que temos uma função de uma variável aleatória, isto é, uma situação em que associamos ao valor observado da variável uma recompensa ou um custo. Doravante consideraremos nesta seção funções de variáveis aleatórias discretas.

EXEMPLO 2.7.2 Considere o jogo da roleta. Há 37 números inteiros do 0 ao 36, sendo 18 dos quais de cor vermelha e 18 de cor preta, sendo que o 0 é branco. Há várias opções de apostas, entre as quais destacamos duas: pode-se apostar em preto ou vermelho e, em caso de acerto, para cada real apostado recebe-se um outro real; pode-se apostar na primeira dúzia, isto é, no conjunto dos números de 1 a 12; caso ocorra um desses números, recebe-se R$ 2,00 para cada real apostado. Suponha que um jogador aposta R$ 10,00 na primeira dúzia. Designemos por X a variável aleatória que é igual ao número observado quando se gira a roleta. Designemos por Y o seu ganho nessa rodada da roleta. Vamos determinar a distribuição de probabilidade e a esperança de Y.

Pela definição de Y tem-se:

Se $1 \leq X \leq 12$ então $Y = 20$

Se $13 \leq X \leq 36$ ou se $X = 0$ então $Y = -10$.

Como todos os pontos do espaço amostral são equiprováveis, obtemos para a distribuição de probabilidade de Y:

valor de Y	Probabilidade
20	12/37
−10	25/37

Podemos calcular a esperança de Y de duas maneiras. Usando a definição de esperança obtemos a partir da tabela:

$$E(Y) = 20.(12/37) - 10.(25/37) = -10/37.$$

Agora consideremos a variável aleatória $Y(X)$ definida no espaço amostral: $\Omega = \{0, 1, 2, \ldots, 36\}$. Cada um desses pontos tem probabilidade 1/37. $Y(X(\omega)) = 20$, para ω= 1, 2, ... , 12 e $Y(X(\omega))$ = 0 para ω =0, 13, 14, ... , 36. Calculando-se a esperança dessa variável aleatória $Y(X)$ definida nesse espaço amostral, temos:

$$E[Y(X)] = \sum_{\omega=0}^{36} Y(X(\omega))P(\omega) = 20(12/37) - 10(25/37) = -10/37$$

EXEMPLO 2.7.3 Seja X o número observado quando se lança um dado e seja $g(x) = (x - E(X))^2$. O valor de E(X) é 3,5. Seja $Y = g(X) = (X - E(X))^2$. Na tabela abaixo listamos os valores de Y e os valores de X em cada um dos pontos do espaço amostral e calculamos $E(Y)$. Note que desta tabela podemos determinar também a distribuição de probabilidade de Y.

Espaço amostral	X	Y	Probabilidade
1	1	6,25	1/6
2	2	2,25	1/6
3	3	0,25	1/6
4	4	0,25	1/6
5	5	2,25	1/6
6	6	6,25	1/6

Calculando-se a esperança de Y definida no espaço amostral $\Omega = \{1,2,3,4,5,6\}$, tem-se:

$$\begin{aligned} E(Y) &= 6,25.1/6 + 2,25.1/6 + 0,25.1/6 + \\ &\quad +0,25.1/6 + 2,25.1/6 + 6,25.1/6 \\ &= 2,91. \end{aligned} \tag{2.16}$$

Da tabela obtemos a distribuição de probabilidade de Y:

$Y = (X - E(X))^2$	Probabilidade
0,25	1/3
2,25	1/3
6,25	1/3

Calculando-se a esperança de $Y = (X - EX)^2$, tem-se:

$$E(X - E(X))^2 = 0,25.1/3 + 2,25.1/3 + 6,25.1/3 = 2,91. \tag{2.17}$$

Note a diferença conceitual que existe entre essas duas maneiras de calcular o valor da esperança nesses dois exemplos. Na expressão (2.16) para calcular a esperança de $(X - E(X))^2$, utilizamos a distribuição de probabilidade de X e não precisamos determinar a distribuição de probabilidade de $(X - E(X))^2$, enquanto para calcular essa esperança pela expressão (2.17), precisamos determinar a distribuição de probabilidade dessa função de X.

O seguinte lema mostra que esta situação vale no caso geral para toda função de uma variável aleatória.

LEMA 2.7.1 *Seja X uma variável aleatória discreta definida em um espaço amostral Ω, g uma função real e seja $Y = g(X)$.*

A distibuição de probabilidade de Y é dada por:

$$\text{a)} \qquad P[Y = y] = \sum_{x:g(x)=y} P[X = x] \tag{2.18}$$

$$\text{b)} \qquad E[g(X)] = \sum_{x} g(x)P[X = x]. \tag{2.19}$$

DEMONSTRAÇÃO Note que $[Y = y] = \bigcup_{x:g(x)=y}[X = x]$. Dado um valor y de Y considera-se a imagem inversa desse valor pela função g, ou seja, todos os valores x tais que $g(x) = y$ e toma-se a reunião desses eventos. Para valores distintos de x, os eventos $[X = x]$ são mutuamente exclusivos e portanto a probabilidade da reunião é igual à soma das probabilidades, o que é expresso pela expressão (2.18).

Para verificar b note que $g(X)$ é uma função composta de g e de X, que é uma variável aleatória definida no espaço amostral Ω. Para $\omega \in \Omega$ seu valor é $g(X(\omega))$. O resto da demonstração é análogo à demonstração do lema 2.5.1. De fato,

$$\begin{aligned} E[g(X)] &= \sum_{\omega} g(X(\omega))P(\omega) = \sum_{x} \sum_{\omega:X(\omega)=x} g(X(\omega))P(\omega) = \\ &= \sum_{x} g(x) \sum_{\omega:X(\omega)=x} P(\omega) = \sum_{x} g(x)P[X = x]\,. \end{aligned}$$

DEFINIÇÃO 2.7.1 *A variância de uma variável aleatória X, denotada $\sigma^2(X)$, é definida por:*

$$\sigma^2(X) = E(X - E(X))^2. \tag{2.20}$$

Note que se a variável aleatória X é discreta, assumindo valores $x_1, x_2, \ldots$ então a variância será dada por:

$$\sigma^2(X) = \sum_{i=1}^{\infty} (x_i - E(X))^2 P[X = x_i]. \tag{2.21}$$

Se a variavel aleatória é contínua, então a variância é dada por:

$$\sigma^2(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx. \tag{2.22}$$

Uma outra forma de se exprimir a variância de uma variável aleatória X é a seguinte:

$$\sigma^2(X) = E(X^2) - (E(X))^2. \tag{2.23}$$

De fato, por definição $\sigma^2(X) = E(X - E(X))^2$. Desenvolvendo-se o quadrado nessa expressão, tem-se:

$$\sigma^2(X) = E(X^2 - 2XE(X) + (E(X))^2) = E(X^2) - 2(E(X))^2 + (E(X))^2$$

simplificando, obtém-se (2.23).

A variância fornece uma medida de dispersão dos valores da variável aleatória em relação à sua esperança. Ela, a exemplo da esperança, também pode ser interpretada fisicamente. Representando-se a distribuição de probabilidades por meio de uma distribuição de massas, colocando-se em cada valor da variável uma massa igual a sua probabilidade, então o valor da variância coincide com o valor do momento de inércia dessa distribuição de massas em torno de seu centro de gravidade.

DEFINIÇÃO 2.7.2 *A raiz quadrada da variância de uma variável aleatória X, denotada $\sigma(X)$, denomina-se desvio padrão de X.*

EXEMPLO 2.7.4 Calcular a variância da variável aleatória X do exemplo 2.4.1

Vamos transcrever aqui a distribuição de probabilidade de X.

Valores de X	Probabilidades
0	1/4
1	1/2
2	1/4

$E(X) = 0(\frac{1}{4}) + 1(\frac{1}{2}) + 2(\frac{1}{4}) = 1$, pela fórmula (2.17) tem-se:

$$\sigma^2(X) = (-1)^2(\frac{1}{4} + 0(\frac{1}{2}) + 1(\frac{1}{4}) = \frac{1}{2}.$$

2.8 MOMENTOS E SUA FUNÇÃO GERADORA

Vimos que a esperança de uma variável aleatória nos dá idéia da posição da distribuição de probabilidade dessa variável aleatória e que a variância nos dá idéia da dispersão dos valores da variável em torno da esperança. Da expressão (2.23) vemos que a variância pode ser expressa como uma função da esperança das duas primeiras potências de X. Pode-se mostrar que outras características da distribuição de probabilidade de X podem ser expressas por meio das esperanças das potências de X. As esperanças das sucessivas potências de X constituem o conceito de momentos dessa variável

aleatória. Introduziremos também o conceito de função geradora de momentos e mostraremos que esta é uma ferramenta muito útil para determinarem-se distribuições de funções de variáveis aleatórias, principalmente de somas de variáveis aleatórias independentes.

DEFINIÇÃO 2.8.1 *A esperança de X^k é denominada momento de ordem k da variável aleatória X, para $k = 1, 2, 3, \ldots$*

+ Observe que a esperança de X é o momento de ordem 1 de X e que a variância é igual à diferença entre o segundo momento de X e o quadrado do primeiro momento.

Se em vez de considerarmos a esperança de X^k, considerarmos a esperança de $|X|^k$ o momento é denominado momento absoluto de ordem k.

Para uma variável aleatória discreta X com distribuição de probabilidade que associa aos valores x_1, x_2, ... as probabilidades $P[X = x_1]$, $P[X = x_2]$, ... o momento de ordem k, para $k = 1, 2, \ldots$ é dado por:

$$E(X^k) = \sum_{i=1}^{\infty} x_i^k P[X = x_i] \; . \tag{2.24}$$

Para uma variável contínua X com densidade de probabilidade $f(x)$, temos para o momento de ordem k, para $k = 1, 2, \ldots$

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx. \tag{2.25}$$

Vimos que a toda variável aleatória está associada sua distribuição de probabilidade. Vamos associar a X uma outra função definida da seguinte maneira:

DEFINIÇÃO 2.8.2 *A função geradora de momentos da variável aleatória X, denotada $\phi_X(t)$, é definida por:*

$$\phi_X(t) = E(e^{tX}) \tag{2.26}$$

para todo t, $-\infty < t < \infty$, em que a esperança seja finita.

No caso de uma variável contínua, a expressão (2.26) se exprime da seguinte maneira:

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx. \tag{2.27}$$

Para variáveis aleatórias discretas, a função geradora de momentos se exprime da seguinte maneira:

$$\phi_X(t) = \sum_{i=1}^{\infty} e^{tx_i} P[X = x_i]. \tag{2.28}$$

Em vários pontos impusemos a condição da esperança de uma variável aleatória ser finita. Isto ocorreu, por exemplo, na definição da função geradora de momentos. Vamos aproveitar a ocasião para dar exemplos de variáveis aleatórias contínuas e discretas que têm esperança infinita.

EXEMPLO 2.8.1 Seja X uma variável aleatória com distribuição de Cauchy, isto é, cuja densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \text{ para } -\infty < x < \infty.$$

Calculemos $E|X|$.

$$E|X| = \lim_{a\to\infty}\int_{-a}^{a}\frac{|x|}{\pi(1+x^2)}dx = \lim_{a\to\infty}\int_{0}^{a}\frac{x}{\pi(1+x^2)}dx =$$

$$= \lim_{a\to\infty}\frac{1}{\pi}(\log(a^2+1)) = \infty.$$

EXEMPLO 2.8.2 Seja X uma variável aleatória discreta que assume os valores $1, 2, \ldots, 2^n, \ldots$, com probabilidades:

$$P[X = 2^n] = \frac{1}{2^n}.$$

A esperança de X é dada por:

$$EX = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \infty.$$

EXEMPLO 2.8.3 Seja X uma variável aleatória que assume os valores zero e um, com probabilidades iguais a $1-p$ e p, respectivamente. Vamos determinar a função geradora de momentos de X.

$$\phi_X(t) = E(e^{tX}) = 1 - p + pe^t \tag{2.29}$$

para todo t tal que $-\infty < t < \infty$.

EXEMPLO 2.8.4 Seja X uma variável aleatória com densidade de probabilidade $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$, onde λ é uma constante positiva. Determinar a função geradora de momentos de X.

$$\phi(t) = E(e^{tX}) = \int_0^{\infty} e^{tX}\lambda e^{-\lambda x}dx = \int_0^{\infty}\lambda e^{-(\lambda-t)x}dx = \frac{\lambda}{\lambda - t}.$$

Observe que esta integral só existe, ou melhor, é finita para os valores de t tais que $0 < t < \lambda$.

Mostraremos a seguir a razão de $\phi_X(t)$ ser denominada função geradora de momentos. Consideremos, a título de exemplo, uma variável aleatória contínua X, cuja função geradora de momentos $\phi_X(t)$ é dada pela expressão (2.27). Derivando-se em relação a t e trocando o símbolo de derivada com o de integral temos:

$$\frac{d}{dt}\phi_X(t) = \int_{-\infty}^{\infty} xe^{tx}f(x)dx.$$

Calculando-se o valor dessa derivada no ponto t=0, temos:

$$\phi_X'(0) = \int_{-\infty}^{\infty} xf(x)dx = E(X). \tag{2.30}$$

Calculando-se a segunda derivada de $\phi_X(t)$ tem-se:

$$\frac{d^2}{dt^2}\phi_X(t) = \int_{-\infty}^{\infty} x^2e^{tx}f(x)dx. \tag{2.31}$$

O valor dessa derivada no ponto zero é:

$$\phi_X''(0) = \int_{-\infty}^{\infty} x^2f(x)dx = E(X^2). \tag{2.32}$$

Prosseguindo-se para $k = 3, 4, \ldots$ nós obtemos:

$$\phi_X^k(0) = E(X^k), \tag{2.33}$$

onde $\phi_X^k(0)$ designa a k-ésima derivada da função $\phi_X(t)$ no ponto $t = 0$.

Vimos que, tendo a função geradora de momentos de uma variável aleatória, podemos obter todos os momentos dessa variável aleatória. Vamos mostrar que, reciprocamente, se tivermos todos os momentos, podemos obter a função geradora de momentos.

Nós estamos tratando esta matéria heuristicamente, isto é, estamos fornecendo um esboço do desenvolvimento matemático e não estamos apresentando as condições e justificações para executar certas operações. Para fazê-lo, precisaríamos trabalhar com um nível de pré-requisitos que não são exigidos neste texto.

Vamos retomar nosso objetivo, que é mostrar que os momentos determinam a função geradora de momentos.

Consideremos a expansão da função exponencial em série de potências:

$$e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} .$$

Para a variável aleatória X, a função exponencial correspondente é dada por:

$$e^{tX} = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!} .$$

Supondo que a esperança de e^{tX} é finita temos:

$$E[e^{tX}] = E[\sum_{n=0}^{\infty} \frac{(tX)^n}{n!}] = \sum_{n=0}^{\infty} \frac{t^n E[X^n]}{n!} .$$

Nós admitimos na igualdade acima que as condições que permitem trocar o símbolo de esperança com o da série estão satisfeitas. Assim, obtemos:

$$\phi_X(t) = E[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^n E[X^n]}{n!} . \qquad (2.34)$$

A expressão (2.34) nos mostra que se conhecermos todos os momentos da variável aleatória X, podemos determinar a função geradora de momentos de X somando os termos da série nessa expressão.

A seguir enunciamos um resultado muito importante, como veremos mais adiante. Não apresentaremos, contudo, sua demonstração por exigir conhecimentos mais avançados do que os adotados para a leitura deste livro. A compreensão da aplicação desse resultado é essencial.

LEMA 2.8.1 *Existe uma correspondência biunívoca entre funções de distribuição e funções geradoras de momento de variáveis aleatórias.*

No lema que acabamos de enunciar chamamos a atenção do leitor que a função de distribuição está sempre definida, ao passo que a função geradora de momentos de uma variável aleatória pode não estar definida, por não ser finita a soma ou a integral conforme a variável seja discreta ou contínua.

Considerando os casos em que as funções geradoras de momento são finitas, o que dissemos anteriormente implica que podemos colocar em uma coluna as funções de distribuição de variáveis aleatórias e em outra coluna correspondente as funções geradoras de momento correspondentes. Se for dada uma função de distribuição, podemos identificar qual a função geradora de momentos que lhe corresponde e vice-versa.

Vamos utilizar esse resultado nos capítulos seguintes na determinação de distribuições de probabilidade de somas de variáveis aleatórias independentes e no Capítulo 7 na demonstração do teorema do limite central.

2.9 EXERCÍCIOS

1. Considere uma variável aleatória discreta T cuja distribuição de probabilidade é apresentada a seguir:

Valores de T	2	3	4	5	6	7
Probabilidades	1/10	1/10	4/10	2/10	1/10	1/10

Determine: (a) $P(T \geq 6)$; (b) $P(|T - 4| > 2)$; (c) $P(T$ ser um número primo).

2. Seja X uma variável aleatória discreta com distribuição de probabilidade dada por $P[X = x] = c2^{-x}$ para $x \in \mathbb{N} = \{0, 1, 2, \ldots\}$ (conjunto dos números naturais incluindo-se o zero) e zero no complementar. Determine: (a) O valor da constante c; (b) $P(X \leq 2)$; (c) $P(X > 5)$; (d) $P(X$ ser ímpar).

3. Suponha que uma variável aleatória discreta tenha a seguinte distribuição de probabilidade: $P[X = x] = cx$ para $x \in \{0, 1, \ldots, N\}$ e zero fora desse conjunto. Determine: (a) O valor da constante c quando $N = 4$; (b) O valor de c para um valor qualquer de N; (c) $P(X \leq a)$, $a \leq N$; (d) $P(X$ ser par).

4. Considere o lançamento de dois dados simultaneamente (admita que os dados são equilibrados). Para cada um dos itens a seguir, determine o campo de variação da variável aleatória X, bem como sua distribuição de probabilidades: (a) X é o maior valor observado no lançamento dos dois dados. (b) X é a soma dos valores observados. (c) X é o produto dos valores observados. (d) X é a diferença entre o maior valor observado e o menor valor observado.

5. Seja X uma variável aleatória com densidade $f(x) = cx^2$, $-1 \leq x \leq 1$, e $f(x) = 0$ caso contrário. (a) Determine o valor da constante c. (b) Calcule $P(|X| > \frac{1}{2})$. (c) Ache o valor de α tal que $F_X(\alpha) = P(X \leq \alpha) = \frac{1}{4}$ (o valor de α que satisfaz esta relação é denominado primeiro quartil da distribuição de X).

6. Seja X uma variável aleatória contínua cuja função densidade de probabilidade é dada por: $f(x) = \lambda e^{-\lambda x}$ para $x > 0$ e zero no complementar. (a) Verifique que $\int_{-\infty}^{\infty} f(x)dx = 1$. (b) Calcule $P(X > 6)$. (c) Esboce o gráfico de $f(x)$. (A variável aleatória X apresentada neste exercício possui distribuição exponencial de parâmetro λ. A distribuição exponencial é de grande importância em teoria da confiabilidade e será estudada de maneira mais detalhada no Capítulo 5.)

7. Dizemos que uma variável aleatória tem distribuição triangular no intervalo $[0, 1]$ se sua densidade é dada por: $f(x) = cx$ para $0 \leq x \leq \frac{1}{2}$, $f(x) = c(1 - x)$ para $\frac{1}{2} < x < 1$, e $f(x) = 0$ para os demais valores de x. (a) Determine o valor da constante c. (b) Esboce o gráfico de $f(x)$. (c) Calcule $P(X > \frac{8}{10})$ e $P(\frac{1}{4} < X < \frac{3}{4})$.

8. Considere a variável aleatória cuja função densidade de probabilidade é: $f(x) = c.\cos x$, $\frac{\pi}{2} \leq x \leq \pi$, e $f(x) = 0$ caso contrário. (a) Determine o valor de c. (b) Calcule $P(X < \frac{3\pi}{4})$.

9. Considere novamente as questões de 1 a 8. Determine, em cada caso, $E(X)$ e $\mathrm{Var}(X)$ (para o exercício 3, utilize as igualdades $\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}$ e $\sum_{k=1}^{N} k^3 = [\frac{N(N+1)}{2}]^2$).

10. Demonstre a generalização do resultado apresentado no lema 2.5.2, isto é, se $X_1, \ldots, X_n$ são variáveis aleatórias com esperança finita e $\alpha_1, \ldots, \alpha_n$ constantes reais, mostre que

    $E\left[\sum_{i=1}^{n} \alpha_i X_i\right] = \sum_{i=1}^{n} \alpha_i E[X_i]$.

11. Mostre, para toda variável aleatória X contínua, que: (a) Se existe uma constante α tal que $P(X \geq \alpha) = 1$, então $E(X) \geq \alpha$. (b) Se existe uma constante β tal que $P(X \leq \beta) = 1$, então $E(X) \leq \beta$. (c) Se $X \geq Y$ (isto é, $X(\omega) \geq Y(\omega)$, $\forall \omega \in \Omega$), então $E(X) \geq E(Y)$, onde Y é uma variável aleatória contínua. (Verifique se os resultados em (a), (b) e (c) continuam válidos se X e Y são discretas.)

12. Seja X uma variável aleatória com $E(X^2) < \infty$ e sejam α e β constantes reais. (a) Mostre que $\mathrm{Var}(\alpha X + \beta) = \alpha^2 \mathrm{Var}(X)$. (b) Calcule $E[(\beta X + 4)^2]$ se $E(X) = 10$ e $\mathrm{Var}(X) = 3$.

13. Considere uma variável aleatória X tal que $E(X^2) < \infty$. Seja $f(a) = E[(X - a)^2]$ uma função de $\mathbb{R}$ em $\mathbb{R}$. (a) Determine o valor a^* de A que minimiza $f(a)$. (b) Calcule $f(a^*)$, onde a^* é o valor de a obtido no item anterior. (Interprete estes resultados.)

14. Considere dois lançamentos consecutivos de um dado que suporemos equilibrado. Sejam X: número de vezes em que é obtida

a face 1, $x = 0, 1, 2$. Y: número de vezes em que é obtida a face 6, $y = 0, 1, 2$. $Z = X + Y$: número de vezes em que aparece ou uma face 1 ou uma face 6, $z = 0, 1, 2$. Determine: $\text{Var}(X)$, $\text{Var}(Y)$ e $\text{Var}(Z)$.

É verdade que $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$?

15. Seja X uma variável aleatória cuja função densidade é dada por: $f(x) = \frac{2}{x^3}$ para $1 \leq x \leq \infty$ e zero no complementar. Verifique que $E(X) = 2$, mas não existe $\text{Var}(X)$ (isto é, $\text{Var}(X) = \infty$). (Note que é possível construir uma variável aleatória discreta X_1' semelhante à descrita acima, tal que $E(X') < \infty$ e $\text{Var}(X') = \infty$.)

16. Dizemos que a distribuição da variável aleatória X é simétrica (em torno do zero) se $P(X \geq x) = P(X \leq -x)$ para todo $x \in \mathbb{R}$. Mostre que: (a) Se X é simétrica e $E(|X|) < \infty$, então $E(X) = 0$. (b) Se X é simétrica em torno de μ e $E(|X|) < \infty$, então $E(X) = \mu$ (analogamente a definição apresentada, X é simétrica em torno de μ se $P(X \geq \mu + x) = P(X \leq \mu - x)$, $\forall x \in \mathbb{R}$.

17. Considere novamente os exercícios de 5 a 8. Em cada caso, determine a função geradora de momentos de X. Obtenha então, a partir da função geradora, $E(X)$ e $\text{Var}(X)$.

18. Calcule a função geradora de momentos da variável aleatória X (discreta) quando: (a) $f(x) = 1/6$, se $x = 1$ ou $x = -1$, e $f(x) = 4/6$, se $x = 0$. (b) $f(x) = (1/2)^x$, $x = 1, 2, 3, \ldots$ (c) $f(x) = (1-p)^{x-1}p$, $x = 1, 2, \ldots$, $0 < p < 1$. (A variável aleatória X com distribuição dada em (c) é denominada geométrica (ou possui distribuição geométrica) de parâmetro p. Esta variável será estudada de forma mais detalhada no capítulo 4.)

19. Definimos a função geradora de probabilidades da variável aleatória X (inteira não-negativa) como sendo: $\varphi_X(t) = E(t^X)$, $t \in \mathbb{R}$. (a) Verifique que $\varphi_X(t) = \phi_X(\log t)$ e que $\phi_X(t) = \varphi_X(e^t)$, onde $\phi_X(t)$ é a função geradora de momentos de X. (b) Mostre que $\frac{d\varphi_X(t)}{dt}|_{t=1} = E(X)$. (c) $\frac{d^{(n)}\varphi_X(t)}{dt^{(n)}}|_{t=0} = n!P(X = n)$.

20. Seja X uma variável aleatória. A função característica de X é a função $\psi : \mathbb{R} \to \mathbf{C}$ dada por $\psi_X(t) = E(e^{itX})$, onde $i = \sqrt{-1}$. Lembrando que $e^{itX} = \cos tX + i\text{sen}tX$ e $|a + bi|^2 = a^2 + b^2$, verifique que a função característica é limitada por 1, isto é, $|\psi_X(t)| \leq 1$, $\forall t \in \mathbb{R}$. A função característica de variáveis aleatórias é de grande importância em teoria das probabilidades, dentre outros motivos, por ser limitada e portanto sempre existir. Os pré-requisitos adotados não nos permitem tratar da função característica nesse texto.

21. Seja X uma variável aleatória cuja função geradora de momentos é ϕ_1; seja $Y = aX + b$, onde a e b são constantes reais e, por fim, seja ϕ_2 a função geradora de momentos de Y. Então, para todo valor de t tal que $\phi_1(at)$ existe $\phi_2(t) = e^{bt}\phi_1(at)$.

22. Suponha que X é uma variável aleatória com função geradora de momentos ϕ, média μ e variância σ^2. Seja $c(t) = \log \phi(t)$. Verifique que $c'(0) = \mu$ e $c''(0) = \sigma^2$.

23. Seja X uma variável aleatória cuja função de distribuição é dada por: $F(x) = 0$, $x < 0$, $F(x) = x/3$, $0 \leq x < 1$, $F(x) = x/2$, $1 \leq x < 2$, $F(x) = 1$, $x \geq 2$. (a) Esboce o gráfico de F e verifique se X é uma variável aleatória contínua. (b) Calcule $P(\frac{1}{2} \leq X \leq \frac{3}{2})$, $P(\frac{1}{2} \leq X < \frac{3}{2})$, $P(\frac{1}{2} \leq X \leq 1)$, $P(\frac{1}{2} \leq X < 1)$, $P(1 < X < 2)$ e $P(1 \leq X \leq 2)$.

24. Se a função de distribuição F da variável aleatória X fosse definida de uma das formas abaixo, descreva como as condições (a), (b) e (c) do lema 2.4.1 deveriam ser modificadas em cada caso: i) $F(x) = P(x < x)$, ii) $F(x) = P(X > x)$, iii) $F(x) = P(X \geq x)$.

25. Sejam $f_1, f_2, \ldots, f_n$ funções densidade de probabilidade em $\mathbb{R}$. Verifique se as funções apresentadas abaixo são funções densidade de probabilidade.
(a) $\alpha f_1(x) + (1-\alpha)f_2(x)$, $0 < \alpha < 1$.
(b) $f_i(x) + f_j(x)$, $i, j = 1, 2, \ldots, n$ e $i \neq j$.
(c) $\alpha_1 f_1(x) + \cdots + \alpha_n f_n(x)$, $\alpha_i \geq 0$, $i = 1, \ldots, n$ e $\sum_{i=1}^n \alpha_i = 1$.

26. Cinco bolas são selecionadas aleatoriamente sem reposição de uma urna contendo N bolas numeradas de 1 até N, $N > 5$. Seja X a variável aleatória que denota o maior valor selecionado, determine a função de distribuição de X.

27. De um conjunto de N elementos, um subconjunto não-vazio é escolhido aleatoriamente (considere que todos os subconjuntos não-vazios têm a mesma probabilidade de serem escolhidos). Seja X o número de elementos contidos no subconjunto escolhido, determine a função de distribuição de X, bem como $E(X)$ e $\text{Var}(X)$. Verifique que

$$\lim_{N\to\infty} \frac{E(X)}{N} = \frac{1}{2} \quad \text{e} \quad \lim_{N\to\infty} \frac{\text{Var}(X)}{N} = \frac{1}{4}.$$

28. Para qualquer valor de $p > 1$, seja $c(p) = \sum_{x=1}^{\infty} \frac{1}{x^p}$. Suponha que X é uma variável aleatória discreta com a seguinte distribuição de probabilidade: $f(x) = \frac{1}{c(p)x^p}$, $x = 1, 2, \ldots$ (a) Para qualquer inteiro positivo n, determine a probabilidade de X ser divisível por n. (b) A probabilidade de X ser ímpar.

29. Um livro para apostadores recomenda a seguinte estratégia que afirma ser vencedora para o jogo de roleta. Aposte R$ 1,00 nas vermelhas. Se o resultado do giro da roleta for vermelha (cuja probabilidade é $\frac{18}{38}$), então retire seu benefício de R$ 1,00 e saia do jogo. Se perder a aposta (cuja probabilidade é $\frac{20}{38}$), faça uma aposta adicional de R$ 1,00 em vermelha em cada um dos próximos dois giros da roleta, e então saia. Seja X o seu ganho ao deixar o jogo de roleta. (a) Determine $P(X > 0)$. (b) Você acha que, de fato, esta é uma estratégia vencedora? Justifique. (c) Calcule $E(X)$.

30. A variável aleatória Y tem distribuição de probabilidade dada por

$$P(Y = k) = \frac{1}{k!} \int_0^{\infty} f(t) t^k e^{-t} dt, \quad k = 0, 1, 2, \ldots$$

onde $f(t)$ é a função densidade de probabilidade de uma variável aleatória com média μ e variância σ^2. Calcule $E(Y)$ e $\text{Var}(Y)$.

31. Uma urna contém bolas numeradas de 1 a N. Uma pessoa retira uma bola e a devolve, retira uma segunda bola e a devolve, e procede desta forma até obter uma bola pela segunda vez, isto é, até obter uma bola já retirada anteriormente. Seja X o número total de extrações necessárias para obter esta repetição,
(a) Obtenha a distribuição de X (dica: calcule $P(X > k)$).
(b) Mostre que $E(X) = 2 + \left(1 - \frac{1}{n}\right) + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right)$.

32. Considere um ponto escolhido aleatoriamente no interior de um disco de raio R no plano. Seja X o quadrado da distância do centro do disco ao ponto escolhido, obtenha a função de distribuição de X.

33. Considere um ponto escolhido uniformemente no interior de um triângulo de base B e altura H. Seja X a distância da base ao ponto escolhido, obtenha a função densidade de probabilidade de X.

34. Suponha que um ponto é escolhido ao acaso em um palito de comprimento unitário e que o palito é quebrado em dois pedaços neste ponto. Determine o valor esperado do comprimento do maior pedaço.

35. Suponha que X denote a posição de um ponto escolhido uniformemente no intervalo $(0, 5)$. Seja Y uma variável aleatória tal que $Y = 0$ se $X \leq 1$, $Y = 5$ se $X \geq 3$ e $Y = X$ caso contrário, determine a função de distribuição de Y.

36. Seja X uma variável aleatória com função densidade de probabilidade f. Seja c um número real fixo. Determine a função de distribuição de $Y = \max\{X, c\}$ quando:
(a) $f(x) = x$, $0 \leq x \leq 1$ e $c = \frac{1}{2}$,
(b) $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$ e $c = \lambda$.

37. Suponha que a densidade da variável aleatória X é dada por $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$. Seja Y a variável aleatória inteira definida por $Y = m$, se $m < X \leq m + 1$, onde m é inteiro não-negativo, (a) Obtenha a distribuição de Y. (b) Escreva $P(Y = m)$ da forma $P(Y = m) = p^m(1 - p)$, para algum p, $0 < p < 1$.

38. Para um grupo de n pessoas, determine o número esperado de dias do ano que são aniversário de exatamente k pessoas, $k \leq n$. (Suponha que o ano tem 365 dias e que todos os arranjos são equiprováveis.)

39. Um homem possui em seu chaveiro n chaves e deseja abrir a porta de sua casa experimentando as chaves ao acaso e independentemente. Determine a média e a variância do número de tentativas se (a) As chaves incorretas são descartadas e, conseqüentemente, não mais selecionadas. (b) As chaves incorretas não são separadas, podendo ser escolhidas novamente. Admita que apenas uma chave consegue abrir a porta.

40. Um produto de venda sazonal traz um lucro de B reais por unidade vendida e um prejuízo de L reais por cada unidade estocada e não vendida. O número de unidades deste produto vendidas em um certo supermercado (X) é uma variável aleatória com distribuição de probabilidade $p(i) = P(X = i)$, $i \geq 0$. Supondo que o estoque é feito no início da estação, que não pode haver reposição do produto durante a mesma e que não há devolução das unidades estocadas, determine o nível de estoque que maximiza o lucro esperado do supermercado.

41. Um florista faz estoque de uma flor de curta duração que lhe custa R$ 0,50 e que ele vende a R$ 1,50 no primeiro dia em que a flor está na loja. Toda flor que não é vendida nesse primeiro dia é jogada fora. Seja X a variável aleatória que denota o número de flores vendidas por este florista em um dia. Sabendo-se que a função de probabilidade de X é dada por

X	0	1	2	3
$P(X = x)$	1/10	4/10	3/10	2/10

determine quantas flores deveria o florista ter em estoque a fim de maximizar o lucro esperado.

42. A cada noite, diferentes meteorologistas nos fornecem a "probabilidade" de que irá chover no dia seguinte. Para avaliar o quão boas são estas previsões, é possível atribuir-se escores a cada um desses meteorologistas, como segue: se o meteorologista diz que irá chover no dia seguinte com probabilidade p, então ele receberá um escore de $1 - (1 - p)^2$ se chover no dia seguinte e de $1 - p^2$ se não chover.

    Acompanhando os escores obtidos pelos meteorologistas durante um certo período de tempo, podemos concluir que o meteorologista com maior escore médio é aquele que melhor prediz o tempo. Suponha agora que um certo meteorologista esteja ciente deste procedimento de escores e deseje maximizar seu escore esperado num determinado dia. Se este meteorologista acredita que de fato irá chover no dia seguinte com probabilidade p^*, qual o valor de p que ele deve dizer para maximizar seu escore esperado? (Interprete o resultado.)

3

Variáveis Aleatórias Discretas Multidimensionais

3.1 DISTRIBUIÇÕES DE PROBABILIDADE

Muitas vezes, em vez de observarmos uma única característica numérica associada aos resultados de um experimento, observamos várias. No Capítulo 2 definimos uma variável aleatória como uma função definida num espaço amostral que associa a cada resultado do experimento um número. Isto nos leva a definir variável aleatória multidimensional da seguinte maneira:

Definição 3.1.1 *Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias definidas no espaço amostral S. À n-upla $(X_1, X_2, \ldots, X_n)$ denominaremos variável aleatória n-dimensional. Ela associa a cada ponto ω de S a n-upla $(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$.*

Indicamos a seguir alguns exemplos. Estamos certos de que o leitor poderá facilmente exibir outros exemplos.

EXEMPLO 3.1.1 Suponha que num determinado dia sejam observados os pousos de aviões nos aeroportos de Guarulhos, Brasília e no Galeão. Seja X o número de aviões que pousam em Guarulhos, Y o número de aviões que pousam em Brasília e Z o número daqueles que pousam no Galeão. Temos assim uma variável aleatória tridimensional (X, Y, Z).

EXEMPLO 3.1.2 Numa pesquisa de opinião pública é escolhida uma amostra de 20 pessoas e a cada uma delas é perguntado se ela concorda com a criação de um imposto para custear despesas de educação e saúde. Podemos associar o número 1 a uma resposta positiva e zero a uma resposta negativa. Fica assim definida uma variável aleatória 20-dimensional, $(X_1, X_2, ..., X_{20})$, onde para i=1, 2,. .. , 20, $X_i = 1$ ou 0, conforme a resposta do i-ésimo elemento da amostra seja positiva ou negativa.

EXEMPLO 3.1.3 Na praça do pedágio em uma rodovia em uma certa hora existem cinco faixas de passagem abertas. Considere nesse instante o número de carros que se encontram em cada uma das faixas. Conte nesse número aqueles carros que estão efetuando o pagamento da taxa do pedágio. Temos uma variável aleatória de dimensão 5, dada por $(X_1, X_2, X_3, X_4, X_5)$, onde X_i é igual ao número de carros na faixa i, para $i = 1, 2, 3, 4$ e 5.

EXEMPLO 3.1.4 Um dado é lançado e simultaneamente é girada a roda de uma roleta. Seja X o resultado do lançamento do dado, Y o número observado na roleta. O par (X, Y) é uma variável bidimensional.

EXEMPLO 3.1.5 O questionário aplicado em cada uma das dez residências em uma pesquisa de domicílios contém as seguintes quatro perguntas: a) Quantos banheiros há na casa? b)Quantos carros os moradores possuem? c)Quantas televisões? Quantas geladeiras?

A cada casa fica associada uma quádrupla de inteiros: X_1 = número de banheiros, X_2 = número de carros, X_3 = número de televisões e X_4 = número de geladeiras. (X_1, X_2, X_3, X_4) é uma variável aleatória de dimensão 4.

Uma variável aleatória discreta multidimensional analogamente a uma variável aleatória unidimensional só fica determinada quando além de seus valores também são conhecidas as probabilidades com que esses valores são assumidos. A definição de distribuição de probabilidade que foi dada para uma variável aleatória unidimensional se estende de maneira natural para uma variável aleatória multidimensional.

DEFINIÇÃO 3.1.2 *A distribuição de probabilidade da variável aleatória $(X_1, X_2, \ldots, X_n)$ é uma tabela que associa a cada valor $(x_1, x_2, \ldots, x_n)$ dessa variável sua correspondente probabilidade $P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]$. Ela é denominada também distribuição conjunta de $X_1, X_2, \ldots, X_n$.*

Observemos que a notação $[X_1 = x_1, X_2 = x_2, ..., X_n = x_n]$ representa a interseção dos eventos $[X_1 = x_1], [X_2 = x_2], ...,$ $[X_n = x_n]$, ou seja,

$$[X_1 = x_1, \ldots, X_n = x_n] =$$

$$= \{\omega : X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n\}$$

$$= \{\omega : X_1(\omega) = x_1\} \cap \{\omega : X_2(\omega) = x_2\} \cap \cdots \cap \{\omega : X_n(\omega) = x_n\}.$$

EXEMPLO 3.1.6 Consideremos o conjunto das 32 cartas utilizadas para jogar pôquer, isto é, sete, oito, nove, dez, valete, dama, rei e ás, dos quatro naipes. Estas cartas são bem embaralhadas e duas cartas são dadas a uma pessoa. Considere as variáveis X igual ao número de ases que a pessoa recebe e Y igual ao número de cartas de copas que ela recebe. Vamos determinar a distribuição de probabilidade de (X, Y).

Mostremos que:

$$P[X = 0, Y = 0] = \binom{21}{2} \Big/ \binom{32}{2} = \frac{210}{496}.$$

O número de pontos do espaço amostral é $\binom{32}{2}$, pois com 32 cartas podemos formar esse número de subconjuntos com 2 elementos. Como não queremos que haja ases, nem cartas de copas, as duas

cartas devem ser retiradas das 21 restantes que não são nem copas nem ases. Esta escolha pode ser feita de $\binom{21}{2}$ maneiras. Mostremos agora que:

$$P[X=1,Y=1]=\frac{\binom{21}{1}+\binom{3}{1}\binom{7}{1}}{\binom{32}{2}}=\frac{42}{496}\,.$$

Basta justificarmos a quantia que aparece no numerador. Queremos calcular o número de pares de cartas que contem um ás e uma carta de copas. Se o ás for de copas, então a outra carta não pode ser de copas. Como existem 21 cartas entre as 32, que não são nem ases nem carta de copas, podemos escolher essa carta de $\binom{21}{1}$ maneiras. Por outro lado podemos receber um ás que não seja de copas de $\binom{3}{1}$ maneiras e uma carta de copas, que não pode ser ás de $\binom{7}{1}$ maneiras. Pelo princípio fundamental da contagem obtemos o numerador.

Os valores das demais probabilidades são dados abaixo e deixamos a cargo do leitor verificá-los.

$$\begin{aligned}
P[X=0,Y=1] &= \binom{7}{1}\binom{21}{1}\Big/\binom{32}{2}=\frac{147}{496}\\
P[X=0,Y=2] &= \binom{7}{2}\Big/\binom{32}{2}=\frac{21}{496}\\
P[X=1,Y=0] &= \binom{3}{1}\binom{21}{1}\Big/\binom{32}{2}=\frac{63}{496}\\
P[X=1,Y=2] &= \binom{7}{1}\Big/\binom{32}{2}=\frac{7}{496}\\
P[X=2,Y=0] &= \binom{3}{2}\Big/\binom{32}{2}=\frac{3}{496}\\
P[X=2,Y=1] &= \binom{3}{1}\Big/\binom{32}{2}=\frac{3}{496}\,.
\end{aligned}$$

Note que $P[X=2,Y=2]=0$, pois só há um ás de copas num baralho.

Definição

Sejam $p(x_i,y_j)$ para $i=1,2,\ldots$ e $j=1,2,\ldots$ números reais não negativos tais que sua soma seja igual a 1, isto é,

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p(x_i,y_j)=1\,.$$

Estes números podem ser considerados como a distribuição de probabilidade de uma variável aleatória bidimensional (X,Y), pondo-se: $P[X=x_i,Y=y_j]=p(x_i,y_j)$ para todo $i=1,2,\ldots$ e todo $j=1,2,\ldots$.

A noção de função de distribuição que foi introduzida para variáveis aleatórias unidimensionais na seção 2.4 pode ser estendida para variáveis aleatórias bidimensionais.

DEFINIÇÃO 3.1.3 *A função de distribuição do par de variáveis aleatórias discretas (X,Y) é dada por:*

$$F(x,y)=P[X\le x,Y\le y]=\sum_{i:x_i\le x}\sum_{j:y_j\le y}P[X=x_i,Y=y_j] \quad (3.1)$$

3.2 DISTRIBUIÇÕES MARGINAIS

Consideremos inicialmente uma variável aleatória bidimensional (X,Y). Suponhamos que X assume os valores $x_1,x_2,\ldots,x_n$ e que Y assume os valores $y_1,y_2,\ldots,y_m$. Seja $P[X=x_i,Y=y_j]$, para $i=1,2,\ldots,n$ e $j=1,2,\ldots,m$ a distribuição conjunta de (X,Y).

A distribuição de probabilidade de X é obtida calculando-se para $i=1,2,\ldots,n$:

$$P[X=x_i]=\sum_{j=1}^{m}P[X=x_i,Y=y_j]\,. \quad (3.2)$$

Ela é denominada distribuição marginal de X.

A distribuição marginal de Y é obtida de modo análogo. Para $j=1,2,\ldots,m$:

$$P[Y=y_j]=\sum_{i=1}^{n}P[X=x_i,Y=y_j]\,. \quad (3.3)$$

EXEMPLO 3.2.1 No exemplo 3.1.6 determinamos a distribuição conjunta de (X, Y), onde X é o número de ases e Y o número de copas em duas cartas retiradas do conjunto das 32.Vamos determinar as distribuições marginais de X e de Y.

X e Y assumem cada um deles os valores 0, 1 e 2. Utilizando as expressões (3.2) e (3.3) e os valores calculados no exemplo 3.1.6 obtemos as distribuições de probabilidade de X e de Y. Utilizando a expressão (3.2) e os valores das probabilidades calculadas no exemplo 3.1.6 tem-se

$$P[X=0] = P[X=0,Y=0]+P[X=0,Y=1]+P[X=0,Y=2]= \\ = \frac{210}{496}+\frac{147}{496}+\frac{21}{496} = \frac{378}{496}$$

$$P[X=1] = P[X=1,Y=0]+P[X=1,Y=1]+P[X=1,Y=2]= \\ = \frac{63}{496}+\frac{42}{496}+\frac{7}{496} = \frac{112}{496}$$

$$P[X=2] = P[X=2,Y=0]+P[X=2,Y=1]+P[X=2,Y=2]= \\ = \frac{3}{496}+\frac{3}{496} = \frac{6}{496}.$$

A distribuição marginal de Y é obtida de modo análogo. Essas distribuições são dispostas nas tabelas abaixo.

X	Probabilidade
0	378/496
1	112/496
2	6/496

Y	Probabilidade
0	276/496
1	192/496
2	28/496

EXEMPLO 3.2.2 Seja (X, Y) uma variável aleatória bidimensional cuja distribuição de probabilidade é dada na tabela de dupla entrada apresentada a seguir. Na primeira linha são apresentados os valores de Y e na primeira coluna os valores de X. No cruzamento de uma

linha com uma coluna, a tabela apresenta a probabilidade do valor de X correspondente à linha e do valor de Y correspondente à coluna.

$x \backslash y$	1	2	3	4	
1	0,10	0,05	0,02	0,07	0,24
2	0,08	0,05	0,10	0,19	0,42
3	0,10	0,20	0,04	–	0,34
	0,28	0,30	0,16	0,26	

Vamos determinar as distribuições marginais de X e de Y. Observe que no caso de uma variável aleatória bidimensional cuja distribuição conjunta é apresentada através de uma tabela de dupla entrada, a aplicação fórmula (3.2) para cada valor de X, por exemplo 1, corresponde a somar-se as probabilidades na linha correspondente a esse valor. Aplicando-se também esse procedimento para os valores 2 e 3 de X obtém-se sua distribuição na coluna que corresponde à margem da tabela. Procedimento análogo para Y nos dá na linha da margem inferior da tabela a distribuição marginal de Y.

Para uma variável aleatória bidimensional as fórmulas (3.2) e (3.3) nos dão a maneira de calcular as distribuições marginais. Este procedimento aplica-se a uma variável aleatória n-dimensional, onde n é um inteiro qualquer.

Seja $(X_1, X_2, ..., X_n)$ uma variável aleatória com distribuição conjunta dada por: $P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]$, onde $x_1, x_2, \ldots, x_n$, percorrem o conjunto de valores de $X_1, X_2, \ldots, X_n$, respectivamente.

A distribuição marginal de X_1 é obtida calculando-se para cada valor x_1 de X_1:

$$P[X_1 = x_1] = \\ = \sum_{x_2} \cdots \sum_{x_n} P[X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_n = x_n]. \quad (3.4)$$

Notemos ainda que podemos determinar a distribuição marginal de qualquer subconjunto de $(X_1, X_2, ..X_n)$. A título de ilustração vamos determinar a distribuição de probabilidade conjunta

de (X_1, X_2). Para x_1 e x_2 percorrendo o conjunto de valores de X_1 e X_2 tem-se:

$$P[X_1 = x_1, X_2 = x_2] =$$

$$= \sum_{x_3}\sum_{x_4}\cdots\sum_{x_n} P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] \quad (3.5)$$

EXEMPLO 3.2.3 Consideremos o conjunto das 32 cartas do jogo de pôquer definidas no exemplo 3.1.6. Duas cartas são distribuídas a um jogador. Considere as variáveis aleatórias X = número de ases, Y= número de copas e Z = número de ouros que o jogador recebe. Determine a distribuição conjunta de X, Y e Z.

Vamos calcular as probabilidades:

$$P[X = 0, Y = 0, Z = 0] \qquad \text{e} \qquad P[X = 1, Y = 0, Z = 1]$$

e deixar o cálculo das demais probabilidades como exercício para o leitor.

O número de pontos do espaço amostral, como já vimos, é $\binom{32}{2}$ =496. Determinemos o número de pontos do evento $[X = 0, Y = 0, Z = 0]$. A ocorrência desse evento implica que o jogador não recebe ases, nem cartas de copas e nem cartas de ouros. Como existem oito cartas de cada naipe entre as 32, o jogador deve receber duas cartas entre as 14 que são de espadas e paus, excluídos os ases. Isto pode ser feito de $\binom{14}{2}$ maneiras. Assim, tem-se:

$$P[X = 0, Y = 0, Z = 0] = \frac{\binom{14}{2}}{\binom{32}{2}} = \frac{91}{496} = 0,18\,.$$

O evento $[X = 1, Y = 0, Z = 1]$ ocorre de uma das duas maneiras seguintes:

a) O jogador recebe o ás de ouros e uma carta que é de paus ou de espadas, mas que não pode ser ás. Há 14 cartas que satisfazem essa condição, logo esta carta pode ser escolhida de $\binom{14}{1}$ maneiras.

b) O jogador recebe um ás que não é de ouros e uma carta de ouros, que não pode ser ás. Isto pode ser feito de $\binom{2}{1}$. $\binom{7}{1}$ maneiras.

Como os eventos descritos em a e b são mutuamente exclusivos tem-se:

$$P[X = 1, Y = 0, Z = 1] = \frac{\binom{14}{1} + \binom{2}{1}\binom{7}{1}}{\binom{32}{2}} = \frac{28}{496} = 0,06\,.$$

Vimos que a distribuição conjunta de uma variável aleatória n-dimensional determina as distribuições marginais de cada uma de suas componentes $X_1, X_2, \ldots, X_n$ e também de qualquer subconjunto dessas variáveis.

Uma questão de interesse consiste em saber se o conhecimento das distribuições marginais das variáveis aleatórias $X_1, X_2, \ldots, X_n$ permite determinar sua distribuição conjunta.

A resposta a essa questão é negativa. Sem informações adicionais sobre a relação entre essas variáveis não se pode determinar a distribuição conjunta a partir das marginais.

3.3 VARIÁVEIS ALEATÓRIAS INDEPENDENTES

DEFINIÇÃO 3.3.1 *As variáveis aleatórias X_1, X_2, ... X_n são ditas independentes se para todos os valores $x_1, x_2, \ldots, x_n$ das mesmas tivermos:*

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] =$$

$$= P[X_1 = x_1]P[X_2 = x_2]\cdots P[X_n = x_n]\,. \quad (3.6)$$

Observe na expressão (3.6) que, se as variáveis são independentes, então a distribuição conjunta é obtida como produto das marginais, logo, nesse caso as marginais determinam a conjunta.

Nos modelos de urnas quando as retiradas de bolas da urna são feitas com reposição, as variáveis aleatórias associadas a cada retirada são independentes e podemos obter a distribuição conjunta como produto das marginais.

EXEMPLO 3.3.1 Uma urna contém três bolas vermelhas e cinco brancas. Retiram-se sucessivamente três bolas da urna, repondo-se a bola selecionada após cada retirada. Vamos definir as variáveis $X_i = 1$ se

a i-ésima bola retirada for vermelha e 0 se for branca, para $i = 1, 2$ ou 3 e determinar a distribuição conjunta de X_1, X_2 e X_3.

A fórmula (3.6) exprime a distribuição conjunta como produto das marginais, já que X_1, X_2 e X_3 são independentes. Temos por exemplo:

$$P[X_1 = 1, X_2 = 1, X_3 = 1] = P[X_1 = 1]P[X_2 = 1]P[X_3 = 1] = (\frac{3}{8})^3.$$

Usando-se esta fórmula para calcular as probabilidades dos demais valores de X_1, X_2 e X_3, obtemos a distribuição conjunta de X_1, X_2 e X_3.

(X_1, X_2, X_3)	*Probabilidades*
$(1,1,1)$	$(\frac{3}{8})^3$
$(1,1,0)$	$(\frac{3}{8})^2\frac{5}{8}$
$(1,0,1)$	$(\frac{3}{8})^2\frac{5}{8}$
$(1,0,0)$	$\frac{3}{8}(\frac{5}{8})^2$

(X_1, X_2, X_3)	*Probabilidades*
$(0,1,1)$	$(\frac{3}{8})^2\frac{5}{8}$
$(0,1,0)$	$\frac{3}{8}(\frac{5}{8})^2$
$(0,0,1)$	$\frac{3}{8}(\frac{5}{8})^2$
$(0,0,0)$	$(\frac{5}{8})^3$

A distribuição conjunta das variáveis aleatórias $X_1, X_2, \ldots, X_n$ determina não só as distribuições marginais de qualquer subconjunto dessas variáveis, mas também as distribuições de funções de $X_1, X_2, \ldots, X_n$.

Vamos ilustrar essa afirmação retomando o exemplo 3.3.1 que acabamos de apresentar.

EXEMPLO 3.3.2 Consideremos o exemplo 3.3.1 e definamos a variável aleatória $S_3 = X_1 + X_2 + X_3$, que é igual ao número de bolas vermelhas entre as três retiradas. Determinemos a distribuição de probabilidade de S_3.

S_3 assume os valores 0, 1, 2 e 3. Seja k o valor de S_3. Para calcular $P[S_3 = k]$ determinamos as triplas (x_1, x_2, x_3) tais que $S_3 = x_1 + x_2 + x_3 = k$ e somamos as probabilidades dessas triplas. Por exemplo para $k = 2$ temos:

$$P[S_3 = 2] = P[X_1 = 1, X_2 = 1, X_3 = 0] + P[X_1 = 1, X_2 = 0, X_3 = 1] + \\ + P[X_1 = 0, X_2 = 1, X_3 = 1] = 3(\frac{3}{8})^2\frac{5}{8}.$$

Observe que podemos calcular essas probabilidades a partir da tabela obtida no exemplo 3.3.1 que dá a distribuição conjunta de X_1, X_2, X_3. Basta identificar as triplas onde S_3 é igual ao valor desejado e somar as probabilidades. Assim obtemos:

S_3	Probabilidade
0	$(\frac{5}{8})^3$
1	$3(\frac{3}{8})(\frac{5}{8})^2$
2	$3(\frac{3}{8})^2(\frac{5}{8})$
3	$(\frac{3}{8})^3$

LEMA 3.3.1 *As variáveis aleatórias X e Y, com função de distribuição $F(x, y)$ são independentes se e somente se*

$$F(x, y) = F_X(x)F_Y(y) \tag{3.7}$$

onde $F_X(x)$ e $F_Y(y)$ são as funções de distribuição marginais de X e de Y.

DEMONSTRAÇÃO Suponhamos que X e Y são independentes, isto é, que a expressão (3.6) é válida.

$$\begin{aligned} F(x,y) &= \sum_{i:x_i \le x} \sum_{j:y_j \le y} P[X = x_i, Y = y_j] = \\ &= \sum_{i:x_i \le x} \sum_{j:y_j \le y} P[X = x_i]P[Y = y_j] = \\ &= \sum_{i:x_i \le x} P[X = x_i] \sum_{j:y_j \le y} P[Y = y_j] = F_X(x)F_Y(y)\,. \end{aligned}$$

A demonstração da recíproca fica a cargo do leitor.

LEMA 3.3.2 *Se X e Y são variáveis aleatórias independentes tem-se:*

$$E(XY) = E(X)E(Y)\,. \tag{3.8}$$

DEMONSTRAÇÃO Suponhamos que a distribuição conjunta de (X,Y) seja dada por

$$P[X=x_i, Y=y_j] \text{ para } i=1,2,3,\ldots \text{ e } j=1,2,3\ldots$$

$$\begin{aligned} E(XY) &= \sum_1^\infty \sum_1^\infty x_i y_j P[X=x_i, Y=y_j] = \\ &= \sum_1^\infty \sum_1^\infty x_i y_j P[X=x_i]P[Y=y_j] = \\ &= \sum_1^\infty x_i P[X=x_i] \sum_1^\infty y_j P[Y=y_j] = E(X)E(Y)\,. \end{aligned}$$

Uma das razões fundamentais para a introdução da função geradora de momentos de uma variável aleatória é a propriedade que apresentaremos a seguir, que mostra que a função geradora de momentos de uma soma de variáveis aleatórias independentes é igual ao produto das funções geradoras de momento das parcelas.

LEMA 3.3.3 *Sejam X_1, X_2, .. , X_n variáveis aleatórias independentes cujas funções geradoras de momentos são denotadas ϕ_{X_i}, para $i=1,2,...,n$. A função geradora de momentos de $X_1+X_2+\cdots+X_n$ que designaremos por $\phi_{X_1+X_2+\cdots+X_n}$ satisfaz, para todo t:*

$$\phi_{X_1+X_2+..+X_n}(t) = \phi_{X_1}(t)\phi_{X_2}(t)\cdots\phi_{X_n}(t). \tag{3.9}$$

DEMONSTRAÇÃO

$$\begin{aligned} \phi_{X_1+X_2+\cdots+X_n}(t) &= E(e^{t(X_1+X_2+\cdots+X_n)}) = E(e^{tX_1}e^{tX_2}\cdots e^{tX_n}) = \\ &= E(e^{tX_1})E(e^{tX_2})..E(e^{tX_n}) = \\ &= \phi_{X_1}(t)\phi_{X_2}(t)\cdots\phi_{X_n}(t)\,. \end{aligned}$$

A última igualdade decorre do lema 3.3.2 já que funções de variáveis aleatórias independentes são também variáveis aleatórias independentes.

3.4 COVARIÂNCIA E COEFICIENTE DE CORRELAÇÃO

Seja (X,Y) uma variável aleatória bidimensional. As esperanças das componentes X e Y nos fornecem uma medida de posição das respectivas distribuições nos respectivos eixos das coordenadas do plano. As variâncias de X e de Y dão uma medida da dispersão dos valores de X e de Y em torno das respectivas médias $E(X)$ e $E(Y)$.

A covariância, cuja definição damos a seguir, dá uma idéia da dispersão dos valores da variável bidimensional (X,Y) em relação ao ponto $(E(X),E(Y))$.

DEFINIÇÃO 3.4.1 *Seja (X,Y) uma variável aleatória bidimensional. A covariância de X e Y que denotaremos* $\mathrm{Cov}(X,Y)$ *é definida por:*

$$\mathrm{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]\,. \tag{3.10}$$

Observe que a esperança na expressão da definição da covariância refere-se à distribuição conjunta de (X,Y). Explicitamente temos:

$$\mathrm{Cov}(X,Y) = \sum_1^\infty \sum_1^\infty (x_i - E(X))(y_j - E(Y))P[X=x_i, Y=y_j]\,. \tag{3.11}$$

Vê-se da definição que a covariância é a esperança dos produtos dos desvios dos valores de cada uma das duas variáveis em relação a suas médias. Se as duas variáveis tendem a variar no mesmo sentido, isto é, valores de X acima de sua média estão associados a valores de Y acima de sua média, o mesmo ocorrendo para valores de ambos inferiores à média, então a covariância será positiva. Caso contrário, se valores acima da média de uma variável estão associados a valores inferiores à média da outra, então a covariância será negativa.

A expessão (3.10) pode ser reescrita da seguinte maneira:

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)\,. \tag{3.12}$$

De fato, efetuando-se o produto na expessão (3.10) tem-se:

$$\begin{aligned} \text{Cov}(X,Y) &= E(XY - XE(Y) - YE(X) + E(X)E(Y)) = \\ &= E(XY) - E(X)E(Y) . \end{aligned}$$

LEMA 3.4.1 *Se X e Y são variáveis aleatórias independentes* $\text{Cov}(X,Y) = 0$.

DEMONSTRAÇÃO A demonstração é imediata, pois observe que, se X e Y são independentes, o lema 3.3.1 garante que a esperança do produto das variáveis aleatórias é igual ao produto das esperanças, e substituindo em (3.12) tem-se $\text{Cov}(X,Y) = 0$.

EXEMPLO 3.4.1 Considere uma urna com três bolas brancas e duas bolas vermelhas. Retiram-se duas bolas da urna, uma após a outra sem reposição. Defina a variável aleatória X igual a 1 se a primeira bola retirada é branca, e igual a 0 se esta é vermelha. Analogamente, defina Y igual a 1 se a segunda bola é branca, e 0 se é vermelha. Vamos calcular a covariância entre X e Y.

Vamos inicialmente determinar a distribuição conjunta de X e Y.

$x \backslash y$	0	1
0	$\frac{2}{20}$	$\frac{6}{20}$
1	$\frac{6}{20}$	$\frac{6}{20}$

Temos $E(X) = E(Y) = 3/5$ e $E(XY) = 6/20$. Substituindo-se esses valores na expressão (3.10), obtemos $\text{Cov}(X,Y) = -3/50$.

Observemos que o valor da covariância entre duas variáveis aleatórias X e Y depende das unidades de medida adotadas para medir essas variáveis. De fato se fizermos uma mudança de escala pondo $X' = aX$ e $Y' = bY$ e calcularmos a covariância de X' e de Y', teremos:

$$\begin{aligned} \text{Cov}(X',Y') &= E[(aX - E(aX))(bY - E(bY))] = \\ &= abE[(X - E(X))(Y - E(Y))] = ab\text{Cov}(X,Y) . \end{aligned}$$

Nestas condições é de interesse introduzir-se um conceito cujo valor seja independente da unidade de medida. Seja $\sigma(X)$ o desvio padrão da variável aleatória X.

DEFINIÇÃO 3.4.2 *O coeficiente de correlação das variáveis aleatórias X e Y, denotado $\rho(X,Y)$, é definido por:*

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\sigma(Y)} . \tag{3.13}$$

LEMA 3.4.2 *O coeficiente de correlação $\rho(X,Y)$ das variáveis aleatórias X e Y satisfaz a desigualdade:*

$$|\rho(X,Y)| \leq 1 . \tag{3.14}$$

Se $\rho(X,Y) = \pm 1$, então os valores de X e Y pertencem a uma reta.

DEMONSTRAÇÃO Consideremos a expressão $a(X - E(X)) + (Y - E(Y))$, onde a é um número real e calculemos:

$$\begin{aligned} E(a(X - E(X)) + (Y - E(Y))^2 &= a^2E(X - E(X))^2 + \\ &+2aE[(X - E(X))(Y - E(Y))] + E(Y - E(Y))^2 = \\ &= a^2\sigma^2(X) + 2a\text{Cov}(X,Y) + \sigma^2(Y) \geq 0 \end{aligned}$$

pois é a esperança de uma variável aleatória não negativa.

Se o trinômio é não negativo então o seu discriminante:

$$\Delta = 4\text{Cov}^2(X,Y) - 4\sigma^2(X)\sigma^2(Y) \leq 0.$$

Dividindo-se por $4\sigma^2(X)\sigma^2(Y)$ tem-se: $\rho^2(X,Y) \leq 1$, logo $|\rho(X,Y)| \leq 1$.

Fica assim demonstrada a desigualdade (3.14).

Analisemos o que ocorre quando $\rho = 1$.

Se $\rho(X,Y) = 1$, decorre de (3.13) que $\text{Cov}(X,Y) = \sigma(X)\sigma(Y)$, substituindo-se, vemos que o discriminante Δ=0 e portanto o trinômio tem uma raiz dupla que é igual a

$$-\frac{2\text{Cov}(X,Y)}{2\sigma^2(X)} = -\frac{\sigma(Y)}{\sigma(X)} .$$

Substituindo esse valor de $a = -\frac{\sigma(Y)}{\sigma(X)}$ na expressão do trinômio, tem-se:

$$E(-\frac{\sigma(Y)}{\sigma(X)}(X - E(X)) + (Y - E(Y)))^2 = 0 .$$

Mas uma variável aleatória não negativa tem esperança nula se e somente se ela for identicamente nula. Portanto:

$$-\frac{\sigma(Y)}{\sigma(X)}(X - E(X)) + (Y - E(Y)) = 0 .$$

Reescrevendo, tem-se: $Y - E(Y) = \frac{\sigma(Y)}{\sigma(X)}(X - E(X))$ que é a equação da reta mencionada.

Se $\rho(X, Y) = -1$ procedemos de modo análogo.

EXEMPLO 3.4.2 Sejam X, Y variáveis aleatórias que assumem os valores -1, 0, 1 cuja distribuição conjunta é dada por

$$\begin{aligned} P[X=-1, Y=0] &= P[X=1, Y=0] = P[X=0, Y=-1] = \\ &= P[X=0, Y=1] = \frac{1}{4} . \end{aligned}$$

Calcular $cov(X, Y)$ e verificar que X e Y são dependentes. É fácil ver que X e Y são identicamente distribuídas. Determinemos a distribuição marginal de X. Tem-se:

$$P[X = -1] = \frac{1}{4}, \quad P[X = 0] = \frac{1}{2}, \quad P[X = 1] = \frac{1}{4}.$$

Daí decorre que $E(X) = 0$ e portanto $E(Y) = 0$. Mas $E(XY) = 0$ pois em todo ponto ou $X = 0$ ou $Y = 0$. Desta maneira, $cov(X, Y) = E(XY) - E(X)E(Y) = 0$.

X e Y são dependentes pois se $X = -1$ então $Y = 0$.

3.5 DISTRIBUIÇÕES CONDICIONAIS

Vamos inicialmente tratar do caso de duas variáveis aleatórias e posteriormente indicar como se estendem os conceitos para uma variável n-dimensional, com $n > 2$.

Seja (X, Y) uma variável aleatória bidimensional cuja distribuição conjunta associa a cada par de valores (x, y) dessas variáveis a probabilidade $P[X = x, Y = y]$.

DEFINIÇÃO 3.5.1 *A distribuição condicional de Y dado X associa a cada valor x de X uma distribuição de probabilidade sobre os valores de Y, da seguinte maneira:*

$$P[Y = y|X = x] = \frac{P[X = x, Y = y]}{P[X = x]} . \qquad (3.15)$$

Observamos que, quando nos referimos à distribuição condicional de Y dado X, estamos nos referindo ao conjunto de todas essas distribuições condicionais associadas a cada um dos valores de X.

EXEMPLO 3.5.1 Consideremos a variável aleatória bidimensional (X, Y) do exemplo 3.2.2. Vamos determinar a distribuição condicional de Y dado X.

Nas três colunas abaixo apresentamos as distribuições condicionais de Y dado $X = 1$, $X = 2$ e $X = 3$. Vamos aplicar a fórmula (3.15) para calcular essas distribuições condicionais. Da primeira linha da tabela do exemplo 3.2.2 obtemos as probabilidades conjuntas de $[X = 1, Y = 1]$, $[X = 1, Y = 2]$, $[X = 1, Y = 3]$ e $[X = 1, Y = 4]$. Dividindo-se essas probabilidades por $P[X = 1]$, que está na margem da tabela, obtemos a primeira coluna que dá a distribuição condicional de Y dado $X = 1$. Prosseguindo-se de modo análogo para $X = 2$ e $X = 3$, completamos a tabela a seguir.

$P[Y = 1|X = 1] = \frac{10}{24}$ $P[Y = 1|X = 2] = \frac{8}{42}$ $P[Y = 1|X = 3] = \frac{10}{34}$

$P[Y = 2|X = 1] = \frac{5}{24}$ $P[Y = 2|X = 2] = \frac{5}{42}$ $P[Y = 2|X = 3] = \frac{20}{34}$

$P[Y = 3|X = 1] = \frac{2}{24}$ $P[Y = 3|X = 2] = \frac{10}{42}$ $P[Y = 3|X = 3] = \frac{4}{34}$

$P[Y = 4|X = 1] = \frac{7}{24}$ $P[Y = 4|X = 2] = \frac{19}{42}$ $P[Y = 4|X = 3] = 0.$

Observemos que, para cada valor x da variável aleatória X, temos a distribuição condicional de Y dado x. Podemos então para

cada x usar a expressão (2.9) para definir a esperança com relação a essa distribuição que é denominada esperança condicional de Y dado $X = x$.

DEFINIÇÃO 3.5.2 *A esperança condicional da variável aleatória Y dada a variável aleatória X, que denotaremos por $E(Y|X)$, é a função cujo valor para $X = x$ é:*

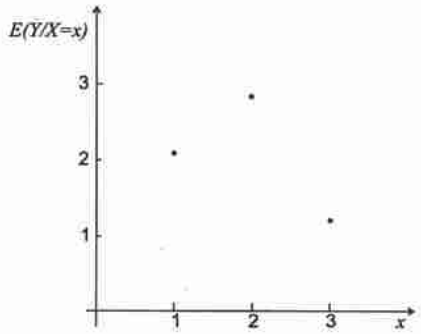
$$E(Y|X = x) = \sum_y yP[Y = y|X = x] \qquad (3.16)$$

onde a soma é estendida a todos os valores de Y.

EXEMPLO 3.5.2 Vamos retomar o exemplo 3.5.1 e determinar os valores de $E(Y|X = 1)$, $E(Y|X = 2)$, $E(Y|X = 3)$. Utilizando-se a expressão (3.16), obtemos:

$$\begin{aligned} E(Y|X=1) &= 1\frac{10}{24} + 2\frac{5}{24} + 3\frac{2}{24} + 4\frac{7}{24} = 2,25 \\ E(Y|X=2) &= 1\frac{8}{42} + 2\frac{5}{42} + 3\frac{10}{42} + 4\frac{19}{42} = 2,95 \\ E(Y|X=3) &= 1\frac{10}{34} + 2\frac{20}{34} + 3\frac{4}{34} = 1,58 \end{aligned}$$

Na figura abaixo representamos $E(Y|X)$.



EXEMPLO 3.5.3 Lança-se um dado sobre uma superfície plana. Denotemos por X o número de pontos observados na face superior do dado. Se $X = n$, então são lançadas n moedas sobre essa superfície. Seja Y o número de caras nos n lançamentos, vamos determinar $E(Y|X)$.

Se $X = n$, $1 \leq n \leq 6$, então são lançadas n moedas. Temos:

$$P[Y = j|X = n] = \binom{n}{j}(\tfrac{1}{2})^n \qquad \text{para } j = 0, 1, \ldots, n.$$

De fato, os lançamentos das moedas são independentes, então a probabilidade de j caras e $n - j$ coroas em posições prefixadas é $(\frac{1}{2})^n$. Como j caras em n lançamentos podem ocorrer de $\binom{n}{j}$ maneiras segue-se a expressão acima.

$$E(Y|X = n) = \sum_{j=0}^{n} jP[Y = j|X = n] = \sum_{j=0}^{n} j\binom{n}{j}(\tfrac{1}{2})^n$$

$$\sum_{j=0}^{n} j\binom{n}{j} = \sum_{j=0}^{n} j\frac{n!}{j!(n-j)!} = \sum_{j=1}^{n} \frac{n!}{(j-1)!(n-j)!} =$$

$$= \sum_{k=0}^{n-1} \frac{n!}{k!(n-(k+1)!} = n\sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} = n2^{n-1} .$$

Note que esta última igualdade decorre de

$$(1+1)^{n-1} = \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} .$$

Substituindo acima vem:

$$E(Y|X = n) = \frac{n}{2} .$$

Analisando a definição de $E(Y|X)$, nota-se que a esperança condicional sendo uma função de X é uma variável aleatória, cuja esperança coincide com $E(Y)$ conforme mostra o seguinte lema.

LEMA 3.5.1 *Sejam X e Y variáveis aleatórias. Tem-se:*

$$E(E(Y|X)) = E(Y) . \qquad (3.17)$$

DEMONSTRAÇÃO

$$E(E(Y|X)) = \sum_x E(Y|X=x)P[X=x] =$$

$$= \sum_x (\sum_y yP[Y=y|X=x])P[X=x] = \sum_x \sum_y yP[X=x, Y=y] =$$

$$= \sum_y y(\sum_x P[X=x, Y=y]) = \sum_y yP[Y=y] = E(Y).$$

EXEMPLO 3.5.4 Vamos retomar o exemplo 3.5.3 e calcular a esperança de Y usando o lema.

Lembramdo que $P[X=n] = \frac{1}{6}$ para $n = 1, 2, \ldots, 6$, usando o lema obtemos:

$$E(Y) = \sum_{n=1}^{6} E(Y|X=n)P[X=n] = \sum_{n=1}^{6} \frac{n}{2} \cdot \frac{1}{6} = 1,75$$

3.6 EXERCÍCIOS

1. Suponha que três bolas são selecionadas ao acaso de uma urna contendo três bolas vermelhas, quatro bolas brancas e cinco bolas azuis. Sejam X e Y respectivamente o número de bolas vermelhas e o número de bolas brancas escolhidas. Determine (a) A distribuição de probabilidade de (X, Y). (b) A distribuição de probabilidade de X e a distribuição de probabilidade de Y. (c) $P(X = Y)$.

2. Dois dados honestos são lançados. Obtenha a distribuição de probabilidade conjunta de (X, Y) quando:

   (a) X é o valor obtido no primeiro dado e Y é o maior valor observado.

   (b) X é o menor valor observado e Y é o maior valor observado.

   (c) X é o maior valor observado e Y é a soma dos valores obtidos.

3. Em um pacote com cinco transistores, dois deles são defeituosos. Os transistores devem ser testados, um de cada vez, até que aqueles defeituosos sejam identificados. Sejam N_1 o número de testes necessários para que o primeiro transistor defeituoso seja identificado e N_2 o número de testes adicionais necessários para se encontrar o segundo transistor defeituoso. Determine a distribuição de probabilidade de (N_1, N_2).

4. Uma urna contém três bolas numeradas 1, 2 e 3. Duas bolas são tiradas sucessivamente da urna, ao acaso e sem reposição. Seja X o número da primeira bola tirada e Y o número da segunda.
   (a) Descreva a distribuição conjunta de (X, Y).
   (b) Calcule $P(X < Y)$.

5. Considere um par de variáveis aleatórias discretas (X, Y) cuja função de distribuição é F, isto é, $F(x, y) = P(X \leq x, Y \leq y)$, $x, y \in \mathbb{R}$. Sejam F_X e F_Y as funções de distribuição das variáveis aleatórias X e Y, respectivamente. Mostre que: $P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F(x, y)$.

6. Considere a situação do exercício 1. Determine a probabilidade de se obter, dentre as três bolas selecionadas, pelo menos uma bola vermelha e pelo menos uma bola branca, através da expressão dada no exercício 5.

7. A distribuição de probabilidade conjunta de (X, Y) é dada por $p(1,1) = 1/8$, $p(1,2) = 1/4$, $p(2,1) = 1/8$, $p(2,2) = 1/2$. (a) Calcule $P(XY \leq 3)$, $P(X + Y > 2)$ e $P(X|Y > 1)$. (b) X e Y são independentes?

8. Considere novamente o exercício 1.
   (a) Calcule $\text{Cov}(X, Y)$.
   (b) As variáveis X e Y são independentes?

9. Um dado honesto é lançado uma única vez. Consideremos as variáveis $X = 1$, se o número observado é maior ou igual a 3, $X = 0$ caso contrário; $Y = 1$, se o número observado é par, $Y = 0$ caso contrário. (a) Determine a covariância entre X e Y. (b) X e Y são independentes?

10. Suponha que X e Y tenham distribuição conjunta dada pela seguinte tabela

$Y \backslash X$	-1	0	1
-1	0	1/5	0
0	1/5	1/5	1/5
1	0	1/5	0

(a) Determine a covariância entre X e Y. (b) X e Y são independentes?

11. Considere o seguinte experimento realizado em duas etapas. Primeiro escolhe-se um ponto x uniformemente em $\{0, 1, \ldots, 10\}$. Em seguida, escolhe-se um ponto uniformemente no conjunto $\{-x, -x+1, \ldots, 0, \ldots, x-1, x\}$. Seja (X, Y) o resultado do experimento, determine: (a) A distribuição de probabilidade de (X, Y). (b) A distribuição marginal de Y. (c) X e Y são independentes?

12. Suponha que $F(x)$ é uma função de distribuição. Mostre que: (a) $F^n(x)$ e (b) $1 - [1 - F(x)]^n$ são também funções de distribuição, onde n é um inteiro positivo.

    Dica: Considere $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes com mesma função de distribuição F e defina $Y = \max\{X_1, X_2, \ldots, X_n\}$ e $Z = \min\{X_1, X_2, \ldots, X_n\}$.

13. Sejam X e Y duas variáveis aleatórias discretas e f e g duas funções quaisquer. Mostre que se X e Y são independentes, então $E[g(X)f(Y)] = E[g(X)]E[f(Y)]$.

14. Verifique: (a) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$; (b) $\text{Cov}(X, X) = \text{Var}(X)$; (c) $\text{Cov}(aX, Y) = \text{Cov}(X, aY) = a\text{Cov}(X, Y)$, $a \in \mathbb{R}$; (d) Conclua que $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$, $a, b \in \mathbb{R}$.

15. Sejam $X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias. Mostre que

    (a) $\text{Cov}\left(\sum_{i=1}^n X_i, Y_j\right) = \sum_{i=1}^n \text{Cov}(X_i, Y_j)$, $\forall j = 1, \ldots, m$.

(b) $\text{Cov}\left(X_i, \sum_{j=1}^m Y_j\right) = \sum_{j=1}^m \text{Cov}(X_i, Y_j)$, $\forall i = 1, \ldots, n$.

(c) Conclua com base em (a) e (b) deste exercício e (d) do exercício 14 que

$$\text{Cov}\left(\sum_{i=1}^n \alpha_i X_i, \sum_{j=1}^m \beta_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \text{Cov}(X_i, Y_j) \,,$$

com $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, n$ e $\beta_j \in \mathbb{R}$, $j = 1, \ldots, m$.

16. Mostre que para duas variáveis aleatórias quaisquer, com segundo momento finito, vale que $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$. Este resultado pode ser estendido para n variáveis aleatórias $X_1, X_2, \ldots, X_n$. Mostre que:

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j).$$

17. Considere novamente o exercício 1. (a) Calcule $\text{Var}(X+Y)$. (b) Calcule a variância da variável aleatória Z: número de bolas azuis escolhidas.

18. Em vários experimentos de diversas áreas de aplicação, é muito usado um esquema de amostragem denominado amostragem aleatória simples (A. A. S.) onde os valores a serem observados, $X_1, X_1, \ldots, X_n$, são considerados variáveis aleatórias independentes e identicamente distribuídas. Nestas situações é comum definir-se as seguintes quantidades: $\overline{X} = \dfrac{\sum_{i=1}^n X_i}{n}$: média amostral (como uma medida de posição para os dados); $X_i - \overline{X}$, $i = 1, 2, \ldots, n$: desvios (diferença entre os valores individuais observados e a média amostral). Estes desvios são utilizados para definir uma medida de dispersão dos dados; $S^2 = \dfrac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}$.

    Suponhamos que $\text{Var}(X_i) = \sigma^2$, $\forall i = 1, 2, \ldots, n$. Mostre que: (a) $\text{Cov}(X_i - \overline{X}, \overline{X}) = 0$ (a média amostral e os desvios são não-correlacionados). (b) $E[S^2] = \sigma^2$ (em inferência estatística, diz-se que S^2 é um estimador não-viciado para σ^2).

19. Sejam X e Y variáveis aleatórias com média zero, variância 1 e coeficiente de correlação ρ. Verifique que $X - \rho Y$ e Y são não-correlacionadas.

20. Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias. Muitas vezes, é interessante realizar um estudo destas variáveis conjuntamente, e não apenas de cada variável individualmente. Nestas situações, é comum utilizar-se o conceito de covariância para se verificar o tipo de associação entre pares de variáveis. Para isso, costuma-se definir a matriz de covariância Σ da variável aleatória
n-dimensional $(X_1, X_2, \ldots, X_n)$ da seguinte forma:

$$\sum_{n\times n} = \begin{pmatrix} \mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \ldots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \ldots & \mathrm{Cov}(X_n,X_n) \\ \vdots & & & \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \ldots & \mathrm{Cov}(X_n,X_n) \end{pmatrix}$$

(note que Σ é simétrica). Para $N = 4$, considere a matriz de covariância

$$\sum = \begin{pmatrix} 25 & -3 & 2 & 2 \\ -3 & 36 & 2 & 0 \\ 2 & 2 & 100 & 2 \\ 2 & 0 & 2 & 400 \end{pmatrix}$$

Determine: (a) Quais variáveis aleatórias (par de variáveis aleatórias) apresentam coeficiente de correlação negativo? (b) Quais variávies são não-correlacionadas? (c) Quais variáveis (par de variáveis) apresentam maior coeficiente de correlação? (d) Calcule $\rho(X_1, \frac{X_2+X_3}{2})$, isto é, o coeficiente de correlação de X_1 e $\frac{X_2+X_3}{2}$.

21. O melhor estimador linear de uma variável aleatória Y em relação a uma outra variável aleatória X pode ser definido como $\widehat{Y}(X) = a + bX$, com a e b tais que $E(\widehat{Y}(X) - Y)^2$ assuma valor mínimo. Determine a e b nestas condições.

22. Para qualquer evento aleatório A, pode-se definir a função indicadora de A da seguinte forma: $I_A = 1$ se A ocorre e $I_A = 0$ se A não ocorre. Mostre que $E(I_A) = P(A)$. Definindo I_B de maneira semelhante a I_A, mostre que $E(I_A I_B) = P(A \cap B)$ e que $\mathrm{Cov}(I_A, I_B) = P(A \cap B) - P(A)P(B)$.

23. Uma das distribuições de probabilidade conjuntas mais importantes é a multinomial, que aparece quando uma seqüência de n experimentos idênticos (realizados sob as mesmas condições) e independentes é realizada. Suponha que cada experimento pode resultar em uma de r realizações, com respectivas probabilidades $p_1, p_2, \ldots, p_r, \sum_{i=1}^{r} p_i = 1$. Se X_i denota o número de experimentos, dentre os n, que resultam na realização i, $i = 1, 2, \ldots, r$, onde $\sum_{i=1}^{r} n_i = n$, tem-se:

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r) = \frac{n!}{n_1! n_2! \cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}.$$

Sob o modelo multinomial, determine: (a) $E(X_i)$, $i = 1, \ldots, r$; (b) $\mathrm{Var}(X_i)$, $i = 1, 2, \ldots, r$; (c) $\mathrm{Cov}(X_i, X_j)$, $i \neq j$.

Dica: Defina os indicadores $I_i(k) = 1$, se o k-ésimo experimento resulta na realização i, $I_i(k) = 0$ caso contrário. Em seguida, expresse X_i como função de tais indicadores e aplique os resultados do exercício 22.

24. Considere novamente o exercício 1. Determine: (a) A distribuição condicional de Y dado $X = x$, $x = 0, 1, 2, 3$. (b) A esperança condicional de Y dado $X = x$, $x = 0, 1, 2, 3$. (c) A esperança de Y utilizando-se os valores obtidos no item (b).

25. Considere novamente o exercício 2, item (c). Obtenha a distribuição condicional de X dado $Y = y$, $y \in \{2, 3, \ldots, 12\}$.

26. Considere novamente o exercício 7. Determine: (a) A distribuição condicional de Y dado $X = x$, $x = 1, 2$. (b) A esperança condicional de Y dado $X = x$, $x = 1, 2$. (c) A esperança de Y.

27. No exercício 11, determine as distribuições condicionais de Y dado X e de X dado Y. Calcule, em seguida, $E(Y|X)$ e

$E(X|Y)$, ou seja, para cada x e cada y obtenha $E(Y|X=x)$ e $E(X|Y=y)$.

28. Um prisioneiro está preso em uma cela com três portas. A primeira leva a um túnel que retorna à cela após dois dias; a segunda leva a outro túnel que retorna à cela após quatro dias; a terceira leva a outro túnel que conduz para fora da cela após um dia. Se a cada tentativa o prisioneiro escolhe uma porta ao acaso e independentemente de eventuais tentativas anteriores, qual é o número esperado de dias até que o prisioneiro consiga sair da cela?

29. Suponha que o número de pessoas que entram em uma loja de departamentos em um determinado dia seja uma variável aleatória com média 50. Suponha ainda que a quantia gasta por cada um destes consumidores sejam variáveis aleatórias independentes com média comum R$ 50,00. Assuma também que a quantia gasta por um consumidor seja também independente do número de pessoas que entraram na loja. Qual é a quantia esperada de dinheiro gasta na loja em um determinado dia?

30. Se X e Y são independentes, mostre que $E[X|Y=y]=E[X]$, $\forall y$. Verifique ainda que se $E[X|Y=y]=E[X]$, $\forall y$, então X e Y são não-correlacionadas.

31. Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas. Determine $E[X_1|X_1+\cdots+X_n=x]$.

32. Mostre que:
(a) $E(XY|Y=y)=y.E(X|Y=y)$.
(b) $E[g(X,Y)|Y=y]=E[g(X,y)|Y=y]$.
(c) $E(XY)=E[Y.E(X|Y)]$.

33. Uma ilha tem uma floresta antiga com N seringueiras. Sabe-se que um pirata famoso escondeu seu tesouro dentro de M dessas árvores, porém não há indicação de quais árvores contêm ouro. Pensando em achar parte deste tesouro, um indivíduo derruba as árvores uma a uma, sendo que elas são selecionadas ao acaso, e verifica se há ouro dentro de cada uma delas.

(a) Seja X o número de árvores encontradas com ouro se n árvores são derrubadas. Usando funções indicadoras apropriadas, encontre $E(X)$.

(b) Seja Y o número de árvores derrubadas até encontrar r árvores com ouro, $r \leq M$ encontre $E(Y)$.

34. Uma moeda é lançada continuamente até se obter uma seqüência de R caras consecutivas. Sejam P a probabilidade de se obter cara em um lançamento da moeda e T_R o número de lançamentos da moeda necessários para se obter R caras consecutivas. (a) Determine $E[T_R|T_{R-1}]$. (b) Determine $E[T_R]$ em termos de $E[T_{R-1}]$. (c) Obtenha $E[T_1]$. (d) Obtenha $E[T_R]$.

35. Mostre que:
$$E(X)=E(X|X<a)P(X<a)+E(X|X\geq a)P(X\geq a).$$

36. Considere a seguinte definição:
$$\mathrm{Var}(X|Y)=E[(X-E(X|Y))^2|Y].$$
Mostre que $\mathrm{Var}(X)=E[\mathrm{Var}(X|Y)]+\mathrm{Var}[E(X|Y)]$.

37. Considere uma população que consiste de indivíduos com capacidade de auto-reprodução. Suponha que cada indivíduo, independentemente dos demais, seja capaz de produzir J outros indivíduos, para a próxima geração, com probabilidade P_J, $J=0,1,2,\ldots$, antes de morrer. Denote por X_n o número de indivíduos na N-ésima geração, com $X_0=1$. Sejam $\mu=\sum_{J=0}^{\infty}JP_J$ e $\sigma^2=\sum_{J=0}^{\infty}(J-\mu)^2P_J$ a média e a variância do número de filhos que cada indivíduo coloca na população. (a) Mostre que $E[X_n]=\mu E[X_{n-1}]$. (b) Conclua que $E[X_n]=\mu^n$. (c) Mostre que $\mathrm{Var}(X_n)=\sigma^2\mu^{n-1}+\mu^2\mathrm{Var}(X_{n-1})$. (d) Conclua que $\mathrm{Var}(X_n)=\sigma^2\mu^{n-1}\left(\frac{\mu^n-1}{\mu-1}\right)$, $\mu\neq 1$ e $\mathrm{Var}(X_n)=n\sigma^2$, $\mu=1$.

38. Um conceito importante em processos estocásticos e que apresenta características interessantes é o conceito de martingal. Considere uma seqüência $\{X_n, n\geq 1\}$ de variáveis aleatórias. Considere outra seqüência $\{Y_n, n\geq 1\}$ tal que $Y_n=f(X_1,X_2,$

$\ldots, X_n)$, $\forall n \geq 1$. Dizemos que $\{Y_n, n \geq 1\}$ é um martingal com respeito a $\{X_n, n \geq 1\}$ se:

(a) $E(|Y_n|) < \infty$. (b) $E[Y_{n+1}|X_1, X_2, \ldots, X_n] = Y_n, \forall n \geq 1$.

Mostre a seguinte propriedade dos martingais: $E[Y_n] = E[Y_{n-1}] = \cdots = E[Y_1]$, $\forall n \geq 1$.

39. Uma urna contém inicialmente b bolas brancas e p bolas pretas. Retira-se ao acaso desta urna uma bola, observando-se sua cor. Em seguida, repõe-se esta bola à urna juntamente com R bolas da mesma cor. Em seguida, retira-se novamente ao acaso uma bola da urna e observa-se a sua cor. Esta bola é reposta na urna juntamente com R bolas da mesma cor e, assim, sucessivamente. Considere a seqüência de variáveis $\{X_n, n \geq 1\}$, onde $X_n = 1$, se a n-ésima bola retirada da urna é branca, $X_n = 0$ caso contrário.

    Definindo $Y_n = Y_n(X_1, X_2, \ldots, X_n)$ como sendo a proporção de bolas brancas na urna após a n-ésima etapa do processo, mostre que $\{Y_n, n \geq 1\}$ é um martingal em relação a $\{X_n, n \geq 1\}$. Determine $E[Y_{20}]$. (Este modelo de urna é denominado modelo de urna de Polya. Este modelo foi sugerido por G. Polya para descrever, de maneira bem simplificada, a evolução de moléstias contagiosas em uma população e serviu como protótipo para o desenvolvimento de vários outros modelos de urnas.)

40. Um grande número, N, de pessoas é submetido a um exame de sangue. Este pode ser efetuado de duas maneiras: (i) cada pessoa pode ser testada individualmente; neste caso, são necessários N testes; (ii) as amostras de sangue de k pessoas podem ser misturadas e analisadas em conjunto. Se o teste é negativo, esse único teste é suficiente para as k pessoas. Se o teste é positivo, cada uma das k pessoas deve ser testada separadamente, e ao todo são necessários $k + 1$ testes para as k pessoas. Suponha que p, a probabilidade de que o teste seja positivo, seja a mesma para todas as pessoas e que estas sejam independentes. (a) Qual a probabilidade de que o teste para uma amostra misturada de k pessoas apresente positivo? (b) Qual o valor esperado do número de testes necessários sob o plano (ii)?

41. Designemos por P_x a probabilidade de que uma pessoa com x anos de idade sobreviva um ano. Além disso, designemos por ${}_tP_x$ a probabilidade de que sobreviva por t anos. Mostre que ${}_{t+\mu}P_x = ({}_tP_x)({}_\mu P_{x+t})$ e que quando t é um número inteiro N, temos:

$$ {}_tP_x = P_x \cdot P_{x+1} \cdots P_{x+n-1} \ . $$

    Seja L_0 uma constante positiva arbitrária, e definamos as funções $L_x = L_0 \cdot {}_xP_0$ e $dx = \ell_x - \ell_{x+1}$, para todo $x \geq 0$. Mostre que:

$$ {}_tP_x = \frac{\ell_{x+t}}{\ell_x} \quad \text{e} \quad q_x = \frac{dx}{\ell_x} \ . $$

    Interprete estes resultados. A probabilidade q_x é conhecida como a taxa de mortalidade à idade x.

42. Considere um grupo de L_x pessoas de idade x, que podem ser consideradas mutuamente independentes com respeito às suas possibilidades de vida e morte. Seja λ_{x+t} o número dessas pessoas no grupo que sobreviverão à idade $x + t$. Mostre que $E[\lambda_{x+t}] = L_{x+t}$.

43. Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas. Sejam:

$$
\begin{aligned}
X_{(1)} &= \min\{X_1, X_2, \ldots, X_n\} \\
&\quad \text{(menor valor entre } X_1, X_2, \ldots, X_n) \\
X_{(2)} &= \min\{\{X_1, X_2, \ldots, X_n\} - \{X_{(1)}\}\} \\
&\quad \text{(segundo menor valor entre } X_1, X_2, \ldots, X_n) \\
&\vdots \\
X_{(i)} &= \min\{\{X_1, X_2, \ldots, X_n\} - \{X_{(1)}, \ldots, X_{(i-1)}\}\} \\
&\quad (i\text{-ésimo menor valor entre } X_1, X_2, \ldots, X_n) \\
&\vdots \\
X_{(n)} &= \max\{X_1, X_2, \ldots, X_n\} \\
&\quad \text{(maior valor entre } X_1, X_2, \ldots, X_n) \ .
\end{aligned}
$$

Suponha que X_i tem função de distribuição F, $i = 1, \ldots, n$. Os valores ordenados $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ são denominados estatísticas de ordem das variáveis $X_1, \ldots, X_n$.

(a) Verifique que $F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$ e $F_{X_{(n)}}(x) = [F(x)]^n$ (veja exercício 12).

(b) Obtenha a função distribuição conjunta de $(X_{(1)}, X_{(n)})$, isto é, $P(X_{(1)} \leq a, X_{(n)} \leq b)$, $a, b \in \mathbb{R}$.

44. Considere k bolas numeradas de 1 a k distribuídas em duas urnas: urna I e urna II. Sorteia-se ao acaso um número em $\{1, 2, \ldots, k\}$ e transfere-se a bola cujo número foi sorteado para a outra urna. Por exemplo, se a bola cujo número escolhido está na urna I, ela é transferida para a urna II, e vice-versa. Repete-se este procedimento indefinidamente. Seja X_n a proporção de bolas, dentre as k bolas, contidas na urna I após a n-ésima repetição deste procedimento, $n \geq 1$. Seja X a proporção de bolas na urna I antes de se iniciar o processo.

(a) Determine $E[X_n]$ (expresse $E[X_n]$ em função de $E[X_{n-1}]$ via condicionamento);

(b) Verifique que $\lim_{N \to \infty} E[X_n] = \frac{1}{2}$, para toda configuração inicial das urnas.

(O modelo de urna apresentado aqui é denominado modelo de Ehrenfest. Este modelo descreve, de maneira bem simples, uma situação de troca de calor entre dois corpos isolados. Para isso, considere as urnas como corpos isolados e as bolas como moléculas em movimento dentro de cada corpo. Quanto maior o número de moléculas no corpo (bolas na urna), maior sua energia e, conseqüentemente, mais facilmente ele cederá calor ao outro corpo. É razoável esperar, a longo prazo, que os corpos atinjam o equilíbrio térmico, o que é traduzido pelo resultado em (b).)

4

Modelos Probabilísticos Discretos

Neste capítulo apresentaremos alguns dos principais modelos probabilísticos utilizados para descrever vários fenômenos ou situações que encontramos na natureza ou ainda experimentos por nós construídos. Esses modelos são expressos por uma família de distribuições de probabilidade que dependem de um ou mais parâmetros. Restringiremo-nos a considerar modelos em que as variáveis aleatórias que os expressam são discretas, ou seja, assumem um conjunto enumerável de valores, finito ou infinito.

4.1 DISTRIBUIÇÃO BINOMIAL

Vamos descrever as condições que permitem caracterizar esse modelo.

Consideremos um experimento em que estamos interessados num particular de seus eventos. Registramos em cada repetição do experimento a ocorrência ou não desse evento. A ocorrência do evento de interesse será denominada sucesso e sua não-ocorrência,

falha. Muitas vezes, num abuso de linguagem, o sucesso (S) será identificado com o evento de interesse. Listamos a seguir alguns experimentos e consideramos em cada um deles um evento de interesse.

EXEMPLO 4.1.1 Suspeita-se que uma moléstia contagiosa possa ter atingido os moradores de uma região. Cada morador da região é classificado conforme tenha sido contagiado ou não. O evento de interesse é: a pessoa examinada foi contagiada.

EXEMPLO 4.1.2 Consideremos o conjunto dos televisores produzidos por uma indústria eletrônica. Cada televisor ao sair da linha de produção é testado e verifica-se se ele apresenta defeito ou não. O evento de interesse é: o televisor examinado apresenta defeito.

EXEMPLO 4.1.3 Ao fazer-se um estudo sobre o grau de escolaridade de jovens de 18 a 25 anos, cada jovem nesta faixa etária é classificado conforme tenha concluído ou não o segundo grau. O evento de interesse será "ter concluído o segundo grau".

EXEMPLO 4.1.4 Em uma prova de múltipla escolha cada questão tem cinco alternativas das quais somente uma é correta. Para cada questão, escolher a alternativa correta corresponde a sucesso e uma das erradas, a falha.

Como o leitor deve ter notado, em todas as situações descritas cada elemento da população é classificado segundo possua ou não uma dada característica.

Para construir o modelo binomial vamos introduzir uma seqüência de ensaios que são denominados ensaios de Bernoulli.

Uma seqüência de ensaios de Bernoulli é definida por meio das três condições seguintes:

i) Em cada ensaio considera-se somente a ocorrência ou não-ocorrência de um certo evento que será denominado sucesso (S) e cuja não ocorrência será denominada falha (F).

ii) Os ensaios são independentes.

iii) A probabilidade de sucesso, que denotaremos por p, é a mesma em cada ensaio. A probabilidade de falha será denotada por $1-p$.

Para um experimento que consiste na realização de n ensaios de Bernoulli, o espaço amostral pode ser considerado como o conjunto das n-uplas de comprimento n, em que em cada posição há um sucesso (S) ou uma falha (F).

Pelas condições ii) e iii) vemos que a probabilidade de um ponto amostral com sucessos nos k primeiros ensaios e falhas nos $n-k$ ensaios seguintes é $p^k(1-p)^{n-k}$. Note que esta é a probabilidade de qualquer ponto com k sucessos e $n-k$ falhas em posições fixadas.

Notemos que o evento $\{X = k\}$ ocorre se observarmos um ponto amostral que tenha k sucessos e $n-k$ falhas. O número de pontos do espaço amostral que satisfaz essa condição é igual ao número de maneiras com que podemos escolher k ensaios dentre os n para a ocorrência de sucesso, pois nos $n-k$ restantes deverão ocorrer falhas. Este número é igual ao número de combinações de n elementos tomados k a k, ou seja $\binom{n}{k}$.

Decorre do que foi exposto que, para $k = 0, 1, ..., n$:

$$P[X = k] = \binom{n}{k} p^k (1-p)^{n-k} . \tag{4.1}$$

A fórmula (4.1) é denominada distribuição binomial com parâmetros n e p, onde n é o número de ensaios e p a probabilidade de sucesso em cada ensaio.

O número de sucessos X em n ensaios de Bernoulli pode ser representado por meio de variáveis aleatórias associadas a cada ensaio, que assumem valores zero ou 1.

Seja $X_i = 1$ se ocorre sucesso no i-ésimo ensaio e $X_i = 0$ se ocorre falha, para $i = 1, 2, \ldots, n$. Então X pode ser expresso da seguinte maneira:

$$X = X_1 + X_2 + \cdots + X_n .$$

As variáveis aleatórias que assumem valor 1 ou zero conforme um dado evento ocorra ou não são denominadas variáveis indicadoras ou indicadores.

A distribuição binomial pode ser representada por um modelo de urna. Consideremos uma urna com N bolas das quais M são brancas e $N-M$ são pretas. Retiram-se sucessivamente n bolas da

urna recolocando-se após cada retirada a bola selecionada. Vamos considerar sucesso a ocorrência de branca em uma retirada e falha a ocorrência de preta. Temos uma seqüência de n ensaios de Bernoulli com probabilidade de sucesso igual a $\frac{M}{N}$ e de falha $(1-\frac{M}{N})$ Desta maneira temos pela fórmula (4.1), para $k = 0, 1, ..n$:

$$P[X=k] = \binom{n}{k}(\frac{M}{N})^k(1-\frac{M}{N})^{n-k}\,.$$

Observemos que os modelos de urnas onde as bolas são repostas após cada retirada representam esquemas de amostragem com reposição.

EXEMPLO 4.1.5 Considere um conjunto de n mulheres grávidas internadas em uma maternidade para dar à luz. O número de filhos do sexo feminino é uma variável aleatória com distribuição binomial, pois, cada nascimento pode ser considerado independente dos demais e a probabilidade de um nascimento do sexo feminino é aproximadamente $\frac{1}{2}$. Para $n = 6$, denotando-se por X o número de filhos do sexo feminino, calculemos $P[X = 2]$.

Pela fórmula (4.1) tem-se:

$$P[X=2] = \binom{6}{2}(\frac{1}{2})^2(\frac{1}{2})^4 = 0,23\,.$$

EXEMPLO 4.1.6 Na linha de produção de uma fábrica, em condições normais de funcionamento cada uma das peças pode ser considerada como produzida independentemente das demais. Se retirarmos uma amostra de n peças da linha de produção e se chamarmos de p a fração de peças defeituosas que são produzidas, então X, o número de peças defeituosas na amostra, é uma variável aleatória com distribuição binomial com parâmetros n e p. Para $n = 5$ e $p = 0,1$ calculemos as probabilidades que haja: a) uma peça defeituosa; b) duas peças defeituosas; c) pelo menos uma peça defeituosa.

Nas condições mencionadas no exemplo temos uma seqüência de ensaios de Bernoulli com $n = 5$ e $p = 0,1$. Assim usando-se a fórmula (4.1) temos:

$$P[X=1] = \binom{5}{1}(0,1)(0,9)^4 = 0,328$$

$$P[X=2] = \binom{5}{2}(0,1)^2(0,9)^3 = 0,365$$

$$P[X\geq 1] = 1 - P[X=0] = 1-(0,9)^5 = 0,4095.$$

Vamos determinar a esperança de X, ou o que é equivalente, da distribuição binomial com parâmetros n e p. A partir da definição dada pela expressão (2.9)

$$\begin{aligned}
E(X) &= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} = \\
&= \sum_{k=0}^{n}\frac{kn!}{k!(n-k)!}p^k(1-p)^{n-k} = \\
&= \sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}p^k(1-p)^{n-k} = \\
&= np\sum_{j=0}^{n-1}\frac{(n-1)!}{j!(n-j-1)!}p^j(1-p)^{n-j-1} = \\
&= np\sum_{j=o}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-j-1} = \\
&= np(p+(1-p))^{n-1} = np\,. \qquad (4.2)
\end{aligned}$$

Na terceira igualdade pusemos $j = k-1$ e a penúltima igualdade é conseqüência da expansão do binômio $(p+(1-p))$ à potência $n-1$.

Vamos determinar a função geradora de momentos da distribuição binomial. A partir desta, obteremos novamente a esperança, que é o primeiro momento da distribuição, e a variância, que é o segundo momento em torno da média.

$$\begin{aligned}
\phi(t) &= E(e^{tX}) = \sum_{k=0}^{n} e^{tk}\binom{n}{k}p^k(1-p)^{n-k} = \\
&= \sum_{k=0}^{n}\binom{n}{k}(pe^t)^k(1-p)^{n-k} = (pe^t+(1-p))^n, \quad (4.3)
\end{aligned}$$

De acordo com a fórmula (2.25) temos :

$$E(X) = \frac{d}{dt}\phi(0).$$

Derivando-se $\phi(t)$ tem-se :

$$\frac{d}{dt}\phi(t) = n(pe^t + (1-p))^{n-1}pe^t.$$

Pondo $t = 0$ vem :

$$E(X) = np. \tag{4.4}$$

Vamos agora calcular o segundo momento de X, isto é, $E(X^2)$ a partir das expressões (2.32) e (4.3).

$$\frac{d^2}{dt^2}\phi(t) = n(n-1)(pe^t + (1-p))^{n-2}p^2(e^t)^2 + n(pe^t + (1-p))^{n-1}pe^t$$

Pondo $t = 0$ obtemos:

$$\frac{d^2}{dt^2}\phi(0) = E(X^2) = n(n-1)p^2 + np.$$

Substituindo os valores de $E(X^2)$ e $(E(X))^2$ em $\mathrm{Var}(X) = E(X^2) - (E(X))^2$
temos $\mathrm{Var}(X) = n(n-1)p^2 + np - n^2p^2$, simplificando vem:

$$\mathrm{Var}(X) = np(1-p). \tag{4.5}$$

4.2 DISTRIBUIÇÃO HIPERGEOMÉTRICA

Essa distribuição representa um modelo para amostragem sem reposição de uma população com um número finito de elementos, em que cada elemento pode ser de um de dois tipos. Se a população tem N elementos, M de um tipo e $N - M$ do outro, então podemos representa-lá por meio de uma urna que contenha, digamos, M bolas brancas e $N - M$ bolas vermelhas. Retira-se da urna n bolas sem reposição, isto é, após cada retirada a bola selecionada não é reposta na urna. Vamos designar por X o número de bolas brancas entre as n bolas retiradas da urna. Desejamos determinar a distribuição de probabilidade da variável aleatória X.

Vamos mostrar que :

$$P[X = k] = \binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n} \tag{4.6}$$

onde $max\{0, n-(N-M)\} \leq k \leq min\{M, n\}$.

Para justificar esses limites, notemos que o número de bolas brancas na amostra k é menor ou igual ao número de bolas brancas na urna M e também menor ou igual ao número de bolas na amostra n, portanto menor ou igual ao menor deles. Isso justifica o limite superior. Se o tamanho da amostra n é menor ou igual ao número de bolas vermelhas $N-M$, então na amostra todas podem ser vermelhas e portanto k=0. Se $n \geq (N-M)$, então mesmo que todas as $(N-M)$ vermelhas pertençam à amostra, haverá $n - (N - M)$ brancas na amostra.

O espaço amostral para esse experimento é formado pelo conjunto das amostras não ordenadas de n bolas retiradas das N, ou o que é o mesmo, pelo conjunto das combinações de N elementos tomados n a n, cujo número é igual a $\binom{N}{n}$.

Existem $\binom{M}{k}$ combinações de k bolas brancas retiradas das M e $\binom{N-M}{n-k}$ combinações de $n-k$ vermelhas retiradas das $N-M$. Assim o número de combinações com k brancas e $n-k$ vermelhas é o produto $\binom{M}{k}\binom{N-M}{n-k}$. Aplicando a definição de probabilidade obtemos a fórmula (4.6).

EXEMPLO 4.2.1 Uma caixa contém 15 peças das quais quatro estão com defeito. Retira-se uma amostra de três peças da caixa sem reposição. Calcular a probabilidade que haja duas ou três peças defeituosas na amostra.

A distribuição de probabilidade do número de peças defeituosas na amostra é hipergeométrica com $N = 15$, $M = 4$ e $n = 3$. Se designarmos por X o número de peças defeituosas na amostra, desejamos calcular $P[X = 2] + P[X = 3]$.

$$P[X = 2] + P[X = 3] = \frac{\binom{4}{2}\binom{11}{1}}{\binom{15}{3}} + \frac{\binom{4}{3}\binom{11}{0}}{\binom{15}{3}} = 0,153\,.$$

Vamos, a título de ilustração, calcular a esperança da distribuição hipergeométrica diretamente a partir da definição, supondo por comodidade que $0 \leq k \leq n$.

$$E(X) = \sum_{k=0}^{n} k\binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n} =$$

$$= n.\frac{M}{N}.\sum_{k=1}^{n}\binom{M-1}{k-1}\binom{N-M}{n-k}\Big/\binom{N-1}{n-1} =$$

$$= n.\frac{M}{N}.\sum_{j=0}^{n-1}\binom{M-1}{j}\binom{N-1-M+1}{n-1-j}\Big/\binom{N-1}{n-1} .$$

Simplificando: usando-se a relação $\sum_{i=0}^{m}\binom{a}{i}\binom{b}{m-i} = \binom{a+b}{m}$ ou o que é equivalente, o fato de que a soma acima é a soma das probabilidades de uma distribuição hipergeométrica com parâmetros $N-1$, $M-1$ e $n-1$, vem:

$$E(X) = n.\frac{M}{N} . \tag{4.7}$$

Sugerimos que o leitor mostre, ver exercício 19, que a variância da distribuição hipergeométrica é dada por :

$$\mathrm{Var}(X) = n\frac{M}{N}\frac{N-M}{N}(1-\frac{n-1}{N-1}) . \tag{4.8}$$

EXEMPLO 4.2.2 Foram impressos vinte exemplares de um boletim. Em quatro exemplares falta uma página. Você escolhe cinco exemplares ao acaso entre os vinte. Qual é a probabilidade que haja dois exemplares com a página faltando entre os cinco? Qual é a probabilidade que pelo menos um tenha a página faltando? Qual a média e a variância do número de exemplares com a página faltando entre os cinco ?

Seja X a variável aleatória que dá o número de exemplares com a página faltando na amostra de 5. X tem distribuição hipergeométrica com N=20, M =4 e n=5. Desta maneira pela expressão (4.6) tem-se:

$$P[X=2] = \binom{4}{2}\binom{16}{3}\Big/\binom{20}{5} = 0,216 .$$

O evento pelo menos um boletim tem página faltando ocorre se X=1 ou X=2 ou X=3 ou X=4, logo sua probabilidade é igual a $P[X=1]+P[X=2]+P[X=3]+P[X=4]$. É mais fácil calcular a probabilidade do evento complementar, isto é, $1-P[X=0]$. Pela fórmula (4.6) tem-se:

$$P[X=0] = \binom{16}{5}\Big/\binom{20}{5} = \frac{91}{323} .$$

Logo a probabilidade procurada é $1-P[X=0] = \frac{232}{323}$.

Da expressão (4.7) vem: $E(X) = n\frac{M}{N} = 1$.

Para a variância pela expressão (4.8) tem-se:

$$\mathrm{Var}(X) = n(\frac{M}{N})(\frac{N-M}{N})(1-\frac{n-1}{N-1}) = \frac{12}{19} .$$

4.3 DISTRIBUIÇÃO GEOMÉTRICA

Consideremos uma seqüência ilimitada de ensaios de Bernoulli, com probabilidade de sucesso p em cada ensaio. Designemos sucesso por S e falha por F. Realizamos os ensaios até que ocorra o primeiro sucesso.

O espaço amostral para esse experimento é o conjunto:

$$\{S, FS, FFS, ..., FF, ..., FS, ...\} .$$

Um elemento típico desse espaço amostral é uma seqüência de comprimento n em que nas primeiras $n-1$ posições temos F e na n-ésima temos S.

Seja X a variável aleatória que dá o número de falhas que precedem o primeiro sucesso.

A distribuição de probabilidade de X para $j = 0, 1, ...$ é dada por:

$$P[X=j] = (1-p)^j p . \tag{4.9}$$

O evento $[X=j]$ ocorre se e somente se ocorrem somente falhas nos j primeiros ensaios e sucesso no $(j+1)$-ésimo ensaio. A expressão (4.9) segue da independência dos ensaios.

Vamos calcular $E(X)$ a partir da definição.

No cálculo de $E(X)$, utilizaremos uma expressão que vale a pena destacar, pois é de interesse geral.

Para todo número real x no intervalo (0,1) consideremos a série geométrica cuja soma é dada a seguir:

$$\sum_{i=1}^{\infty} x^i = \frac{x}{1-x} . \tag{4.10}$$

Derivando-se ambos os membros dessa igualdade, temos:

$$\frac{d}{dx}\sum_{i=1}^{\infty} x^i = \sum_{i=1}^{\infty} ix^{(i-1)} = \frac{1}{(1-x)^2} . \tag{4.11}$$

Usando-se a definição de esperança dada pela expressão (2.9) tem-se:

$$\begin{aligned} E(X) &= \sum_{j=0}^{\infty} j(1-p)^j p = p\sum_{j=1}^{\infty} j(1-p)^j = \\ &= p(1-p)\sum_{j=1}^{\infty} j(1-p)^{j-1} = \frac{p(1-p)}{p^2} . \end{aligned}$$

Observe que utilizamos (4.11) com $x = 1-p$ para obter a última igualdade acima. Simplificando vem:

$$E(X) = \frac{1-p}{p} . \tag{4.12}$$

Usando a expressão (2.32) podemos calcular $E(X^2)$ e obter a variância de X usando (2.23). Sugerimos ao leitor que faça esse cálculo que fornecerá :

$$var(X) = \frac{1-p}{p^2} . \tag{4.13}$$

EXEMPLO 4.3.1 Um banco de sangue necessita sangue do tipo O-Rh negativo. Seja p a proporção de indivíduos na população com esse tipo de sangue. Suponha que as pessoas são escolhidas da população ao acaso para serem examinadas se tem esse tipo de sangue. O número de pessoas examinadas antes de se encontrar a primeira pessoa com esse tipo de sangue tem distribuição geométrica com parâmetro p. Calcule para $p = 0,1$ a probabilidade de a primeira pessoa a ser encontrada com esse tipo de sangue ser a quinta. Se

X designa o número de pessoas examinadas antes de se encontrar a primeira com tipo de sangue O-Rh negativo, queremos calcular $P[X = 4]$.

$$P[X = 4] = (1-p)^4 p = (0,9)^4 0,1 = 0,0656.$$

A distribuição geométrica tem uma propriedade que serve para caracterizá-la no conjunto das distribuições discretas, que é expressa no seguinte lema.

LEMA 4.3.1 *Se X é variável aleatória discreta com distribuição geométrica dada por (4.9) então, para todo $j, k = 1, 2, ...$ tem-se:*

$$P[X \geq j+k | X \geq j] = P[X \geq k] . \tag{4.14}$$

DEMONSTRAÇÃO

$$\begin{aligned} P[X \geq j+k | X \geq j] &= \frac{P[X \geq j+k, X \geq j]}{P[X \geq j]} = \frac{P[X \geq j+k]}{P[X \geq j]} = \\ &= \frac{\sum_{i=j+k}^{\infty}(1-p)^i p}{\sum_{i=j}^{\infty}(1-p)^i p} = \frac{\frac{(1-p)^{j+k}}{p}}{\frac{(1-p)^j}{p}} = \\ &= (1-p)^k = P[X \geq k] . \end{aligned}$$

Esta propriedade reflete a falta de memória ou de desgaste da distribuição geométrica. Considere um objeto cujo tempo de vida é uma variável aleatória X com distribuição geométrica com parâmetro p. Suponha que esse tempo de vida é medido em unidades discretas, isto é, ao final de cada unidade de tempo verifica-se se o objeto está funcionando ou se deixou de funcionar. A expressão (4.14) mostra que, se o objeto funcionou até o instante j, então a probabilidade de que ele funcione por mais k unidades de tempo é a mesma que a probabilidade de que um objeto novo funcione por k unidades de tempo.

Para uma seqüência ilimitada de ensaios de Bernoulli é comum definir-se a variável aleatória Y que é igual ao número de ensaios até

a ocorrência do primeiro sucesso. A relação entre Y e X é portanto: $Y = X + 1$.

4.4 DISTRIBUIÇÃO DE POISSON

A distribuição de Poisson que estudaremos nesta seção tem grande importância tanto teórica como aplicada. Optamos por introduzir essa distribuição como uma aproximação para a distribuição binomial quando o número de ensaios de Bernoulli tende a infinito, porém de modo que o produto $np_n = \lambda$ permanece constante. Nós de fato consideramos para cada inteiro n um conjunto de n ensaios de Bernoulli em que a probabilidade de sucesso em cada um dos n ensaios é p_n.

Associado aos n ensaios de Bernoulli, definimos a variável aleatória X_n que é igual ao número de sucessos nesses n ensaios. Então X_n tem distribuição binomial com parâmetros n e p_n.

$$P[X_n = k] = \binom{n}{k} {p_n}^k (1-p_n)^{n-k} =$$

$$= \frac{n(n-1)..(n-(k-1))}{k!} (\frac{\lambda}{n})^k (1-\frac{\lambda}{n})^n (1-(\frac{\lambda}{n})^{-k} =$$

$$= 1.(1-\frac{1}{n})...(1-\frac{k-1}{n})\frac{1}{k!}.(\lambda)^k (1-\frac{\lambda}{n})^n (1-\frac{\lambda}{n})^{-k} .$$

O primeiro termo do produto é 1 e os $k-1$ seguintes são iguais a $1-\frac{1}{n}, 1-\frac{2}{n}, \ldots, 1-\frac{k-1}{n}$, respectivamente, que tendem a 1 quando $n \to \infty$. O termo $(1-\frac{\lambda}{n})^{-k}$ também tende a 1.

Levando-se em conta que no cálculo mostra-se que:

$$\lim_{n\to\infty} (1-\frac{\lambda}{n})^n = e^{-\lambda}.$$

Passando-se ao limite na expressão acima para n tendendo a infinito obtemos:

$$\lim_{n\to\infty} P[X_n = k] = \frac{e^{-\lambda}\lambda^k}{k!} .$$

Observemos que decorre de:

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \qquad \text{que} \qquad \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1 .$$

DEFINIÇÃO 4.4.1 *A variável aleatória X tem distribuição de Poisson com parâmetro λ se para $k = 0, 1, 2...$:*

$$P[X = k] = \frac{e^{-\lambda}\lambda^k}{k!} . \tag{4.15}$$

Temos na realidade não uma distribuição de Poisson, mas sim uma família de distribuições de Poisson dependendo do parâmetro λ.

Seja X uma variável aleatória com distribuição de Poisson com parâmetro λ. Vamos mostrar que:

$$E(X) = \lambda \tag{4.16}$$

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} \\ &= \lambda.e^{-\lambda}.\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda , \quad \text{onde } j = k-1. \end{aligned}$$

De maneira análoga, podemos calcular $E(X^2)$ e obter a variância de X.

$$\text{Var}(X) = E(X^2) - E^2(X) = \lambda. \tag{4.17}$$

Outro caminho seria determinar a função geradora de momentos de X e então obter, a partir desta, a média e a variância de X. Obteremos a função geradora $\phi(t)$ para X e deixaremos a cargo do leitor calcular a média e variância.

$$\phi(t) = E(e^{tX}) = \sum_{k=0}^{\infty} e^{tk}\frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=0}^{\infty}\frac{e^{-\lambda}(\lambda e^t)^k}{k!} = e^{\lambda(e^t-1)}. \tag{4.18}$$

4.5 EXERCÍCIOS

1. Sabe-se que os parafusos produzidos por uma certa companhia são defeituosos com probabilidade $0,01$, independentemente uns dos outros (isto é, a fração não-conforme de parafusos na produção é $0,01$). A companhia vende os parafusos em pacotes de dez unidades e oferece uma garantia de devolução do dinheiro caso existam dois ou mais parafusos defeituosos no pacote com dez parafusos. (a) Qual a proporção de pacotes vendidos para os quais a companhia deve efetuar devolução de dinheiro? (b) Supondo que o número de parafusos defeituosos num determinado pacote é independente dos demais pacotes, qual a probabilidade de que uma pessoa que compra dez pacotes de parafusos tenha que retornar à companhia para devolução do dinheiro?

2. Suponha que o número de erros tipográficos em uma única página de um livro tem distribuição de Poisson com parâmetro $\lambda = \frac{1}{2}$. (a) Calcule a probabilidade de existir exatamente dois erros tipográficos em uma página. (b) Calcule a probabilidade de que exista pelo menos um erro em uma página. (c) Suponha agora que o livro em questão possui 200 páginas. Qual a probabilidade de não existir erros tipográficos neste livro?

3. Uma moeda, cuja probabilidade de cara em um lançamento é p, $0 < p < 1$, é lançada indefinidamente até que ocorra uma seqüência de cinco caras (seguidas). Um estudante resolve modelar esse experimento definindo a variável aleatória X igual ao número de lançamentos da moeda até que ocorra a seqüência de cinco caras. Ele então raciocina do seguinte modo: para $k = 5$, $P(X = k) = p^5$, pois devemos ter cara nos cinco primeiros lançamentos. Para $k = 6$, $P(X = k) = (1-p)p^5$, pois devemos ter coroa no primeiro lançamento e cara nos cinco lançamentos seguintes.

   Assim, concluiu que

$$P(X = k) = \begin{cases} 0 & , k < 5, \\ (1-p)^{k-5}p^5 & , k \geq 5, \end{cases}$$

e, portanto, $X^* = X - 5$ tem distribuição geométrica de parâmetro p^5. Discuta a proposta do estudante abordando as características do modelo proposto e a validade da expressão $P(X = k)$.

4. Suponha que uma particular característica de uma pessoa (por exemplo, cor dos olhos) é determinada por um par de genes. Suponha ainda que D representa o gene dominante e R o gene recessivo, de tal forma que uma pessoa com genes DD é dominante pura, uma com RR é recessiva pura e uma com RD ou DR é híbrida, e que uma pessoa dominante pura e uma pessoa híbrida têm aparências semelhantes quanto a esta característica. Se ambos os membros de um casal são híbridos quanto a uma particular característica e se o casal possui quatro filhos, qual a probabilidade de que três das quatro crianças tenham fenótipo determinado pelo gene dominante? (Lembre-se que cada criança recebe um gene de cada um dos pais.)

5. Suponha que a probabilidade de que um item produzido por uma máquina seja defeituoso é $0,1$. Determine a probabilidade de que uma amostra de dez itens conterá no máximo um item defeituoso. Compare os resultados obtidos pelas distribuições binomial e Poisson.

6. Com a finalidade de aumentar a arrecadação do seu sistema de loterias, a Caixa Econômica Federal implantou um novo jogo denominado Loto II, no qual o apostador escolhe seis dezenas do conjunto $\{01, 02, \ldots, 50\}$. Toda semana, a Caixa sorteia seis dezenas desse mesmo conjunto e atribui prêmios aos acertadores da: (a) Sena – as seis dezenas sorteadas, (b) Quina – cinco das dezenas sorteadas, (c) Quadra – quatro das dezenas sorteadas. Determine a probabilidade de que uma pessoa que aposta na Loto II ganhe algum dos prêmios oferecidos.

7. Um apostador da Loto II (descrita no exercício 6) é acusado de conhecer previamente o resultado das dezenas sorteadas, uma vez que ganhou doze prêmios da Loto II em um único ano, fato considerado extremamente raro. Construa um modelo

probabilístico adequado ao problema e verifique a validade de tal acusação.

8. Considere um experimento que consiste em contar o número de partículas alfa emitidas, num intervalo de tempo de um segundo, por um grama de material radioativo. Sabe-se por experiência passada que, em média, 3,2 de tais partículas são emitidas por segundo. Determine a probabilidade (aproximada) de que não mais que duas partículas alfa sejam emitidas.

9. Um comprador adquire componentes eletrônicos em lotes de tamanho dez. Este comprador adota a seguinte postura: inspeciona três componentes aleatoriamente de um lote e o aceita apenas se todos os três componentes não forem defeituosos. Se 30% dos lotes têm quatro componentes defeituosos e 70% têm apenas um, qual a proporção de lotes que o comprador rejeita?

10. Dois amigos, Paulo e Pedro, resolvem ir juntos a uma casa lotérica para apostar na Loto II. Paulo apostou nas dez últimas semanas, ao passo que Pedro nunca apostou na loteria em questão. Paulo argumenta que Pedro deve apostar nas mesmas dezenas que ele, pois julga que o fato de já ter participado do jogo algumas vezes e nunca ter ganho prêmio algum torna mais eminente sua vitória no próximo sorteio, relativamente a Pedro. Em outras palavras, Paulo argumenta que o fato de já ter apostado dez vezes na Loto II e nunca ter ganho prêmio algum aumenta a probabilidade de obter algum prêmio no próximo sorteio. Discuta a afirmação feita por Paulo propondo um modelo probabilístico para a situação descrita.

11. Considere um júri constituído por dez jurados que adota a seguinte regra para determinar o veredito: se pelo menos oito dos jurados julgarem o réu culpado, este será condenado, caso contrário, o réu não será condenado. Suponha que cada jurado dá seu veredito independentemente dos demais e que cada um toma a decisão correta com probabilidade θ. Suponha ainda que α representa a probabilidade do réu ser culpado. Determine a probabilidade do júri tomar uma decisão correta.

12. Um jornaleiro compra jornais a dez centavos e os vende a quinze centavos. Contudo, ele não pode devolver os jornais não vendidos. Se a demanda diária de jornais é uma variável aleatória binomial com $n = 10$ e $p = \frac{1}{3}$, quantos jornais, aproximadamente, ele deve comprar de maneira a maximizar seu lucro esperado?

13. Suponha que um motor de um avião falha, em vôo, com probabilidade $1 - p$, independentemente dos demais motores. Se um avião precisa de que a maioria dos seus motores funcione para realizar um vôo bem-sucedido, para que valores de p um avião com cinco motores é preferível a um avião com três motores?

14. Um sistema de satélites consiste de N componentes e funciona em qualquer dia se pelo menos k dos N componentes funcionarem nesse dia. Em dias chuvosos, cada componente funciona independentemente com probabilidade p_1, ao passo que em dias não chuvosos os componentes funcionam independentemente com probabilidade p_2. Se a probabilidade de chover amanhã é α, qual é a probabilidade de que o sistema de satélites irá funcionar amanhã?

15. Mostre que $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$.

    Dica: defina uma variável aleatória com distribuição hipergeométrica de parâmetros $2n$, n, n.

16. Para uma variável aleatória N assumindo valores inteiros não-negativos, mostre que $E(N) = \sum_{i=1}^{\infty} P\{N \geq i\}$.

17. Seja N uma variável aleatória que assume valores inteiros não-negativos. Mostre que $\sum_{i=0}^{\infty} iP\{N > i\} = \frac{1}{2}(E[N^2] - E[N])$.

18. Seja X uma variável aleatória com valor esperado μ e variância σ^2. Seja $Y = \frac{X-\mu}{\sigma}$. Determine: (a) $E[Y]$. (b) $\mathrm{Var}[Y]$.

    A variável Y assim obtida é denominada variável padronizada (neste caso, Y é obtida padronizando-se X) e é de grande importância em várias técnicas de análise estatística.

19. Mostre que para uma variável aleatória X com distribuição hipergeométrica de parâmetros N, m, n, a variância é dada pela fórmula (4.8).

20. Mostre que:

$$\binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n} \to \binom{n}{k}p^k(1-p)^{n-k},$$

quando $N, M \to \infty$ tais que $\frac{M}{N} \to p$. Interprete o resultado.

21. Seja X uma variável aleatória Poisson com parâmetro λ. Mostre que $P\{X = k\}$ atinge o valor máximo para $k = [\lambda]$, onde $[\lambda]$ denota o maior inteiro menor ou igual a λ.

22. Seja X uma variável aleatória binomial com parâmetros n e p. Mostre que $P\{X = k\}$ atinge o valor máximo para $k = [(n+1)p]$.

23. Seja X uma variável aleatória com distribuição de Poisson com parâmetro λ. Determine $P(X \in A)$, onde $A = \{0, 2, 4, 6, 8, \ldots\}$.

24. Considere X uma variável aleatória Poisson com parâmetro λ. (a) Mostre que $E[X^n] = \lambda E[(X+1)^{n-1}]$. (b) Calcule $E[X^4]$. (c) Determine $E[X!]$, para $0 < \lambda < 1$. (d) Determine $E[\cos \pi X]$ e $\mathrm{Var}[\cos \pi X]$.

25. Suponha que o número de eventos que ocorrem em um intervalo de tempo especificado é uma variável aleatória de Poisson com parâmetro λ. Se cada evento é contado com probabilidade p, independentemente de qualquer outro evento, mostre que o número de eventos que são contados é uma variável aleatória Poisson com parâmetro λp.

26. Para uma variável aleatória com distribuição hipergeométrica com parâmetros N, m e n, determine: $\frac{P\{X=k+1\}}{P\{X=k\}}$.

27. Se X e Y são variáveis aleatórias de Poisson independentes com parâmetros λ_1 e λ_2, respectivamente, calcule a distribuição de $X + Y$.

28. Sejam X e Y variáveis aleatórias binomiais independentes com parâmetros (n, p) e (m, p), respectivamente. Determine a distribuição de $X + Y$.

29. Considere novamente as variáveis do exercício 27. Calcule a distribuição condicional de X dado que $X + Y = n$.

30. Considere novamente as variáveis X e Y do exercício 28, com a restrição $n = m$, isto é, X e Y têm a mesma distribuição binomial (n, p). Determine a distribuição condicional de X dado que $X + Y = m$.

31. Suponha que X e Y são variáveis aleatórias geométricas independentes com mesmo parâmetro p. Calcule

$$P\{X = i | X + Y = n\}.$$

32. Suponha que ensaios independentes, cada um tendo probabilidade p de sucesso, $0 < p < 1$, são realizados até que um total de R sucessos seja acumulado. Seja X o número de ensaios necessários para se obter o total de R sucessos. Determine a distribuição de X (dizemos, neste caso, que X possui distribuição binomial negativa com parâmetros R e p).

33. Considere a variável aleatória X do exercício anterior. Determine $E(X)$ e $\mathrm{Var}(X)$.

Dica: Observe que $X = Y_1 + Y_2 + \cdots + Y_R$, onde Y_i tem distribuição geométrica de parâmetro p, $1 \leq i \leq R$.

34. Se ensaios independentes, cada um deles resultando em sucesso com probabilidade p, são realizados indefinidamente, qual a probabilidade de que R sucessos ocorram antes de M fracassos?

35. Suponha que existam N diferentes figurinhas e que um colecionador compre sucessivamente envelopes fechados que contêm apenas uma delas. Admita que a probabilidade de um envelope conter uma dada figurinha é $1/N$ independentemente uns dos outros. Seja X o número de envelopes que o colecionador deve comprar até conseguir completar sua coleção. (a) X tem distribuição binomial negativa? (b) Determine $E(X)$.

36. Sejam X, Y e Z variáveis aleatórias independentes com a mesma distribuição geométrica de parâmetro p. Determine:

    (a) $P(X = Y)$. (b) $P(X \geq 2Y)$. (c) $P(X + Y \leq Z)$.

37. Sejam X uma variável aleatória binomial negativa com parâmetros R e p, e Y uma variável aleatória binomial com parâmetros n e p. Mostre que $P\{X > n\} = P\{Y < r\}$.

38. Mostre que, se a variável aleatória X tem distribuição de Poisson com parâmetro λ, sua variância é dada pela expressão (4.17).

39. Seja X uma variável aleatória binomial com parâmetros n e p. Qual o valor de p que maximiza $P\{X = k\}$, $k = 0, 1, 2, \ldots, n$? Este é um exemplo de um método estatístico usado para estimar p quando a observação de uma variável aleatória binomial (n, p) é igual a k. Se o valor de n é conhecido e o de p desconhecido, então estimamos p escolhendo o valor de p que maximiza $P\{X = k\}$. Este método é conhecido como estimação de máxima verossimilhança.

40. Seja X uma variável aleatória Poisson com parâmetro λ. Qual o valor de λ que maximiza $P\{X = k\}$, $k \geq 0$? Em outras palavras, determine o estimador de máxima verossimilhança para λ.

41. Um número desconhecido, digamos N, de animais de uma certa espécie habitam uma dada região. Para obter alguma informação sobre o tamanho dessa população animal, ecologistas normalmente realizam o seguinte experimento: inicialmente capturam um certo número de animais, digamos m. Em seguida, os animais capturados são marcados de alguma maneira e então devolvidos (soltos) na região em questão. Após um certo período de tempo, necessário para que os animais marcados se dispersem por toda a região, uma nova captura de n animais é feita. Seja X o número de animais marcados que são apreendidos na segunda captura. Se admitirmos que a população de animais na região se manteve inalterada (fixa) entre os dois instantes de captura e que em cada um destes instantes os N animais tinham a mesma probabilidade de serem capturados, determine:

    (a) $P\{X = i\}$ em função de N (desconhecido).

    (b) O estimador de máxima verossimilhança para N se são capturados i animais marcados, isto é, se $X = i$.

5

Modelos Probabilísticos Contínuos

No Capítulo 4, apresentamos os modelos probabilísticos discretos, isto é, aqueles descritos por variáveis aleatórias discretas. Neste capítulo apresentaremos os modelos probabilísticos descritos por variáveis aleatórias que possuem uma densidade de probabilidade. Cada modelo corresponde a uma família de distribuições de probabilidade, expressa por densidades de probabilidade que dependem de um ou mais parâmetros.

5.1 DISTRIBUIÇÃO UNIFORME

Definição 5.1.1 *A variável aleatória X tem distribuição uniforme no intervalo $[a, b]$ se sua densidade de probabilidade for dada por:*

$$f(x) = \frac{1}{b-a} \tag{5.1}$$

para $a \leq x \leq b$ e $f(x) = 0$ fora desse intervalo.

Vamos calcular a média de X, utilizando a expressão (2.15).

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2} . \tag{5.2}$$

O segundo momento de X é dado por:

$$E(X^2) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{a^2+ab+b^2}{3} . \tag{5.3}$$

Substituindo os valores dados por (5.2) e (5.3) na expressão (2.23) nós obtemos a variância de X

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{(b-a)^2}{12} . \tag{5.4}$$

Vamos descrever um experimento cujo resultado nos dá a distribuição uniforme no intervalo $(0, 2\pi)$. Consideremos um segmento de comprimento 2π. Vamos unir as duas pontas desse segmento e formar um círculo de raio unitário. O comprimento desse círculo é precisamente 2π. Vamos fixar um ponteiro no centro desse círculo e vamos então girá-lo, observando-o até que ele venha a parar. Por razões de simetria nós vemos que a chance do ponteiro parar em qualquer arco do círculo é a mesma para qualquer arco de um comprimento dado. Seja X o comprimento do arco determinado pela origem e pelo ponto onde o ponteiro pára. Assim temos uma variável aleatória com distribuição uniforme no intervalo $(0, 2\pi)$.

Se quisermos obter a distribuição uniforme no intervalo $[a, b]$ basta pôr $b-a = 2\pi r$, construir um círculo de raio $r = \frac{b-a}{2\pi}$ e proceder da maneira descrita.

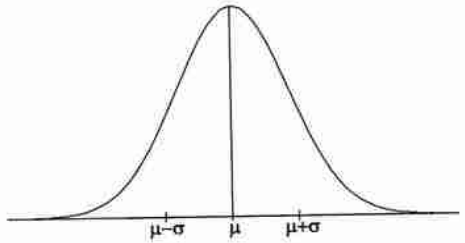
5.2 DISTRIBUIÇÃO NORMAL

A distribuição normal tem grande importância em probabilidade e na estatística em decorrência de um resultado que demonstraremos no Capítulo 7 que é denominado "Teorema do Limite Central". Este afirma que, se realizarmos uma seqüência de experimentos independentes, o resultado de cada um deles expresso pela mesma distribuição de probabilidade ou, em outras palavras, por variáveis aleatórias independentes e identicamente distribuídas com média μ e variância σ^2, então para um número n de parcelas suficientemente grande $S_n = X_1 + X_2 + \cdots + X_n$ tem distribuição de probabilidade aproximadamente normal com média $n\mu$ e variância $n\sigma^2$.

DEFINIÇÃO 5.2.1 *Uma variável aleatória X tem distribuição normal com parâmetros μ e σ^2 se sua densidade de probabilidade é dada por:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \tag{5.5}$$

para $-\infty < x < \infty$. As constantes μ e σ^2 satisfazem as condições: $-\infty < \mu < \infty$ e $\sigma^2 > 0$.



Observando-se a função de densidade $f(x)$, vemos que a mesma é simétrica em relação a μ, daí decorrendo que $E(X) = \mu$. Vamos, no entanto, a título de exercício, determinar esse valor a partir da definição, pela fórmula (2.15).

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx =$$

Fazendo-se a transformação de variáveis $z = \frac{x-\mu}{\sigma}$ temos:

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (\mu + z\sigma) e^{-z^2/2} \sigma dz =$$

$$= \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \sigma \int_{-\infty}^{\infty} z e^{-z^2/2} dz.$$

Na primeira parcela na expressão acima temos:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = 1.$$

Na segunda parcela a integral é nula, pois trata-se da integral de uma função ímpar. Temos então:

$$E(X) = \mu. \tag{5.6}$$

Calculemos a variância de X que tem distribuição normal com densidade dada por (5.5)

$$\begin{aligned} \text{Var}(X) &= E(X - E(X))^2 = \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx. \end{aligned}$$

Fazendo-se a transformação de variáveis $z = \frac{x-\mu}{\sigma}$, a expressão acima transforma- se em:

$$\int_{-\infty}^{\infty} z^2\sigma^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/2}\sigma dz = \sigma^2 \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.$$

Integrando-se por partes, pondo $u = z$ e $v = ze^{-z^2/2}$, é fácil verificar que a integral do lado direito é igual a 1 e portanto

$$\text{Var}(X) = \sigma^2. \tag{5.7}$$

Obtivemos assim o significado dos parâmetros μ e σ^2 na fórmula (5.1). μ é a média dessa distribuição e σ^2 é a variância.

Mostraremos a seguir que a transformação linear, que consiste em subtrair de X sua média e dividir pelo desvio padrão, isto é, por $Z = \frac{X-\mu}{\sigma}$, nos dá uma variável aleatória Z que tem distribuição normal com média zero e variância 1. Na realidade qualquer transformação linear de uma variável com distribuição normal tem também distribuição normal. Sugerimos ao leitor que verifique esta afirmação.

Seja X uma variável aleatória com distribuição normal de média μ e variância σ^2, que doravante indicaremos $N(\mu, \sigma^2)$, e seja $Z = \frac{X-\mu}{\sigma}$. A função de distribuição de X no ponto y é dada por:

$$F(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

Seja F_Z a função de distribuição de Z:

$$\begin{aligned} F_Z(z) &= P[Z \le z] = P[\frac{X-\mu}{\sigma} \le z] = \\ &= P[X \le \mu + \sigma z] = \int_{-\infty}^{\mu+\sigma z} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx. \end{aligned}$$

Fazendo-se nesta integral a transformação de variáveis $\frac{x-\mu}{\sigma} = t$, a expressão acima se transforma em:

$$\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2}\sigma dt = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \tag{5.8}$$

Mas esta última expressão é a função de distribuição de uma variável aleatória com distribuição normal com média zero e variância 1, logo Z tem distribuição $N(0,1)$.

Lembramos ao leitor que a função $e^{-x^2/2}$ não tem primitiva, isto é, não existe uma função $G(x)$ tal que $G'(x) = e^{-x^2/2}$. Deste modo a integral em (5.8), que fornece a função de distribuição da distribuição $N(0,1)$, tem que ser calculada numericamente.

Por esta razão é que se constroem tabelas para obter as probabilidades associadas à distribuição $N(0,1)$. Sabendo-se que uma tranformação linear de uma variavel aleatória com distribui ção $N(0,1)$ conduz a uma outra variável aleatória que tem distribuição normal, obtém-se as probabilidades para qualquer distribuição normal.

Como a densidade da distribuição $N(0,1)$ é simétrica em torno da origem, basta que se tabule, por exemplo, os valores das probabilidades $P[0 \le Z \le x]$ para $x \ge 0$. No apêndice apresentamos um resumo dessa tabela extraído de Bussab e Severo que apresentam a tabela nesse formato.

A área hachurada dada na tabela para o valor x é igual a $P[0 \le Z \le x]$.

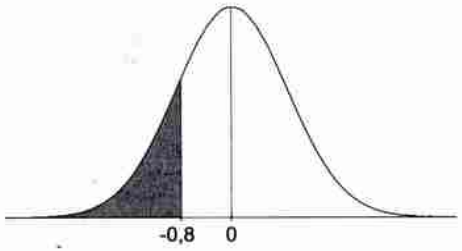
Seja $F_Z(z) = P[Z \le z]$ a função de distribuição de Z que tem distribuição $N(0,1)$.

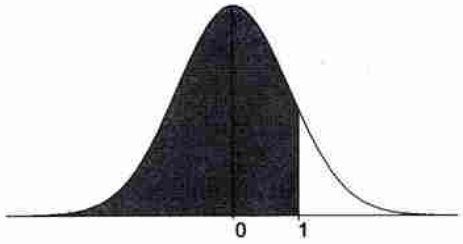
Vamos utilizar a tabela para determinar os valores de $F_Z(z)$ para alguns valores de z. Para $z = -0,8$, tem-se:

$$P[Z < -0,8] = P[Z < 0] - P[-0,8 < Z < 0] =$$

$$= 0,5 - P[0 < Z < 0,8] =$$

$$= 0,5 - 0,288 = 0,212$$





Para $z = 1$

$$P[Z \leq 1] = P[Z < 0] + P[0 < Z < 1] = 0,5 + 0,341 = 0,841.$$

Calculemos agora $P[Z > -0,5]$.

$$P[Z > -0,5] = P[-0,5 < Z < 0] + P[Z \geq 0] =$$

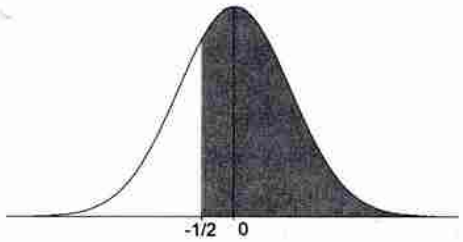
$$= P[0 < Z < 0,5] + 0,5 = 0,191 + 0,5 = 0,691.$$

EXEMPLO 5.2.1 Seja X uma variável aleatória com distribuição $N(30, 9)$. Calcular $P[25 < X < 31]$.

Utilizando-se o fato de que $Z = \frac{X-30}{3}$ tem distribuição $N(0, 1)$ e utilizando-se a tabela da página 244, no apêndice, tem-se:

$$P[25 < X < 31] = P[\frac{25-30}{3} < \frac{X-30}{3} < \frac{31-30}{3}] =$$

$$= P[-1,667 < Z < \frac{1}{3}] = 0,581 .$$



EXEMPLO 5.2.2 Seja X uma variável aleatória com distribuição $N(10, 100)$. Calcule (a) $P(X > 30)$; (b) $P(10 < X < 20)$.

(a) $P(X > 30) = P(\frac{X-10}{10} > \frac{30-10}{10})$. Mas $Z = \frac{X-10}{10}$ tem distribuição $N(0, 1)$. Assim,

$$P(X > 30) = P\left(\frac{X-10}{10} > 2\right) = P(Z > 2) =$$

$$= \frac{1}{2} - P(0 < Z < 2) = \frac{1}{2} - 0,47725 = 0,02275.$$

$$\text{(b)} \quad P(10 < X < 20) = P\left(\frac{10-10}{10} < \frac{X-10}{10} < \frac{20-10}{10}\right) =$$

$$= P\left(0 < \frac{X-10}{10} < 1\right) =$$

$$= P(0 < Z < 1) = 0,34134.$$

EXEMPLO 5.2.3 Seja X uma variável aleatória com distribuição normal (5, 16). Calcule (a) $P(X < 3)$; (b) $P(3 < X < 9)$.

(a) Pondo $Z = \frac{X-5}{4}$, então Z tem distribuição $N(0, 1)$

$$P(X < 3) = P\left(\frac{X-5}{4} < \frac{3-5}{4}\right) = P(Z < -0,5) =$$

$$= P(Z > 0,5) = \frac{1}{2} - P(0 < Z < 0,5) =$$

$$= \frac{1}{2} - 0,191 = 0,309.$$

$$\begin{aligned}
\text{(b) } P(3<X<9) &= P\left(\frac{3-5}{4}<\frac{X-5}{4}<\frac{9-5}{4}\right)= \\
&= P\left(-\frac{1}{2}<Z<1\right)= \\
&= P\left(-\frac{1}{2}<Z<0\right)+P(0<Z<1)= \\
&= P\left(0<Z<\frac{1}{2}\right)+P(0<Z<1)= \\
&= 0,19146+0,34134=0,5328.
\end{aligned}$$

Vamos agora obter a função geradora de momentos da distribuição normal com média zero e variância 1. Denotando esta função geradora por $\phi_Z(t)$ pela expressão (2.27) tem-se:

$$\Phi_Z(t)=\int_{-\infty}^{\infty} e^{tx}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx.$$

No expoente da exponencial, nesta última integral vamos completar o quadrado, isto é, vamos pôr:

$$\begin{aligned}
\exp\left\{tx-\frac{x^2}{2}\right\} &= \exp\left\{\frac{-x^2+2tx-t^2}{2}\right\}e^{t^2/2}= \\
&= \exp\left\{-\frac{(x-t)^2}{2}\right\}e^{t^2/2}.
\end{aligned}$$

Substituindo-se na integral acima, tem-se

$$\Phi_Z(t)=e^{t^2/2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-t)^2}{2}\right\}dx.$$

Nesta última expressão a integral em questão é a integral de uma densidade normal com média t e variância 1 e é portanto igual a 1, logo:

$$\Phi_Z(t)=e^{t^2/2}. \tag{5.9}$$

Seja X uma variável aleatória com distribuição $N(\mu,\sigma^2)$. Vimos que $Z=\dfrac{X-\mu}{\sigma}$ tem distribuição $N(0,1)$. Dessa expressão tem-se $X=\sigma Z+\mu$.

Vamos exprimir a função geradora de momentos de X, que denotamos por $\phi_X(t)$, em termos da função geradora de momentos de Z, denotada $\phi_Z(t)$.

$$\begin{aligned}
\phi_X(t) &= E(e^{tX})=E(e^{t(\sigma Z+\mu)})= \\
&= E(e^{t\mu}e^{t\sigma Z})=e^{t\mu}E(e^{(t\sigma)Z})
\end{aligned}$$

ou seja,

$$\phi_X(t)=e^{t\mu}\phi_Z(t\sigma). \tag{5.10}$$

Pela expressão (5.9), $\phi_Z(t)=e^{t^2/2}$ e substituindo-se nessa expressão tem-se

$$\phi_X(t)=e^{t\mu}e^{t\sigma^2/2}=e^{t\mu+(t^2\sigma^2)/2}. \tag{5.11}$$

Chamamos a atenção do leitor que (5.10) expressa a função geradora de momentos de uma variável qualquer X obtida por uma transformação linear de outra variável aleatória Z.

Sugerimos ao leitor que obtenha $E(X)$ e $\mathrm{Var}(X)$ a partir da função geradora de momentos usando a expressão (5.11).

5.3 MODELOS PROBABILÍSTICOS PARA TEMPOS DE VIDA

Nesta seção apresentaremos alguns modelos que são utilizados para descrever o tempo de vida de um componente ou de um sistema. Entendemos sistema num sentido amplo, podendo ser, por exemplo, um organismo vivo, um conjunto de seres vivos, um conjunto de componentes interligados que formam um sistema de controle em uma indústria, um conjunto de semáforos articulados.

Seja X uma variável aleatória cujos valores representam os tempos de vida de um componente. Denotaremos por $f(x)$ a densidade de probabilidade de X e por $F(x)$ sua função de distribuição. Um dos problemas de interesse consiste em determinar o tempo de vida de um sistema, conhecendo-se os tempos de vida de seus componentes. Para certos tipos especiais de sistemas este problema é tratado em textos de teoria da confiabilidade.

Vamos introduzir o conceito de confiabilidade do sistema (ou de um componente) no instante x como a probabilidade que o sistema não tenha falhado até o instante x.

DEFINIÇÃO 5.3.1 *A confiabilidade de um sistema cuja função de distribuição do tempo de vida é $F(x)$ é por definição:*

$$R(x) = 1 - F(x). \tag{5.12}$$

Um outro conceito importante no estudo da confiabilidade de sistemas é o de taxa de falha, cuja definição damos a seguir.

DEFINIÇÃO 5.3.2 *A taxa de falha de um sistema cujo tempo de vida tem função de distribuição $F(x)$, com densidade de probabilidade $f(x)$ é dada por:*

$$h(x) = \frac{f(x)}{1 - F(x)}. \tag{5.13}$$

A taxa de falha, que é uma função do tempo, é aproximadamente igual à probabilidade de que ocorra falha num intervalo de tempo Δx, pequeno em relação a x, após o instante x, dado que não houve falha até esse instante. De fato:

$$\lim_{\Delta x \to 0} \frac{P[x < X \leq x + \Delta x | X > x]}{\Delta x} = \lim_{\Delta x \to 0} \frac{P[x < X \leq x + \Delta x]}{(\Delta x) P[X > x]} =$$

$$= \frac{1}{1 - F(x)} \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{f(x)}{1 - F(x)}.$$

5.3.1 *Distribuição Exponencial*

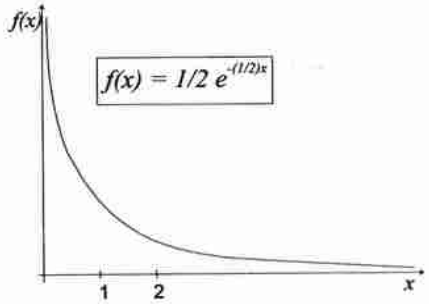
O modelo exponencial que apresentaremos a seguir é um modelo para um tempo de vida com taxa de falha constante. Aqui nos restringiremos a apresentar algumas propriedades probabilísticas do modelo.

DEFINIÇÃO 5.3.3 *Uma variável aleatória X tem distribuição exponencial com parâmetro λ se sua densidade de probabilidade é da forma:*

$$f(x) = \lambda e^{-\lambda x} \tag{5.14}$$

para $x > 0$ e zero no complementar, onde λ é uma constante positiva.

O gráfico dessa densidade para $\lambda = 1/2$ é:



A função de distribuição $F(x)$ da variável aleatória X, com essa densidade é dada por:

$$F(x) = \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x} \tag{5.15}$$

para $0 < x < \infty$ e $F(x) = 0$ para $x < 0$.

A função de confiabilidade da variável aleatória X é dada por:

$$R(x) = 1 - F(x) = e^{-\lambda x} \tag{5.16}$$

para $0 < x < \infty$ e $R(x) = 0$ para $x < 0$.

Utilizando-se a fórmula (5.13), obtemos a taxa de falha $h(x)$ da distribuição exponencial.

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$$

Vemos que esta é uma constante, como afirmamos de início. A interpretação da taxa de falha ser constante é a seguinte: Se um componente tem taxa de falha constante, então não importa por quanto tempo ele já esteja funcionando, a chance de ele vir a falhar num

pequeno intervalo de tempo subseqüente é independente do tempo decorrido.

Uma forma equivalente de expressar essa propriedade consiste em mostrarmos que a distribuição exponencial não sofre desgaste ou, em outra linguagem, não tem memória. Considere um componente que tem distribuição de tempo de vida exponencial. Se ele durou até o instante t, então a probabilidade condicional dele durar mais s unidades de tempo além do instante t, é a mesma que um componente novo venha a durar s unidades de tempo. De fato, seja X o tempo de vida desse componente:

$$P[X > t+s|X > t] = \frac{P[X > t+s, X > t]}{P[X > t]} = \frac{P[X > t+s]}{P[X > t]} =$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P[X > s].$$

Vamos agora determinar a função geradora de momentos, a média e a variância da distribuição exponencial.

$$\phi(t) = E(e^{tX}) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^\infty e^{-(\lambda-t)x}dx = \frac{\lambda}{\lambda - t}\,.$$

Observe que a integral acima existe para t tal que $0 < t < \lambda$. Derivando-se $\phi(t)$ nós obtemos:

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2}\,.$$

Pondo-se $t = 0$ nós obtemos:

$$E(X) = \phi'(0) = \frac{1}{\lambda}\,. \tag{5.17}$$

Deixamos a cargo do leitor verificar que $E(X^2) = \frac{2}{\lambda^2}$ e, portanto, que $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

O tempo de vida de componentes eletrônicos em geral pode ser modelado por uma distribuição exponencial, pois, dentro de certos limites, a taxa de falha dos mesmos pode ser considerada constante. O mesmo não ocorre com componentes mecânicos de um sistema, pois estes em geral sofrem desgaste e têm taxa de falha crescente.

EXEMPLO 5.3.1 Seja X uma variável aleatória exponencial de parâmetro 2. Calcule: (a) $P(X < 1,5)$; (b) $P(X > 0,4)$.

A densidade de probabilidade é dada por $f(x) = 2e^{-2x}$ para $x > 0$ e $f(x) = 0$ para $x < 0$. (a) Utilizando-se a expressão (2.1) com $a = 0$ e $b = 10$ tem-se:

$$P(X < 1,5) = \int_0^{1,5} 2e^{-2x}dx = e^{-2x}|_0^{1,5} = 1 - e^{-3} = 0,67.$$

(b) Utilizando-se novamente a expressão (2.1), agora com $a = 4$ e $b = \infty$, tem-se:

$$P(X > 0,4) = \int_{0,4}^\infty 2e^{-2x}dx = -e^{-2x}|_{0,4}^\infty = e^{-0,8} = 0,45.$$

EXEMPLO 5.3.2 Seja X uma variável aleatória exponencial tal que $E(X) = 3$. Calcule: (a) $P(3 < X < 6)$; (b) $P(X > 3)$; (c) $P(X > 6|X > 3)$. Se $E(X) = 3$ segue-se da expressão (5.17) que $\lambda = 1/3$ e portanto a densidade de probabilidade de X, $f_X(x) = \frac{1}{3}e^{-\frac{1}{3}x}$, para $x > 0$ e $f_X(x) = 0$ para $x < 0$.

(a) $P(3 < X < 6) = \int_3^6 \frac{1}{3}e^{-\frac{1}{3}x}dx = -e^{-\frac{1}{3}x}|_3^6 =$

$= -e^{-\frac{1}{3}.6} - (-e^{-\frac{1}{3}.3}) = -e^{-2} + e^{-1} \Rightarrow P(3 < X < 6) =$

$= e^{-1} - e^{-2} = 0,232.$

(b) $P(X > 3) = \int_3^\infty \frac{1}{3}e^{-\frac{1}{3}x}dx = e^{-\frac{1}{3}x}|_3^\infty = 0 - (-e^{-\frac{1}{3}.3}) =$

$= e^{-1} = 0.3679.$

(c) $P(X > 6|X > 3) = \frac{P(X > 6, X > 3)}{P(X > 3)} = \frac{P(X > 6)}{P(X > 3)} =$

$$= \frac{\int_6^\infty \frac{1}{3}e^{-\frac{1}{3}x}dx}{\int_3^\infty \frac{1}{3}e^{-\frac{1}{3}x}dx} = \frac{-e^{-\frac{1}{3}x}|_6^\infty}{-e^{-\frac{1}{3}x}|_3^\infty} = \frac{e^{-2}}{e^{-1}} = e^{-1} \Rightarrow$$

$\Rightarrow P(X > 6|X > 3) = e^{-1} = 0,3679.$

Esta igualdade é casual entre os resultados de (b) e (c)? Procure justificar sua afirmação.

5.3.2 Distribuição Gama

DEFINIÇÃO 5.3.4 *Uma variável aleatória X tem distribuição gama com parâmetros $r > 0$ e $\lambda > 0$ se sua densidade de probabilidade é dada por:*

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} \tag{5.18}$$

para $x > 0$ e $f(x) = 0$ para $x < 0$.

Se o parâmetro r for inteiro, esta família de distribuições de probabilidade recebe o nome de distribuições de Erlang. Essa distribuição tem papel de destaque em telefonia e recebeu esse nome em homenagem ao engenheiro dinamarquês Erlang, que empregou essas distribuições pela primeira vez para modelar duração de chamadas telefônicas por volta de 1917. Vamos determinar a função geradora de momentos da distribuição gama:

$$\phi(t) = \int_0^\infty e^{tx} \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} dx = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty e^{-(\lambda - t)x} x^{r-1} dx =$$

(fazendo a transformação de variável $y = (\lambda - t)x$)

$$= \frac{\lambda^r}{\Gamma(r)} \int_0^\infty \frac{e^{-y} y^{r-1} dy}{(\lambda - t)^{r-1}(\lambda - t)} = (\frac{\lambda}{\lambda - t})^r \frac{1}{\Gamma(r)} \int_0^\infty e^{-y} y^{r-1} dy.$$

Levando em conta que a integral no lado direito é o valor de $\Gamma(r)$ tem-se:

$$\phi(t) = (\frac{\lambda}{\lambda - t})^r . \tag{5.19}$$

Notemos que, se r for inteiro, a função geradora de momentos obtida é a r-ésima potência da função geradora de momentos da distribuição exponencial com parâmetro λ. Pela proposição 3.3.3 isto implica que uma variável aleatória com distribuição de Erlang de parâmetros r e λ é igual a soma de r variáveis aleatórias independentes e identicamente distribuídas, com distribuição exponencial com parâmetro λ.

Derivando-se a função geradora de momentos dada por (5.19), podemos obter a média e a variância da distribuição γ. Deixamos a cargo do leitor mostrar que:

$$E(X) = \frac{r}{\lambda} \qquad \text{e} \qquad \text{Var}X = \frac{r}{\lambda^2}.$$

5.3.3 Distribuição de Weibull

A distribuição de Weibull é uma das principais distribuições utilizadas para modelar tempos de vida. Vamos definir a distribuição de Weibull por meio de sua densidade de probabilidade e então obter a função de distribuição e a taxa de falha.

DEFINIÇÃO 5.3.5 *Uma variável aleatória tem distribuição de Weibull com parâmetros $\beta > 0$ e $\eta > 0$ se sua densidade de probabilidade, para $x \geq 0$, é dada por:*

$$f(x) = \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta - 1} \exp\left\{-(\frac{x}{\eta})^\beta\right\} \tag{5.20}$$

e $f(x) = 0$ para $x < 0$.

Observemos que se derivarmos $\exp\left\{-(\frac{x}{\eta})^\beta\right\}$ com relação a x nós obtemos $-f(x)$. Daí concluímos que a função de distribuição de X é dada por:

$$F(x) = 1 - \exp\left\{-(\frac{x}{\eta})^\beta\right\}. \tag{5.21}$$

A partir de $f(x)$ e de $F(x)$ otemos a taxa de falha da distribuição de Weibull:

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{\frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1} \exp\left\{-(\frac{x}{\eta})^\beta\right\}}{\exp\left\{-(\frac{x}{\eta})^\beta\right\}}.$$

Simplificando temos:

$$h(x) = \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1} \tag{5.22}$$

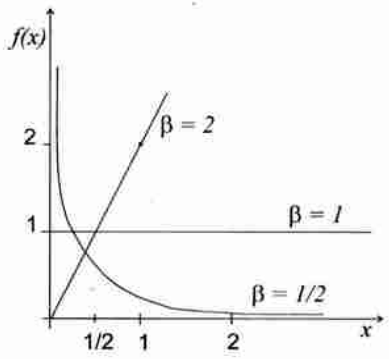
para $x > 0$ e zero para $x \leq 0$.

O parâmetro η é um parâmetro de escala e β é um parâmetro de forma. A taxa de falha da distribuição de Weibull dada por (5.22) é uma função crescente para $\beta > 1$ e decrescente para $\beta < 1$. Para $\beta = 1$ a distribuição de Weibull reduz-se à distribuição exponencial.

Vamos, a título de ilustração, dar alguns exemplos de curvas de taxas de falhas de distribuições de Weibull. Seja η fixado igual a 1.

Para $\beta = 2$ tem-se: $h(x) = 2x$ para $x > 0$ e $h(x) = 0$ para $x < 0$.

Para $\beta = 1$ tem-se $h(x) = 1$ para $x > 0$ e $h(x) = 0$ para $x < 0$.
Para $\beta = \frac{1}{2}$ tem-se: $h(x) = \frac{1}{2}x^{-1/2}$.



Vamos calcular os momentos de ordem r, para $r \geq 1$, da variável aleatória X que tem distribuição de Weibull, cuja densidade é dada por (5.20). Em particular calcularemos a média e o segundo momento obtendo então a variância.

$$E(X^r) = \int_0^\infty x^r \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1} \exp\left\{-(\frac{x}{\eta})^\beta\right\} dx$$

Pondo $\frac{x}{\eta} = y$ vem:

$$\int_0^\infty \frac{\beta}{\eta} y^{\beta-1} y^r \eta^r e^{-y^\beta} \eta dy$$

Pondo $t = y^\beta$ tem-se $y = t^{1/\beta}$ e fazendo a substituição vem:

$$\eta^r \int_0^\infty \frac{\beta y^{\beta-1} y^r e^{-t}}{\beta y^{\beta-1}} dt = \eta^r \int_0^\infty t^{r/\beta} e^{-t} dt$$

Lembrando que $\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$, vem:

$$E(X^r) = \eta^r \Gamma(1 + \frac{r}{\beta}). \qquad (5.23)$$

Para $r = 1$ têm-se: $E(X) = \eta\Gamma(1 + \frac{1}{\beta})$.
Para $r = 2$ obtemos: $E(X^2) = \eta^2\Gamma(1 + \frac{2}{\beta})$.
Utilizando-se a expressão $\mathrm{Var}X = E(X^2) - [E(X)]^2$ obtemos a variância.

Observação: Para as distribuições Weibull e Gama que contêm a exponencial é possível introduzir-se um parâmetro de posição, que representa o valor abaixo do qual nenhuma falha pode ocorrer. Para fazer isso, basta substituir-se nas densidades correspondentes a essas distribuições x por $x - \delta$ e considerar-se a densidade definida para $x > \delta$.

5.4 EXERCÍCIOS

1. Se X é uniformemente distribuída no intervalo $(0, 20)$, calcule a probabilidade de: (a) $X < 3$. (b) $X > 12$. (c) $4 < X < 11$. (d) $|X - 3| < 4$.

2. Se X é uma variável aleatória normal com parâmetros $\mu = 3$ e $\sigma^2 = 9$, determine: (a) $P\{2 < X < 5\}$. (b) $P\{X > 0\}$. (c) $P\{|X - 3| > 6\}$.

3. Se X é uma variável aleatória uniformemente distribuída no intervalo $(-3, 7)$, determine: (a) A função de distribuição de X. (b) $P\{|X - 1| \leq 2\}$. (c) $P(|X| > 3\}$.

4. Ônibus chegam a um determinado ponto de parada em intervalos de tempo de quinze minutos a partir de 7 horas da manhã, isto é, os ônibus chegam ao ponto às 7h00, 7h15, 7h30, 7h45, e assim por diante. Se o instante de chegada de um passageiro ao ponto é uniformemente distribuído entre 7h00 e 7h30, determine a probabilidade: (a) De que ele espere menos que 5 minutos até a chegada de um ônibus. (b) De que ele espere mais de 10 minutos até a chegada de um ônibus.

5. Seja X o número de caras observadas em 40 lançamentos de uma moeda honesta. Determine a probabilidade de $X = 20$. Utilize a aproximação normal e então compare com o resultado obtido através da distribuição binomial.

6. Suponha que a duração de uma ligação telefônica em uma cabine pública (em minutos) é uma variável aleatória exponencial com parâmetro $\lambda = \frac{1}{10}$. Se uma pessoa chega imediatamente

a sua frente na cabine telefônica pública, ache a probabilidade de que você terá que esperar
(a) mais que 10 minutos; (b) entre 10 e 20 minutos.

7. Para deteminar a eficiência de uma certa dieta na redução da quantidade de colesterol na corrente sanguínea, 100 pessoas são submetidas a esta dieta por um intervalo de tempo bastante prolongado. Em seguida, são registrados os níveis de colesterol destas pessoas. O nutricionista responsável pelo experimento decidiu endossar a dieta se pelo menos 65% das pessoas apresentarem um nível de colesterol menor após serem submetidas à dieta. Qual é a probabilidade de que o nutricionista endosse a nova dieta se, na verdade, ela não tem efeito algum sobre o nível de colesterol? (Admita que se a dieta não tem efeito algum sobre a quantidade de colesterol, então o nível de colesterol de cada pessoa será menor após a dieta com probabilidade 1/2.)

8. Suponha que um componente eletrônico tenha um tempo de vida X (em unidades de 1000 horas) que é considerado uma variável aleatória com função densidade de probabilidade $f(x) = e^{-x}$, $x > 0$. Suponha também que o custo de fabricação de um item seja R$ 2,00 e o preço de venda seja R$ 5,00. O fabricante garante devolução total do dinheiro se $x \leq 0,9$. Qual o lucro esperado do fabricante por item produzido?

9. Um ponto é escolhido ao acaso em um segmento de reta de comprimento L. Determine a probabilidade de que a razão entre o menor e o maior segmentos obtidos após a escolha deste ponto seja menor que 1/4.

10. Um ônibus viaja entre duas cidades, A e B, separadas por uma distância de 100 quilômetros. Se o ônibus sofre uma avaria durante a viagem, a distância do local da avaria à cidade A tem uma distribuição uniforme no intervalo $(0, 100)$. Existe uma estação de reparo na cidade A, uma estação de reparo na cidade B, e uma outra estação no centro da rota entre A e B. Foi sugerido que seria mais eficiente ter as três estações de reparo localizadas a 25, 50 e 75 quilômetros, respectivamente, de A. Você concorda com esta sugestão? Justifique.

11. A confiabilidade de um mecanismo eletrônico é a probabilidade de que ele funcione sob as condições para as quais foi planejado. Uma amostra de 1000 destes itens é escolhida ao acaso, e estes itens selecionados são testados. Calcule a probabilidade de se obter pelo menos 30 itens defeituosos, supondo que a confiabilidade de cada item é 0,95. Indique que suposições está fazendo ao efetuar tal cálculo.

12. A quantidade de chuva anual em uma certa região é normalmente distribuída com $\mu = 40$ e $\sigma = 4$. Qual é a probabilidade de que, começando a registrar os índices pluviométricos este ano, serão necessários mais do que 10 anos para se registrar uma quantidade de chuva anual maior que 50? Que suposições você está fazendo?

13. Um corpo de bombeiros está para ser construído ao longo de uma estrada de comprimento A, $A < \infty$. Se incêndios ocorrem de maneira uniforme ao longo desta estrada, onde o corpo de bombeiros deveria ser instalado para minimizar a distância esperada ao incêndio? Isto é, determine a que minimize $E[|X - a|]$, onde X, que denota a posição da ocorrência de um incêndio nesta estrada, é uniformemente distribuído em $(0, A)$.

14. O tempo (em horas) necessário para reparar uma máquina é uma variável aleatória exponencialmente distribuída com parâmetro $\lambda = \frac{1}{2}$. Determine:
    (a) A probabilidade de que o tempo de reparo exceda duas horas.
    (b) A probabilidade condicional de que o tempo de reparo será maior que onze horas dado que a duração do reparo excede nove horas.

15. Uma empresa produz automóveis e garante a restituição da quantia paga se qualquer automóvel apresentar algum defeito grave no prazo de seis meses. A empresa produz automóveis do

tipo A comum e do tipo B de luxo, com um lucro de R$ 1.000,00 e R$ 2.000,00, respectivamente, caso não haja restituição, e com um prejuízo de R$ 3.000,00 e R$ 8.000,00, respectivamente, se houver restituição. Suponha que o tempo para a ocorrência de algum defeito grave seja, em ambos os casos, uma variável aleatória com distribuição normal, respectivamente com médias 9 meses e 12 meses, e variâncias 4 meses e 9 meses. Se tivesse que planejar uma estratégia de marketing para a empresa, você incentivaria as vendas dos automóveis do tipo A ou do tipo B. Justifique.

16. Uma máquina funciona se pelo menos três de cinco dos seus componentes estiverem funcionando. Cada componente, independentemente dos demais, funciona por um tempo que é uma variável aleatória com função densidade de probabilidade dada por $f(x) = \frac{1}{5}e^{-\frac{1}{5}x}$, $x > 0$, x em horas. (a) Determine a probabilidade de que a máquina funcione por mais de cinco horas. (b) Determine o número médio de componentes funcionando dez horas após a máquina ser ligada. Deixe claras as suposições que está fazendo para resolver os itens (a) e (b).

17. Suponha que o número de horas X que uma máquina opera antes de falhar tem uma distribuição contínua com função densidade de probabilidade $f(x) = \frac{1}{10}e^{-\frac{x}{10}}$, $x > 0$. Suponha que no momento em que a máquina é colocada para funcionar, você deve decidir quando retornar para inspecioná-la. Se você retorna antes de a máquina falhar, você incorre num custo de B reais por ter perdido a inspeção. Se você retorna após a quebra da máquina, você incorre num custo de C reais por hora pelo período de tempo durante o qual a máquina não estava operando após a falha. Qual o número ótimo de horas que você deve esperar antes de retornar para a inspeção de forma a minimizar seu custo esperado?

18. Uma fração desconhecida p de uma certa população é de fumantes, e uma amostragem aleatória com reposição será usada para avaliar o valor de p. Deseja-se encontrar p com um erro que não exceda 0,005, com alta probabilidade. Em outras palavras, deseja-se que, com alta probabilidade, a proporção amostral a ser obtida não difira mais de 0,005 da proporção p real de fumantes. Qual deve ser o tamanho da amostra n para que isto ocorra? Explicite as suposições que está fazendo.
Admitindo-se agora que $p \leq 20\%$. Qual deve ser o tamanho da amostra n para que a condição exposta acima seja satisfeita?

19. Com base na definição de taxa de falha, determine a função taxa de falha de uma variável aleatória gama com parâmetros t e λ.

20. Calcule a função taxa de falha de X quando X é uniformemente distribuída em $(0, a)$, $a > 0$.

21. Costuma-se dizer que a taxa de mortalidade de pessoas fumantes é, em cada idade, o dobro da taxa de mortalidade de pessoas não-fumantes. O que isto significa? Isto significa que a probabilidade de uma pessoa não-fumante sobreviver um determinado número de anos corresponde a duas vezes a probabilidade de uma pessoa fumante, de mesma idade, sobreviver este mesmo número de anos?

22. Seja X uma variável aleatória não-negativa e contínua. Mostre que:
$$E(X) = \int_0^\infty P(X > t)dt \ .$$

23. Para uma variável aleatória não-negativa T, o tempo médio de vida residual é definido por $m(t) = E[T - t|T \geq t]$.
(a) Mostre que $m(t) = \frac{\int_t^\infty R(x)dx}{R(t)}$. Interprete o significado de $m(t)$.
(b) Calcule $m(t)$ quando T tem distribuição exponencial de parâmetro λ.
(c) Calcule $m(t)$ quando T é uniformemente distribuído no intervalo $(0, A)$, $A < \infty$.

24. Determine a função geradora de momentos de uma variável aleatória uniformemente distribuída no intervalo (a, b).

25. Uma variável aleatória X tem distribuição beta com parâmetros α e β, $\alpha, \beta > 0$, se sua função densidade de probabilidade é dada por $f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1$, e zero caso contrário, onde $B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. (a) Determine $E(X)$ e $\text{Var}(X)$. (b) Para $a > 1$ e $b > 1$, determine x^* tal que $f(x^*) = \max_{0<x<1} f(x)$.

    A distribuição beta é comumente utilizada em inferência Bayesiana quando se estudam quantidades de interesse assumindo valores no intervalo $(0, 1)$ como, por exemplo, a proporção de fumantes em uma certa população. Observe ainda que, se $\alpha = \beta = 1$, X é uniformemente distribuída em $(0, 1)$.

26. Uma variável aleatória X tem distribuição de Laplace com parâmetro λ, $\lambda > 0$, se sua função densidade de probabilidade é dada por:

$$f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}, \quad x \in \mathbb{R}.$$

    Determine a função de distribuição de X.

27. Se X é uniformemente distribuída no intervalo $(0, 1)$, qual a distribuição de $Y = -\log X$?

28. Se X é uma variável aleatória exponencial com parâmetro λ e $c > 0$ uma constante real, determine a distribuição de $Y = cX$.

29. A mediana de uma variável aleatória contínua, cuja função de distribuição F é o valor real m tal que $F(m) = \frac{1}{2}$. Determine a mediana da variável aleatória X se X é (a) Uniformemente distribuída no intervalo (a, b). (b) Normal com parâmetros μ e σ^2. (c) Exponencial com parâmetro λ.

30. A moda de uma variável aleatória contínua, cuja função densidade de probabilidade é f, é o valor de x para o qual f atinge valor máximo. Calcule a moda da variável aleatória X nos casos (a), (b) e (c) do exercício anterior.

31. Mostre que se X é uma variável aleatória Weibull com parâmetros α e β, então $Y = (\frac{X}{\alpha})^\beta$ é uma variável aleatória exponencial com parâmetro 1, e vice-versa.

32. Se X é uniformemente distribuída em (a, b), qual variável aleatória que possui uma relação linear com X é uniformemente distribuída no intervalo $(0, 1)$? Em outras palavras, determine $\alpha, \beta \in \mathbb{R}$, tais que $Y = \alpha X + \beta$ tem distribuição uniforme $(0, 1)$.

33. Seja X uma variável aleatória contínua com função de distribuição F qualquer, defina a variável aleatória Y por $Y = F(X)$. Mostre que Y é uniformemente distribuída no intervalo $(0, 1)$. Este resultado é de grande importância em técnicas de simulação de variáveis aleatórias, uma vez que, para obtermos observações de uma variável aleatória contínua com função de distribuição F qualquer, basta gerarmos valores da distribuição uniforme $(0, 1)$ e, em seguida, aplicarmos F^{-1} a estes valores, onde F^{-1} é a função inversa de F.

34. Se X é uniformemente distribuída no intervalo $(-1, 1)$, determine a função densidade de probabilidade da variável aleatória $Y = |X|$.

35. Ache a distribuição de $R = A\text{sen}\theta$, onde A é uma constante fixada e θ é distribuída uniformemente em $(-\frac{\pi}{2}, \frac{\pi}{2})$. Tal variável aleatória aparece na teoria de balística. Se um projétil é lançado de um certo ponto da superfície da Terra, formando um ângulo α em relação a esta superfície e com velocidade v, então o alcance deste lançamento, R, pode ser expresso por $R = \frac{v^2}{g}\text{sen}2\alpha$, onde g é a constante gravitacional.

36. Dizemos que uma variável aleatória X tem distribuição de Pareto com parâmetros α e β, $\alpha, \beta > 0$, se sua função densidade de probabilidade é dada por: $f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1}$, $x > \beta$, $f(x) = 0$, $x < \beta$. Mostre que, para $\alpha > 1$, $E(X) = \frac{\alpha\beta}{\alpha-1}$ e para $\alpha > 2$, $\text{Var}(X) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}$.

Esta distribuição é freqüentemente utilizada em economia, em problemas de distribuição de renda. Neste caso, β pode representar algum nível mínimo de renda de uma população e $\int_a^b f(x)dx$ a proporção de indivíduos desta população com renda entre a e b.

37. Outra distribuição bastante comum em economia é a distribuição lognormal. A variável aleatória X assumindo valores positivos tem uma distribuição lognormal com parâmetros μ e σ^2, $\mu \in \mathbb{R}$ e $\sigma^2 > 0$, se $Y = \log X$ tem distribuição normal com média μ e variância σ^2. Determine: (a) A função densidade de probabilidade de X. (b) $E(X)$ e $\mathrm{Var}(X)$.

38. Seja X uma variável aleatória normal com média 0 e variância 1. Seja $Y = X^2$. (a) Determine a função densidade de probabilidade de Y. (b) Calcule $E(Y)$ e $\mathrm{Var}(Y)$.

    Y construída deste modo é denominada uma variável aleatória com distribuição qui-quadrado (χ^2) com 1 grau de liberdade. De um modo geral pode-se construir variáveis aleatórias χ^2 com N graus de liberdade, cujo papel é fundamental em diversos métodos estatísticos.

39. Seja X uma variável aleatória com distribuição gama com parâmetros α e β. Defina a variável aleatória Z como sendo $Z = 2\beta X$. Determine a função densidade de probabilidade de Z, com $2\alpha \in \mathbb{N}$. Qual a distribuição de Z se $\alpha = \frac{1}{2}$?

40. Seja X uma variável aleatória com distribuição beta de parâmetros α e β. Defina a variável aleatória Y como sendo $Y = \frac{\beta X}{\alpha(1-X)}$. Obtenha a função densidade de probabilidade de Y.

    (A variável aleatória Y construída deste modo possui distribuição denominada distribuição F-Snedecor com 2α e 2β graus de liberdade. Como a distribuição de qui-quadrado, a distribuição F-Snedecor também é de extrema importância em diversos métodos estatísticos.)

41. Seja X uma variável aleatória contínua com função densidade de probabilidade f, que depende apenas de X e de um parâmetro θ, isto é,

    1) $f(x) \geq 0, \forall x \in \mathbb{R}$; 2) $\int_{-\infty}^{\infty} f(x)dx = 1$;

    3) $f(x) = f(x, \theta)$.

    Dizemos que f pertence à família exponencial uniparamétrica de distribuições se f pode ser escrita da seguinte forma:

    $$f(X) = f(X, \theta) = S(X).t(\theta).e^{a(X).b(\theta)},$$

    $\forall x \in D$, onde D não depende de θ e s, t, a e b são funções reais. Verifique quais das distribuições a seguir pertencem à família exponencial: (a) Distribuição uniforme no intervalo $(0, A)$, $A < \infty$. (b) Distribuição normal com média μ e variância 4. (c) Distribuição exponencial com parâmetro λ, $\lambda > 0$. (d) Distribuição beta com parâmetros α e β, α conhecido.

6

Variáveis Aleatórias Contínuas Multidimensionais

6.1 DENSIDADES DE PROBABILIDADE

A exemplo do que fizemos no caso unidimensional, vamos introduzir o modelo probabilístico associado a variáveis aleatórias bidimensionais por meio da densidade de probabilidade dessas variáveis aleatórias. A analogia que fizemos entre uma densidade de probabilidade de uma variável aleatória unidimensional e uma distribuição de massas em um eixo também pode ser feita entre a densidade de uma variável aleatória bidimensional e uma distibuição de massas no plano. Na última seção deste capítulo estenderemos os conceitos apresentados para o caso de variáveis aleatórias de dimensão maior que 2.

Definição 6.1.1 *Uma função f(x, y) definida para* $-\infty < x < +\infty$, $-\infty < y < +\infty$, *não-negativa e satisfazendo a condição*

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)dxdy = 1 \tag{6.1}$$

é denominada uma função densidade de probabilidade da variável aleatória bidimensional (X,Y) *se para todo subconjunto* B *de pontos do* $\mathbb{R}^2$ *tivermos:*

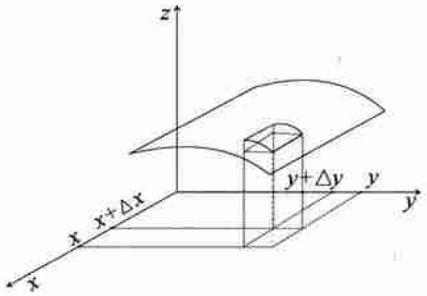
$$P[(X,Y) \in B] = \int\int_B f(x,y)dxdy\ . \tag{6.2}$$

Supõe-se que a função $f(x,y)$ é contínua, exceto possivelmente para um conjunto enumerável de valores de x ou de y.

Doravante designaremos a função densidade de probabilidade simplesmente por densidade de probabilidade.

A interpretação da função densidade de probabilidade no caso bidimensional é a seguinte: seja (x,y) um ponto do plano e consideremos um retângulo de lados Δx e Δy construido a partir do ponto (x,y). A probabilidade de que (X,Y) pertença a esse retângulo é aproximadamente igual ao volume do paralelepípedo de lados Δx e Δy e cuja altura é $f(x,y)$, ou seja,

$$P[x < X \leq x + \Delta x, y < Y \leq y + \Delta y] \simeq f(x,y)\Delta x \Delta y \tag{6.3}$$



Note que o valor exato dessa probabilidade é igual ao volume da figura delimitada pela função $f(x,y)$ e pelo retângulo de lados Δx e Δy, que é igual à integral dupla de $f(x,y)$ nesse retângulo.

Dada a densidade conjunta das variáveis X e Y, podemos determinar as densidades de X e de Y isoladamente, e essas densidades são denominadas densidades marginais de X e de Y e denotadas por $f_X(x)$ e $f_Y(y)$, respectivamente. Elas são obtidas da seguinte maneira:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy \tag{6.4}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx. \tag{6.5}$$

Note que $f_X(x)$ e $f_Y(y)$ são de fato densidades de probabilidade, pois são não-negativas e sua integral na reta é igual a um, pela equação (6.1). Quando não houver menção explícita ao conjunto de valores onde $f(x,y) > 0$, suporemos que esse conjunto é o plano, isto é, o conjunto $R^2 = \{(x,y) : -\infty < x < \infty, -\infty < y < \infty\}$. Neste caso as expressões (6.4) e (6.5) valem para $-\infty < x < \infty$ e para $-\infty < y < \infty$ respectivamente.

EXEMPLO 6.1.1 Seja (X,Y) uma variável aleatória bidimensional cuja densidade de probabilidade é dada por: $f(x,y) = \frac{3}{80}(x^2 + xy)$ para $0 \leq x \leq 2$ e $0 \leq y \leq 4$ sendo zero no complementar desse retângulo. Determinar as densidades marginais de X e de Y e calcular a probabilidade $P[X + Y \leq 3]$.

Para valores de x fora do intervalo (0,2) a densidade conjunta $f(x,y) = 0$ e portanto por (6.4), $f_X(x) = 0$

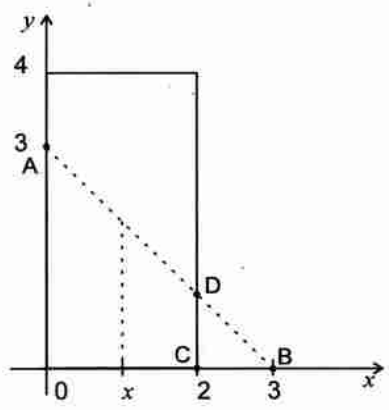
Para x fixado, $0 \leq x \leq 2$, os valores de y tais que $f(x,y) > 0$ pertencem ao intervalo (0,4), logo:

$$\begin{aligned} f_X(x) &= \int_0^4 \frac{3}{80}(x^2 + xy)dy = \int_0^4 \frac{3}{80}x^2 dy + \int_0^4 \frac{3}{80}xydy = \\ &= \frac{3}{80}(x^2 y|_0^4 + \frac{xy^2}{2}|_0^4) = \frac{3}{20}x^2 + \frac{3}{10}x\ . \end{aligned}$$

Analogamente, para valores de y não pertencentes ao intervalo (0,4) temos $f_Y(y) = 0$. Para y tal que $0 \leq y \leq 4$, temos:

$$\begin{aligned} f_Y(y) &= \int_0^2 \frac{3}{80}(x^2 + xy)dx = \frac{3}{80}(x^3/3|_0^2 + yx^2/2|_0^2) = \\ &= \frac{3}{80}(\frac{8}{3} + 2y) = \frac{1}{10} + \frac{3}{40}y\ . \end{aligned}$$

Para calcular $P[X+Y \leq 3]$, notemos que o evento $[X+Y \leq 3]$ é formado pelo conjunto de pontos (x,y) que pertencem ao triângulo OAB. Notando que a densidade conjunta $f(x,y) = 0$ no triângulo DCB segue da expressão (6.2) que



$$P[X+Y\leq 3]=\int_0^2\left\{\int_0^{3-x}\frac{3}{80}(x^2+xy)dy\right\}dx \qquad (6.6)$$

pois, para x fixado tal que $0\leq x\leq 2$, os valores de y, que satisfazem $x+y\leq 3$, pertencem ao intervalo $(0,3-x)$. Vamos calcular a integral interna:

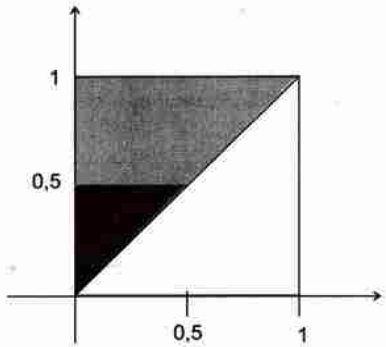
$$\int_0^{3-x}\frac{3}{80}(x^2+xy)dy=\frac{3}{80}\left(x^2y\Big|_0^{3-x}+\frac{xy^2}{2}\Big|_0^{3-x}\right)=$$

$$=\frac{3}{80}(x^2(3-x)+x(3-x)^2/2)=\frac{3}{160}(-x^3+9x).$$

Substituindo em (6.6) vem:

$$P[X+Y\leq 3]=\int_0^2\frac{3}{160}(-x^3+9x)dx=\frac{3}{160}\left(-\frac{x^4}{4}\Big|_0^2+\frac{9x^2}{2}\Big|_0^2\right)=\frac{21}{80}\,.$$

EXEMPLO 6.1.2 Seja (X,Y) um par de variáveis aleatórias cuja densidade conjunta é dada por $f(x,y)=8xy$, para $0<x<y<1$, e $f(x,y)=0$ no complementar. Determinar as densidades marginais de probabilidade de X e de Y e calcular a probabilidade $P[0<X<0,5,\ 0<Y<0,5]$.



Observe que os valores de x variam entre zero e um. Para cada x fixado nesse intervalo os valores de y variam entre x e 1. Fora desse intevalo a densidade conjunta $f(x,y)$ é igual a zero. Assim temos, para x tal que: $0<x<1$,

$$f_X(x)=\int_x^1 8xydy=8xy^2/2|_x^1=4x-4x^3=4x(1-x^2)$$

e $f_X(x)=0$ caso contrário.

Para obter a densidade marginal de Y, observe que para cada valor de y tal que $0<y<1$ a densidade conjunta $f(x,y)$ é positiva para os valores de x tais que $0<x<y$. Deste modo, para y tal que $0<y<1$ temos:

$$f_Y(y)=\int_0^y 8xydx=8yx^2/2|_0^y=4y^3$$

e $f_Y(y)=0$ fora desse intervalo.

Usando a expressão (6.2) com $B=\{(x,y):0<x<0,5,\ 0<y<0,5\}$, tem-se:

$$P[0<X<0,5,\ 0<Y<0,5]=\int_0^{0,5}\int_0^{0,5}f(x,y)dxdy=$$

$$=\int_0^{0,5}\int_x^{0,5}8xydxdy=\int_0^{0,5}\left[\int_x^{0,5}8xydy\right]dx=$$

$$=\int_0^{0,5}[8xy^2/2|_x^{0,5}]dx=\int_0^{0,5}(x-4x^3)dx=0,0625.$$

EXEMPLO 6.1.3 Uma variável aleatória bidimensional (X,Y) tem distribuição normal bivariada com parâmetros $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ e ρ se sua densidade de probabilidade for dada por:

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\}$$

para $-\infty < x < \infty$, $-\infty < y < \infty$ e constantes $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$ e $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ e $|\rho| \leq 1$.

Vamos determinar as densidades marginais de X e de Y.

Designando-se a densidade marginal de X por $f_X(x)$ e usando a expressão (6.4), tem-se:

$$f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\} dy =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\frac{(x-\mu_1)^2}{\sigma_1^2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\} dy.$$

Somando-se e subtraindo-se $\frac{1}{2(1-\rho^2)}\frac{\rho^2(x-\mu_1)^2}{\sigma_1^2}$ no expoente do integrando tem-se

$$\frac{1}{2(1-\rho^2)}\frac{\rho^2(x-\mu_1)^2}{\sigma_1^2} - \frac{1}{2(1-\rho^2)}\left(\frac{\rho^2(x-\mu_1)^2}{\sigma_1} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right) =$$

$$= \frac{\rho^2(x-\mu_1)^2}{2(1-\rho^2)\sigma_1^2} - \frac{1}{2(1-\rho^2)}\left(-\frac{\rho(x-\mu_1)}{\sigma_1} + \frac{y-\mu_2}{\sigma_2}\right)^2 =$$

$$= \frac{\rho^2(x-\mu_1)^2}{2(1-\rho^2)\sigma_1^2} - \frac{1}{2(1-\rho^2)\sigma_2^2}\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2.$$

Substituindo-se na expressão da densidade vem:

$$f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{\rho^2(x-\mu_1)^2}{2(1-\rho^2)\sigma_1^2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho_2)\sigma_2^2}\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2\right\} dy.$$

Simplificando o expoente do termo fora da integral e agrupando convenientemente tem-se:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\} \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\pi(1-\rho^2)\sigma_2^2}(y-\mu_2-\rho\frac{\sigma_2^2}{\sigma_1}(x-\mu_1))^2\right\} dy.$$

Notando que para cada x fixado $\mu_2 - \rho\frac{\sigma_1}{\sigma_2}(x-\mu_1)$ é uma constante e que portanto

$$\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_2^2}(y-\mu_2-\rho\frac{\sigma_2}{\sigma_2}(x-\mu_1))^2\right\}$$

é a densidade de uma variável aleatória com distribuição normal com média $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$ e variância $(1-\rho^2)\sigma_2^2$, concluímos que sua integral na reta toda é igual a um. Assim

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu)^2\right\}$$

que é a densidade marginal de X, que tem portanto distribuição normal com média μ_1 e variância σ_1^2. Analogamente, Y tem distribuição normal com média μ_2 e variância σ_2^2. Veremos que ρ é o coeficiente de correlação entre X e Y.

6.2 FUNÇÕES DE DISTRIBUIÇÃO

A definição 2.4.1 de função de distribuição de uma variável aleatória unidimensional se estende de maneira natural para uma

variável aleatória bidimensional, substituindo-se o conjunto dos valores $[X \leq x]$ pelo conjunto $[X \leq x, Y \leq y]$.

DEFINIÇÃO 6.2.1 *A função de distribuição de uma variável aleatória bidimensional (X,Y), que denotaremos por $F(x,y)$ é dada por:*

$$F(x,y) = P[X \leq x, Y \leq y] \tag{6.7}$$

para $-\infty < x < \infty$ e $-\infty < y < \infty$.

Observemos que para calcular a probabilidade do evento $[X \leq x, Y \leq y]$ no caso de variáveis contínuas, decorre da definição 6.1.1 de densidade que esta probabilidade é igual à integral da densidade conjunta de X e de Y no conjunto $[X \leq x, Y \leq y]$, isto é:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv. \tag{6.8}$$

EXEMPLO 6.2.1 Determinar a função de distribuição da variável aleatória bidimensional do Exemplo 6.1.1 cuja densidade de probabilidade é:

$$f(x,y) = \frac{3}{80}(x^2 + xy)$$

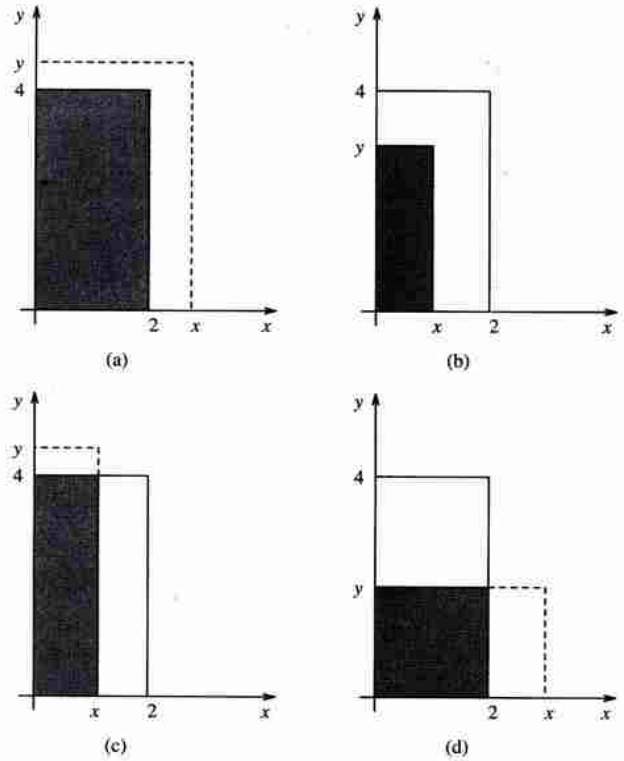
para $0 < x < 2$, $0 < y < 4$ e zero no complementar.

Como $f(x,y) = 0$ fora do retângulo assinalado, segue-se que $F(x,y) = 0$ exceto no primeiro quadrante. Na mesma figura pode-se observar que para $x \geq 2$ e para $y \geq 4$ o evento $[X \leq x, Y \leq y]$ inclui todos os pontos do retângulo assinalado, onde $f(x,y) \geq 0$ e, portanto, integrando-se a densidade conjunta no mesmo, o valor dessa integral é igual a 1. Ver Figura (a).

Resta-nos calcular $F(x,y)$ para $x \geq 2$ e $0 \leq y \leq 4$ e para $0 \leq x \leq 2$ e $y \geq 0$.

Consideremos um ponto (x,y) tal que $0 \leq x \leq 2$ e $0 \leq y \leq 4$. Ver Figura (b). Escrevendo a probabilidade do evento $[X \leq x, Y \leq y]$ como uma integral iterada da densidade conjunta temos:

$$F(x,y) = \int_0^x [\int_0^y \frac{3}{80}(u^2 + uv)dv]du.$$



Calculando-se a integral em relação a v temos;

$$\int_0^y \frac{3}{80}(u^2 + uv)dv = \frac{3}{80}\left[u^2 v|_0^y + u\frac{v^2}{2}|_y^0\right] = \frac{3}{80}\left(u^2 y + u\frac{y^2}{2}\right).$$

Substituindo-se na integral acima obtemos:

$$F(x,y) = \int_0^x \frac{3}{80}(u^2 y + u\frac{y^2}{2})du = \frac{3}{80}[y\frac{u^3}{3}|_0^x + y^2\frac{u^2}{4}|_0^x =$$

$$= \frac{3}{80}(y\frac{x^3}{3} + y^2\frac{x^2}{4}) = \frac{1}{320}x^2 y(4x + 3y).$$

Observe que se considerarmos um ponto (x,y) fixado tal que $0 \leq x \leq 2$ e $y > 4$, então os pontos onde a densidade $f(u,v) > 0$ são os pontos (u,v) tais que $0 \leq u \leq x$ e $0 \leq v \leq 4$. Ver Figura (c). Assim o valor de $F(x,y)$ é:

$$F(x,y) = \int_0^x \left[\int_0^4 \frac{3}{80}(u^2 + uv)dv\right] du.$$

A integral interna em relação a v é a densidade marginal de X que já foi determinada no exemplo 6.1.1 e é igual a $\frac{3}{20}x^2 + \frac{3}{10}x$. Substituindo na expressão de $F(x,y)$ acima obtemos:

$$F(x,y) = \int_0^x (\frac{3}{20}u^2 + \frac{3}{10}u)du = \frac{3}{20}\frac{u^3}{3}|_0^x + \frac{3}{10}\frac{u^2}{2}|_0^x = \frac{x^3}{20} + \frac{3}{20}x^2.$$

Deixamos a cargo do leitor calcular o valor de $F(x,y)$ para $x > 2$ e $0 \leq y \leq 4$, que é obtida integrando a densidade conjunta na região hachurada na Figura (d).

Vimos que, calculando-se a integral dupla da densidade de probabilidade de uma variável aleatória bidimensional (X,Y), obtemos sua função de distribuição. Por outro lado, se calcularmos a derivada parcial de $F(x,y)$ em relação a x e depois em relação a y, obtemos a densidade conjunta de (X,Y). De fato, isto decorre de uma aplicação iterada do teorema fundamental do cálculo.

Calculando-se a derivada parcial em relação a y na expressão (6.8), obtém-se:

$$\frac{\partial}{\partial y}\int_{-\infty}^x \int_{-\infty}^y f(u,v)dudv = \int_{-\infty}^x \frac{\partial}{\partial y}[\int_{-\infty}^y f(u,v)dv]du = \int_{-\infty}^x f(u,y)du.$$

Calculando-se $\frac{\partial}{\partial x}$ de ambos os membros vem:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}\ . \tag{6.9}$$

Faremos agora alguns comentários sobre as propriedades da função de distribuição de uma variável aleatória bidimensional (X,Y). Não apresentaremos demonstrações.

A função de distribuição da variável aleatória (X,Y) satisfaz as seguintes propriedades:

i) $0 \leq F(x,y) \leq 1$

ii) $F(x,y)$ é contínua à direita em cada uma das variáveis x e y.

iii) $\lim_{x\to-\infty} F(x,y) = 0$; $\lim_{y\to-\infty} F(x,y) = 0$; $\lim_{x\to\infty, y\to\infty} F(x,y) = 1$.

Note que estas propriedades são extensões naturais para o caso bidimensional das propriedades da função de distribuição de uma

variável aleatória unidimensional X. O intervalo $(-\infty, x]$ é substituído pelo retângulo $(-\infty, x]$x$(-\infty, y]$. Quando fazemos $x \to -\infty$, este retângulo tende ao conjunto vazio no plano e sua imagem inversa tende ao conjunto vazio no espaço amostral que tem probabilidade zero. Raciocínio análogo explica porque $lim_{y\to-\infty}F(x,y) = 0$. Quando $x \to \infty$ e $y \to \infty$, o retângulo em questão tende para o plano todo, cuja imagem inversa é o espaço amostral que tem probabilidade 1.

Para uma variável aleatória unidimensional X verificamos que:

$$P[a < X \leq b] = F(b) - F(a)\ .$$

Na situação bidimensional vimos que o análogo a um intervalo da forma $(a,b]$ é um retângulo da forma $\{(x,y) : a_1 < x \leq b_1, a_2 < y \leq b_2\}$. Deixamos como exercício ao leitor mostrar que a probabilidade dos valores de X e de Y pertencerem a esse retângulo é dada pela expressão:

$$P[a_1 < X \leq b_1, a_2 < Y \leq b_2] = F(b_1,b_2) - F(a_1,b_2) + \\ -F(a_2,b_1) + F(a_1,a_2). \tag{6.10}$$

Como a expressão no lado esquerdo da igualdade acima é uma probabilidade, tem-se:

$$F(b_1,b_2) - F(a_1,b_2) - F(a_2,b_1) + F(a_1a_2) \geq 0. \tag{6.11}$$

Assim vemos que a função de distribuição $F(x,y)$ satisfaz também à propriedade seguinte:

iv) Quaisquer que sejam $a_1 < b_1$ e $a_2 < b_2$ a expressão (6.11) vale.

6.3 INDEPENDÊNCIA

DEFINIÇÃO 6.3.1 *As variáveis aleatórias X e Y cuja densidade conjunta é $f(x,y)$, para $-\infty < x < \infty$ e $-\infty < y < \infty$, e cujas densidades marginais são denotadas $f_X(x)$ e $f_Y(y)$, são ditas independentes se para todo par de valores (x, y) tivermos:*

$$f(x,y) = f_X(x)f_Y(y). \tag{6.12}$$

A condição de independência de duas variáveis aleatórias pode ser expressa também por meio da função de distribuição conjunta e das funções de distribuição marginais das duas variáveis como mostra o seguinte lema:

LEMA 6.3.1 *Sejam X e Y variáveis aletórias contínuas. Sejam $F(x,y)$, $F_X(x)$ e $F_Y(y)$ as funções de distribuição conjuntas e marginais de X e Y respectivamente e sejam $f(x,y)$, $f_X(x)$ e $f_Y(y)$ as respectivas densidades de probabilidade. X e Y são independentes se e somente se*

$$F(x,y) = F_X(x)F_Y(y) \tag{6.13}$$

para todo par de valores das variáveis X e Y.

DEMONSTRAÇÃO Suponhamos X e Y independentes, isto é, tais que a relação (6.12) é válida. Integrando-se ambos os membros dessa igualdade e notando que a integral dupla do primeiro membro é o valor da função de distribuição conjunta de X e Y, de acordo com (6.8), tem-se:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u)f_Y(v)dudv = \tag{6.14}$$

Integrando iteradamente:

$$= \int_{-\infty}^{x} f_X(u)du \int_{-\infty}^{y} f_Y(v)dv = F_X(x)F_Y(y).$$

Obtivemos assim (6.13).

Suponhamos agora que a relação (6.13) seja válida. Calculemos a segunda derivada parcial de ambos os membros dessa igualdade. Decorre de (6.9) que a derivada segunda do lado esquerdo é $f(x,y)$. A derivada do lado direito nos fornece $f_X(x)f_Y(y)$, que é o que queríamos mostrar.

EXEMPLO 6.3.1 Seja $f(x,y) = xe^{-x(y+1)}$ para $0 < x < \infty$, $0 < y < \infty$ e zero no complementar, a densidade conjunta de X e Y. Verificar se X e Y são independentes.

Determinemos as densidades marginais de X e Y que denotaremos por $f_X(x)$ e $f_Y(y)$ respectivamente.

Para $0 < x < \infty$, usando-se a expressão (6.4) e fazendo-se a transformação de variável $z = y + 1$, tem-se:

$$f_X(x) = \int_0^{\infty} xe^{-x(y+1)}dy = \int_1^{\infty} xe^{-xz}dz = e^{-x}.$$

Para $0 < y < \infty$, usando-se a expressão (6.5) e multiplicando-se e dividindo-se por $y + 1$, obtém-se o produto de $\frac{1}{y+1}$ por uma integral que é igual à esperança de uma variável aleatória com distribuição exponencial com parâmetro $y+1$, cujo valor é $\frac{1}{y+1}$. Assim, o valor de $f_Y(y)$ é $\frac{1}{(y+1)^2}$.

$$f_Y(y) = \int_0^{\infty} xe^{-x(y+1)}dx = \frac{1}{y+1}\int_0^{\infty} x(y+1)e^{-x(y+1)}dx = \frac{1}{(y+1)^2}.$$

Vê-se que X e Y não são independentes pois a densidade conjunta $f(x,y)$ não se fatora no produto das marginais $f_X(x)$ e $f_Y(y)$.

EXEMPLO 6.3.2 Seja $f(x,y) = \lambda\mu^2 ye^{-(\lambda x+\mu y)}$, para $0 < x < \infty$, $0 < y < \infty$ e $f(x,y) = 0$ no complementar, a densidade conjunta da variável aleatória (X,Y). Verificar se X e Y são independentes.

Calculemos $f_X(x)$ e $f_Y(y)$.

$$f_X(x) = \int_0^{\infty} \lambda\mu^2 ye^{-(\lambda x+\mu y)}dy = \lambda\mu e^{-\lambda x}\int_0^{\infty} \mu ye^{-\mu y}dy.$$

Note que a integral nessa expressão é igual à esperança de uma variável aleatória com distribuição exponencial com parâmetro μ e portanto igual a $\frac{1}{\mu}$. Substituindo-se, tem-se:

$$f_X(x) = \lambda e^{-\lambda x}.$$

Calculemos $f_Y(y)$.

$$f_Y(y) = \int_0^{\infty} \lambda\mu^2 ye^{-\mu y}e^{-\lambda x}dx = \mu^2 ye^{-\mu y}\int_0^{\infty} \lambda e^{-\lambda x}dx = \mu^2 ye^{-\mu y},$$

pois a integral acima é a integral da densidade exponencial com parâmetro λ, que é igual a 1.

Calculando-se o produto $f_X(x)f_Y(y)$, verificamos que ele é igual a $f(x,y)$ e portanto X e Y são independentes.

Vimos que a esperança do produto de duas variáveis aleatórias discretas é igual ao produto das esperanças dessas variáveis, veja-se o lema 3.3.2. Esta é uma propriedade que vale para variáveis aleatórias em geral. Aqui verificaremos essa propriedade para duas variáveis contínuas. Para facilidade de referência, enunciaremos essa propriedade como um lema.

LEMA 6.3.2 *Sejam X e Y variáveis aleatórias contínuas. Se elas são independentes, então*

$$E(XY) = E(X)E(Y). \tag{6.15}$$

DEMONSTRAÇÃO Seja $f(x,y)$ a densidade conjunta de X e Y, e sejam $f_X(x)$ e $f_Y(y)$, respectivamente, as densidades marginais de X e de Y.

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy = \\ &= \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy = E(X)E(Y). \end{aligned}$$

A exemplo do que verificamos no caso discreto, a recíproca desse lema não é verdadeira, isto é, a expressão (6.15) pode ser verdadeira e X e Y não serem independentes.

A demonstração da expressão (6.15), ou seja, a esperança do produto de variáveis aleatórias independentes é igual ao produto das esperanças dessas variáveis, foi dada para variáveis discretas no lema 3.3.2 e para variáveis contínuas no lema 6.3.2.

Aplicando-se o lema 3.3.2 obtivemos o lema 3.3.3. Sugerimos ao leitor que, de modo análogo, obtenha, aplicando o lema 6.3.2, o seguinte resultado para variáveis aleatórias contínuas.

LEMA 6.3.3 *Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias contínuas independentes. Seja $X = X_1 + X_2 + \cdots + X_n$ e denotemos por $\phi_X(t)$, $\phi_{X_1}(t)$, ..., $\phi_{X_n}(t)$ as funções geradoras de momentos de X, $X_1, X_2, \ldots, X_n$, respectivamente. Tem-se:*

$$\phi_X(t) = \phi_{X_1}(t)\cdots\phi_{X_n}(t).$$

Note que se $X_1, X_2, \ldots, X_n$ são identicamente distribuídas, suas funções geradoras de modo são iguais a $\phi_{X_1}(t)$ e portanto $\phi_X(t) = [\phi_{X_1}(t)]^n$.

6.4 DISTRIBUIÇÕES DE FUNÇÕES DE VARIÁVEIS ALEATÓRIAS

Seja (X,Y) uma variável aleatória bidimensional da qual é dada a função de distribuição conjunta $F(x,y)$ ou sua densidade conjunta $f(x,y)$. A frase pode parecer estranha, pois dada uma sabemos que podemos obter a outra. No entanto veremos que muitas vezes por razões técnicas é mais conveniente trabalhar com uma delas.

Suponhamos que seja dada a função de distribuição $F(x,y)$ da variável aleatória (X,Y). Seja g uma função definida no $\mathbb{R}^2$ que assume valores reais, que suporemos contínua. A função $Z = g(X,Y)$ é uma variável aleatória definida no mesmo espaço amostral. Vamos determinar a função de distribuição de Z, que denotaremos por $G(z)$.

$$G(z) = P[Z \leq z] = P[g(X,Y) \leq z] = P[(X,Y) \in B_z] \tag{6.16}$$

onde $B_z = \{(x,y) : g(x,y) \leq z\}$.

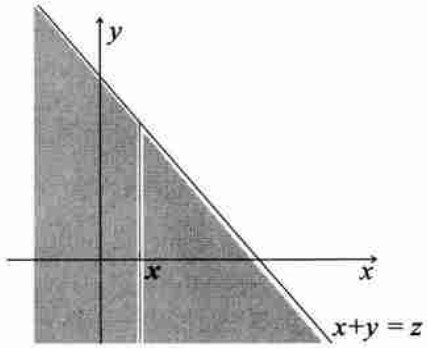
Vamos dar a seguir uma série de exemplos onde esse método é empregado para determinar a distribuição de probabilidade de algumas funções de um par de variáveis aleatórias. Determinaremos, por exemplo, a função de distribuição da soma e do quociente de duas variáveis aleatórias.

EXEMPLO 6.4.1 Seja (X,Y) uma variável aleatória bidimensional com densidade de probabilidade $f(x,y)$. Determinar a função de distribuição e a densidade de probabilidade de $Z = X + Y$.

Denotemos por F_Z a função de distribuição de Z e consideremos o conjunto $B_z = \{(x, y) : x + y \leq z\}$.

Utilizando-se a expressão (6.16) temos:

$$G_Z(z) = \int\int_{B_z} f(x,y)dxdy = \int_{-\infty}^{\infty}[\int_{-\infty}^{z-x} f(x,y)dy]dx.$$



Fazendo-se uma transformação de variáveis na integral interna, pondo-se $u = x + y$, ou seja $y = u - x$ e substituindo, vem:

$$F_Z(z) = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{z} f(x, u-x)du\right]dx.$$

Mudando-se a ordem de integração, o que é permitido, pois o integrando é não negativo, obtém-se:

$$F_Z(z) = \int_{-\infty}^{z}[\int_{-\infty}^{\infty} f(x, u-x)dx]du. \qquad (6.17)$$

Derivando-se em relação a z, obtemos a densidade de probabilidade de Z

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)dx \qquad (6.18)$$

para $-\infty < z < \infty$.

Se as variáveis aleatórias X e Y são independentes $f(x,y) = f_X(x)f_Y(y)$, onde f_X e f_Y são as densidades de probabilidade de X e de Y respectivamente. Substituindo-se em (6.18) obtemos a densidade de $Z = X + Y$ nesse caso.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx. \qquad (6.19)$$

A densidade de Z determinada pela expressão (6.19) é denominada convolução das densidades f_X e f_Y. Esse método é simples conceitualmente, porém requer cuidados especiais na determinação das regiões de integração, e como veremos, há casos que os cálculos são muito complexos. O próximo exemplo ilustra essa afirmação. Nesse exemplo determinaremos a densidade de probalilidade de $Z = X + Y$ através da fórmula (6.19) da convolução e obteremos também essa densidade derivando a função de distribuição obtida usando a expressão (6.17).

EXEMPLO 6.4.2 Sejam X e Y variáveis aleatórias independentes e com a mesma distribuição uniforme no intervalo (a, b), onde a e b são dois números reais tais que $a < b$. Determinemos a densidade de probabilidade de $Z = X + Y$.

Usaremos a expressão (6.19) para determinar a densidade de probabilidade de Z. A densidade de probabilidade comum a X e a Y é a uniforme em (a, b) que é dada por:

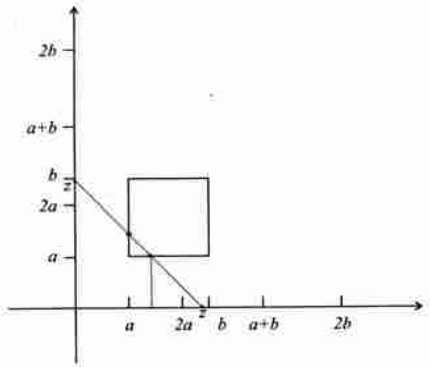
$$f(x) = \frac{1}{b-a}$$

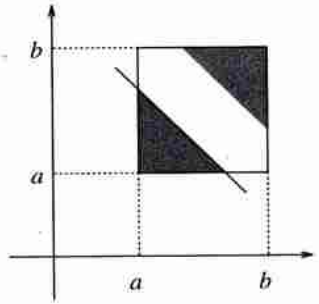
para x tal que $a < x < b$ e $f(x) = 0$ fora do intervalo (a, b).

A densidade de Z será positiva no intervalo dos valores de z tais que $2a < z < 2b$. Aqui temos de tratar separadadmente dois casos, pois os limites de integração na expressão (6.19) são diferentes. De fato, note que valores de $x + y$ constantes correspondem a pontos das retas com inclinação negativa de 45 graus. A reta $x + y = 2a$ toca o retângulo da figura no ponto (a, a) e a reta $x + y = 2b$ toca o retângulo no ponto (b, b). Para valores de z tais que $2a < z < 2b$ as retas $x + y = z$ são paralelas àquelas já mencionadas e delimitam no retângulo da figura os intervalos de integração. Assim para cada z tal que $2a < z < a + b$, tem-se:

$$f(x)f(z-x) = \frac{1}{(b-a)^2}$$

para x pertencente ao intervalo (a, z-a) e igual a zero fora desse intervalo.





Temos portanto:

$$f_Z(z) = \int_a^{z-a} \frac{1}{(b-a)^2} dx = \frac{z-2a}{(b-a)^2}. \qquad (6.20)$$

Obtém-se de modo análogo, para z tal que $a+b < z < 2b$, o seguinte:

$$f_Z(z) = \int_{z-b}^{b} \frac{1}{(b-a)^2} dx = \frac{2b-z}{(b-a)^2}. \qquad (6.21)$$

As expressões (6.20) e (6.21) nos dão os valores da densidade de probalidade de $Z = X + Y$.

Vamos calcular a função de distribuição de Z utilizando a expressão (6.16), isto é, integrando a densidade conjunta de X e Y no conjunto $B_z = \{(x,y) : x + y \leq z\}$. Partiremos de (6.16) e passaremos a (6.19).

Consideremos inicialmente os valores de z tais que $2a \leq z \leq a+b$ e a seguir tais que $a+b \leq z \leq 2b$.

Para z tal que $2a \leq z \leq a+b$, pode-se notar que estaremos integrando no triângulo hachurado do canto inferior.

$$\begin{aligned} F_Z(z) &= \iint_{\{(x,y):2a\leq x+y\leq z\}} \frac{1}{(b-a)^2} dxdy = \\ &= \int_a^{z-a} \left[\int_a^{z-x} \frac{1}{(b-a)^2} dy \right] dx = \frac{(z-2a)^2}{2(b-a)^2}. \end{aligned}$$

Para z tal que $a+b \leq z \leq 2b$, estaremos integrando a densidade conjunta $\frac{1}{(b-a)^2}$ no complementar do triângulo do canto superior em relação ao retângulo já mencionado. Assim, tem-se:

$$\begin{aligned} F_Z(z) &= 1 - \iint_{\{(x,y):z\leq x\leq 2b\}} \frac{1}{(b-a)^2} dxdy = \\ &= 1 - \int_{z-b}^{b} \left[\int_{z-x}^{b} \frac{1}{(b-a)^2} dy \right] dx = 1 - \frac{(2b-z)^2}{2(b-a)^2}. \end{aligned}$$

Deixamos a cargo do leitor verificar que o valor da última integral é aquele indicado e que, derivando-se a função de distribuição, obtém-se a densidade dada por (6.20) e (6.21). Além disso, $F_Z(z) = 0$ para $z < 2a$ e $F_Z(z) = 1$ para $z \geq 2b$.

EXEMPLO 6.4.3 Sejam X e Y variáveis aleatórias contínuas cuja densidade conjunta é denotada $f(x,y)$. Vamos determinar a função de distribuição e a densidade de probabilidade do quociente $Z = \frac{Y}{X}$.

De acordo com a fórmula (6.16) tem-se:

$$F_Z(z) = \int \int_{B_z} f(x,y) dxdy$$

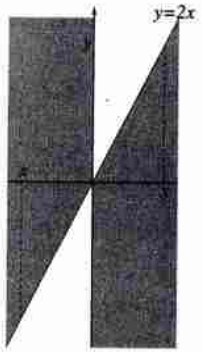
onde $B_z = \{(x,y) : \frac{y}{x} < z\}$.

Observe que, para $z > 0$, por exemplo $z = 2$ na figura a seguir, B_z é igual à reunião das regiões hachuradas.

$$B_z = \{(x,y) : x < 0, y > xz\} \bigcup \{(x,y) : x > 0, y < zx\}.$$

Desse modo, usando-se a expressão (6.16) tem-se:

$$F_Z(z) = \int_{-\infty}^{0} [\int_{xz}^{\infty} f(x,y)dy]dx + \int_{0}^{\infty} [\int_{-\infty}^{xz} f(x,y)dy]dx.$$



Fazendo-se a mudança de variável $y = xu$ na integral entre colchetes, obtém-se:

$$F_Z(z) = \int_{-\infty}^{0}[\int_{z}^{-\infty} xf(x,xu)du]dx + \int_{0}^{\infty}[\int_{-\infty}^{z} xf(x,xu)du]dx =$$

$$= \int_{-\infty}^{0}[\int_{-\infty}^{z}(-x)f(x,xu)du]dx + \int_{0}^{\infty}[\int_{-\infty}^{z} xf(x,xu)du]dx =$$

$$= \int_{-\infty}^{\infty}[\int_{-\infty}^{z}|x|f(x,xu)du]dx.$$

Trocando-se a ordem de integração (que é permitido, pois o integrando é não-negativo), tem-se para $-\infty < z < \infty$:

$$F_Z(z) = \int_{-\infty}^{z}[\int_{-\infty}^{\infty}|x|f(x,ux)dx]du. \qquad (6.22)$$

Esta é a expressão da função de distribuição do quociente Z e derivando-se, obtém-se a densidade de probabilidade de Z, f_Z cujo valor para z tal que $-\infty < z < \infty$ é:

$$f_Z(z) = \int_{-\infty}^{\infty}|x|f(x,xz)dx. \qquad (6.23)$$

EXEMPLO 6.4.4 Sejam X e Y variáveis aleatórias independentes e identicamente distribuídas com distribuição normal com média zero e variância 1. Determinar a distribuição de probabilidade do quociente $\frac{X}{Y}$.

De acordo com a expressão (6.23) a densidade de probabilidade de $Z = \frac{X}{Y}$ é dada por:

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}|x|e^{-\frac{1}{2}(x^2+x^2z^2)}dx =$$

$$= \frac{2}{2\pi}\int_{0}^{\infty}xe^{-\frac{1}{2}x^2(1+z^2)}dx =$$

Fazendo a transformação de variável $y = x\sqrt{1+z^2}$ vem:

$$= \frac{1}{\pi}\int_{0}^{\infty}\frac{y}{\sqrt{1+z^2}}e^{-y^2/2}\frac{dy}{\sqrt{1+z^2}} = \frac{1}{\pi(1+z^2)}\int_{0}^{\infty}ye^{-y^2/2}dy =$$

$$= \frac{1}{\pi(1+z^2)} \text{ para } -\infty < z < \infty$$

pois $\int_{0}^{\infty}ye^{-y^2/2}dy = e^{-y^2/2}\Big|_{0}^{\infty} = 1.$

De modo análogo ao utilizado para determinar a distribuição de probabilidade de uma função de um par de variáveis aleatórias, podemos determinar a distribuição conjunta de uma transformação de uma variável aleatória bidimensional em outra variável aleatória bidimensional.

Seja X_1 e X_2 um par de variáveis aleatórias cuja densidade conjunta é $f(x_1,x_2)$. Sejam g_1 e g_2 funções reais que suporemos contínuas e definamos as variáveis aleatórias Y_1 e Y_2 do seguinte modo:

$$Y_1 = g_1(X_1, X_2)$$

$$Y_2 = g_2(X_1, X_2).$$

Vamos determinar a distribuição conjunta de Y_1 e Y_2 ou, o que é equivalente, da variável aleatória bidimensional $Y = (Y_1, Y_2)$. Utilizando-se a expressão (6.2) tem-se:

$$F_Y(y_1,y_2) = P[Y_1 \le y_1, Y_2 \le y_2] = \int\int_D f(x_1,x_2)dx_1dx_2 \qquad (6.24)$$

onde $D = \{(x_1,x_2) : g_1(x_1,x_2) \le y_1, g_2(x_1,x_2) \le y_2\}$.

Vamos a seguir dar alguns exemplos da aplicação desse método.

EXEMPLO 6.4.5 Sejam X_1 e X_2 variáveis aleatórias independentes e identicamente distribuídas com densidade de probabilidade $f(x) = 2x$ para $0 \leq x \leq 1$, $f(x) = 0$ no complementar.

Determinar a densidade conjunta de $Y_1 = min(X_1, X_2)$ e $Y_2 = max(X_1, X_2)$.

Como X_1 e X_2 são independentes, sua densidade de probabilidade conjunta $f(x_1, x_2) = 4x_1x_2$, para $0 < x_1 < 1$ e para $0 < x_2 < 1$ é igual a zero no complementar desse quadrado.

Denotemos por $F_Y(y_1, y_2)$ a função de distribuição conjunta de Y_1 e Y_2. De acordo com a expressão (6.24) tem-se:

$$F_Y(y_1, y_2) = \int\int_D 4x_1x_2dx_1dx_2$$

$$\begin{aligned} \text{onde } D &= \{(x_1, x_2) : \min(x_1, x_2) \leq y_1, \max(x_1, x_2) \leq y_2\} \\ &= \{(x_1, x_2) : x_1 \leq y_1, x_2 \leq y_1\} \bigcup \\ &\quad \bigcup\{(x_1, x_2) : x_1 \leq y_1, y_1 \leq x_2 \leq y_2\} \bigcup \\ &\quad \bigcup\{(x_1, x_2) : y_1 \leq x_1 \leq y_2, x_2 \leq y_1\}. \end{aligned}$$

Assim

$$\begin{aligned} F_Y(y_1, y_2) &= \int_0^{y_1}\int_0^{y_1} 4x_1x_2dx_1dx_2 + 2\int_0^{y_1}\int_{y_1}^{y_2} 4x_1x_2dx_1dx_2 = \\ &= y_1^4 + 2(y_2^2 - y_1^2)y_1^2 = 2y_1^2y_2^2 - y_1^4 . \end{aligned}$$

A densidade conjunta de (Y_1, Y_2) que denotaremos por f_Y é obtida utilizando-se a expressão (6.9) que reproduzimos aqui para facilidade de referência.

$$f_Y(y_1, y_2) = \frac{\partial^2 F_Y(y_1, y_2)}{\partial y_1 \partial y_2} .$$

Calculando-se esta derivada na expressão acima, obtemos

$$f_Y(y_1, y_2) = 8y_1y_2$$

para $0 \leq y_1 \leq y_2 \leq 1$ e zero no complementar.

O leitor deve ter notado que neste procedimento parte-se da definição de função de distribuição, tanto no caso unidimensional como no bidimensional, e a partir daí calcula-se a probabilidade de interesse, isto é, a probabilidade da imagem inversa do retângulo $(-\infty, x_1)$ x $(-\infty, x_2)$ no espaço amostral das variáveis originais.

Vamos apresentar uma outra técnica que permite obter a densidade de probabilidade conjunta das variáveis transformadas (Y_1, Y_2) quando se conhece a densidade conjunta de (X_1, X_2), utilizando-se um teorema de transformações de variáveis em integrais multiplas.

Relembremos inicialmente a definição do jacobiano de uma transformação do R^2 no R^2. Sejam g_1 e g_2 funções a valores reais definidas no plano e suponhamos que existam e sejam contínuas suas derivadas parciais: $\frac{\partial g_1}{\partial x_1}$, $\frac{\partial g_1}{\partial x_2}$, $\frac{\partial g_2}{\partial x_1}$, $\frac{\partial g_2}{\partial x_2}$. Denotemos por $g = (g_1, g_2)$ a transformação do R^2 no R^2 definida por estas funções.

O jacobiano da transformação g, que denotaremos por Jg, é a função a valores reais que no ponto $x = (x_1, x_2)$ é igual ao determinante da matriz indicada:

$$Jg(x) = \det \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} . \tag{6.25}$$

Seja $X = (X_1, X_2)$ uma variável aleatória bidimensional cuja densidade conjunta de probabilidade é $f_X(x_1, x_2)$ e consideremos as novas variáveis aleatórias dadas por:

$$Y_1 = g_1(X_1, X_2)$$

$$Y_2 = g_2(X_1, X_2)$$

que defininem a transformação $Y = (Y_1, Y_2)$.

Utilizando a expressão (6.2) podemos escrever:

$$\begin{aligned} P[(Y \in A] &= P[g(X) \in A] = P[X \in g^{-1}(A)] = \\ &= \int\int_{g^{-1}(A)} f_X(x_1, x_2)dx_1dx_2. \end{aligned} \tag{6.26}$$

Suponhamos ainda que a transformação $g = (g_1, g_2)$ que leva o ponto (x_1, x_2) no ponto $(g_1(x_1, x_2), g_2(x_1, x_2))$ seja biunívoca e que o jacobiano dessa transformação dado por (6.25) não se anule. Sob essas condições, a transformação inversa g^{-1} é também contínua com derivadas parciais contínuas e seu jacobiano Jg^{-1} dado por

$$Jg^{-1}(y_1, y_2) = \frac{1}{Jg(x_1, x_2)}$$

é diferente de zero, onde $x_1 = g_1^{-1}(y_1, y_2)$ e $x_2 = g_2^{-1}(y_1, y_2)$ e tem-se:

$$\int\int_{g^{-1}(A)} f_X(x_1, x_2)dx_1dx_2 =$$

$$= \int\int_A f_X(g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2)Jg^{-1}(y_1, y_2)dy_1dy_2 \quad (6.27)$$

Vê-se assim que a densidade de probabilidade $f_Y(y_1, y_2)$ de $Y = (Y_1, Y_2)$ é dada por:

$$f_Y(y_1, y_2) = f_X(g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2))Jg^{-1}(y_1, y_2). \quad (6.28)$$

EXEMPLO 6.4.6 Sejam X e Y variáveis aleatórias contínuas cuja densidade de probabilidade conjunta é $f_{X,Y}(x, y)$. Sejam

$$r = \sqrt{x^2 + y^2} \qquad \text{e} \qquad \phi = arctg\frac{y}{x},$$

para $r \geq 0$ e $0 \leq \phi \leq 2\pi$. Determinar a densidade de probabilidade conjunta de (R, Φ).

A transformação que leva (x, y) em (r, ϕ) no domínio mencionado é biunívoca e sua inversa é:

$$x = rcos\phi \qquad \text{e} \qquad y = rsen\phi\ .$$

O jacobiano dessa transformação inversa é igual a:

$$J = det\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{vmatrix} = det\begin{vmatrix} cos\phi & -rsen\phi \\ sen\phi & rcos\phi \end{vmatrix} = r$$

Logo, de acordo com (6.28), tem-se:

$$f_{R,\phi}(r, \phi) = rf_{X,Y}(rcos\phi, rsen\phi). \quad (6.29)$$

EXEMPLO 6.4.7 Sejam X e Y variáveis aleatórias independentes com distribuição N(0,1). Determinar a distribuição conjunta de R e Φ definidas no exemplo anterior e a densidade marginal de R.

A densidade conjunta de X e Y é:

$$f_{X,Y}(x, y) = \frac{1}{2\pi}\exp\left\{-\frac{x^2 + y^2}{2}\right\}.$$

Utilizando-se a expressão (6.29), tem-se:

$$f_{R,\Phi}(r, \phi) = \frac{r}{2\pi}e^{-r^2/2}$$

para $r \geq 0$ e $0 \leq \phi \leq 2\pi$ e zero no complementar.

Denotando-se por f_R a densidade marginal de R e usando-se (6.4), tem-se, para $r \geq 0$:

$$f_R(r) = \int_0^{2\pi} \frac{r}{2\pi}e^{-r^2/2}d\phi = re^{-r^2/2} \quad (6.30)$$

e igual a zero para $r \leq 0$.

EXEMPLO 6.4.8 Sejam X e Y variáveis aleatórias independentes com distribuição exponencial com parâmetro 1, isto é, com densidade de probabilidade $f_X(x) = e^{-x}$, para $x \geq 0$ e igual a zero no complementar. Determinar a densidade conjunta de:

$$U = X + Y \qquad \text{e} \qquad V = \frac{Y}{X}\ . \quad (6.31)$$

Em decorrência da independência, segue-se que a densidade conjunta de X e Y é para $x \geq 0$ e $y \geq 0$ dada por

$$f(x, y) = e^{-(x+y)}$$

e é igual a zero no complementar.

A transformação definida por (6.31) é biunívoca e sua inversa é:

$$x = \frac{u}{1+v}$$

$$y = \frac{uv}{1+v}\ .$$

O jacobiano dessa transformação é:

$$J = det\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = det\begin{vmatrix} \frac{1}{1+v} & \frac{-u}{(1+v)^2} \\ \frac{1}{1+v} & \frac{u}{(1+v)^2} \end{vmatrix} = \frac{2u}{(1+v)^2}\ .$$

Aplicando-se (6.28), obtém-se a densidade conjunta de U e V:

$$f_{U,V}(u,v) = ue^{-u}\frac{1}{(1+v)^2} \tag{6.32}$$

para $u \geq 0$ e $v \geq 0$ e zero no complementar desse intervalo.

Observemos que a densidade conjunta de U e V se fatora em um produto de uma fução só de U e de uma função só de V. Estas funções são respectivamente a densidade de probabilidade de U e de V. Logo U e V são variáveis aleatórias independentes.

EXEMPLO 6.4.9 Sejam X e Y variáveis aleatórias independentes com distribuição gama com parâmetros α e β respectivamente. Determinar a densidade de probabilidade conjunta de (U,V) e a densidade marginal de U, para U e V definidas da seguinte maneira:

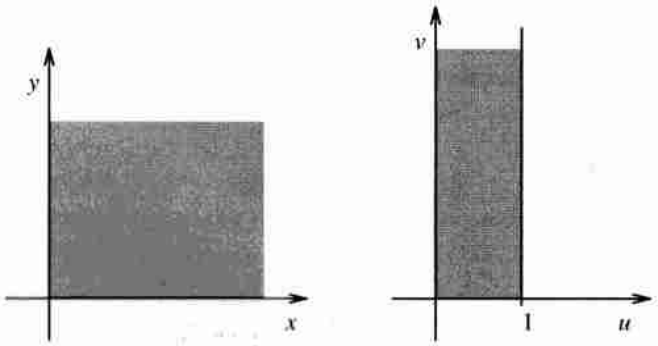
$$U = \frac{X}{X+Y} \qquad V = X+Y.$$

As densidades de probabilidade de X e Y são dadas respectivamente por:

$$f_X(x) = \frac{1}{\Gamma(\alpha)}x^{\alpha-1}e^{-x} \quad \text{e} \quad f_Y(y) = \frac{1}{\Gamma(\beta)}y^{\beta-1}e^{-y}$$

para $x \geq 0$ e $y \geq 0$, respectivamente e zero nos complementares desses intervalos.

A transformação $u = \frac{x}{x+y}$ e $v = x+y$ estabelece uma correspondência biunívoca entre o primeiro quadrante $\{(x,y), x \geq 0, y \geq 0\}$ e o conjunto $\{(u,v) : 0 \leq u \leq 1, \sigma \geq 0\}$.



A transformação inversa é:

$$\begin{aligned} x &= uv \\ y &= (1-u)v\,. \end{aligned}$$

Denominando por J o jacobiano desta transformação inversa, tem-se

$$J = \det\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \det\begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v(1-u)+uv = v.$$

De acordo com (6.28) e levando em conta que X e Y são independentes, a densidade conjunta de U e V é dada por:

$$\begin{aligned} f_{U,V}(u,v) &= f_{X,Y}(uv,(1-u)v)|J| = \\ &= \frac{1}{\Gamma(\alpha)}(uv)^{\alpha-1}e^{-uv}\frac{1}{\Gamma(\beta)}((1-u)v)^{\beta-1}e^{-(1-u)v}v = \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1}v^{\alpha+\beta-1}e^{-v} \end{aligned} \tag{6.33}$$

para $0 \leq u \leq 1$ e $v \geq 0$ e igual a zero no complementar.

A densidade marginal de U é obtida usando-se (6.4). Assim, se $f_U(u)$ designa a densidade de U, tem-se integrando $f_{U,V}(u,v)$ dada por (6.33)

$$\begin{aligned} f_U(u) &= \int_0^\infty \frac{1}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1}v^{\alpha+\beta-1}e^{-v}dv = \\ &= \frac{u^{\alpha-1}(1-u)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}\int_0^\infty v^{\alpha+\beta-1}e^{-v}dv = \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1} \end{aligned}$$

para $0 \leq u \leq 1$ e zero no complementar.

Vê-se assim que U tem distribuição beta com parâmetros α e β.

EXEMPLO 6.4.10 Seja (X,Y) uma variável aleatória com distribuição normal bivariada, cuja densidade é dada no exemplo 6.1.3. Mostramos que $\text{cov}(X,Y) = \rho\sigma_1\sigma_2$ e que portanto $\rho = \frac{\text{cov}(X,Y)}{\sigma_1\sigma_2}$ é o

coeficiente de correlação entre X e Y.

$$E(X-EX)(Y-EY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}(x-\mu_1)(y-\mu_2)$$
$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \right.\right.$$
$$\left.\left. -2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\} dxdy \ .$$

Fazendo-se a transformação de variáveis

$$\frac{x-\mu_1}{\sigma_1} = u \quad \text{e} \quad \frac{y-\mu_2}{\sigma_2} = v$$

obtém-se a transformação inversa $x = \sigma_1 u + \mu_1$ e $y = \sigma_2 v + \mu_2$, cujo jacobiano é

$$\det\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \det\begin{vmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{vmatrix} = \sigma_1\sigma_2 \ .$$

Substituindo-se na integral dupla tem-se:

$$\text{cov}(X,Y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sigma_1 u \sigma_2 v$$
$$\exp\left\{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\right\}\sigma_1\sigma_2 du dv \ .$$

Vamos calcular a integral dupla como uma integral iterada. Preliminarmente completa-se o quadrado no expoente da exponencial

$$u^2-2\rho uv+v^2 = u^2-2\rho uv+v^2+\rho^2v^2-\rho^2v^2 = (u-\rho v)^2+v^2(1-\rho^2)$$

e substituindo-se na integral têm-se:

$$\text{cov}(X,Y) = \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \ \exp\left\{-\frac{1}{2(1-\rho^2)}((u-\rho v)^2+\right.$$
$$\left.+v^2(1-\rho^2))\right\} du dv =$$

$$= \sigma_1\sigma_2\int_{-\infty}^{\infty}\frac{v}{\sqrt{2\pi}}e^{-v^2/2}\left[\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\right.$$
$$\left. u \ \exp\left\{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2\right\} du\right] dv \ .$$

Note que a integral entre colchetes é igual à esperança de uma variável aleatória normal com média ρv e variância $(1-\rho^2)$ cujo valor é ρv.

$$= \rho\sigma_1\sigma_2\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}v^2e^{-v^2/2}dv.$$

Esta integral é igual à variância de uma variável aleatória $N(0,1)$ e portanto é igual a 1.

$$= \rho\sigma_1\sigma_2$$

como queríamos demonstrar.

6.5 DISTRIBUIÇÕES CONDICIONAIS

Apresentamos nesta seção a definição e algumas propriedades da distribuição condicional de uma variável aleatória Y, dada a variável aleatória X quando X e Y são contínuas.

Sejam X e Y variáveis aleatórias contínuas cuja densidade conjunta $f_{X,Y}(x,y)$ é contínua. Suponhamos que, para todo número real x e todo $\Delta x > 0$ arbitrário, o evento $[x < X \leq x+\Delta x]$ tenha probabilidade positiva. Nestas condições está bem definida a probabilidade condicional:

$$P[Y \leq y | x < X \leq x+\Delta x] = \frac{P[Y<y, x<X\leq x+\Delta x]}{P[x<X\leq x+\Delta x]} =$$
$$= \frac{\int_x^{x+\Delta x}\left[\int_{-\infty}^{y} f(u,v)dv\right]du}{\int_x^{x+\Delta x}\left[\int_{-\infty}^{\infty} f(u,v)dv\right]du} \ .$$

Os integrandos $\int_{-\infty}^{y} f(u,v)dv$ e $\int_{-\infty}^{\infty} f(u,v)dv$ são funções contínuas de x.

Dividindo-se o numerador e o denominador por Δx, passando-se ao limite para $\Delta x \to 0$ e aplicando o teorema fundamental do cálculo, obtém-se:

$$\lim_{\Delta x \to 0} P[Y \leq y | x < X \leq x + \Delta x] = \frac{\int_{-\infty}^{y} f(x,v)dv}{\int_{-\infty}^{\infty} f(x,v)dv} = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}dv \tag{6.34}$$

onde $f_X(x) = \int_{-\infty}^{\infty} f(x,v)dv$ é a densidade marginal de X.

O leitor pode verificar que a função de y do lado direito da expressão tem as propriedades de uma função de distribuição. Esta função é definida como a função de distribuição condicional de Y dado $X = x$. Desta maneira definiremos o integrando como a densidade condicional de Y dado $X = x$.

DEFINIÇÃO 6.5.1 *A densidade condicional de Y, dado $X = x$ denotada por $f(y|x)$, é dada para cada x fixo por:*

$$f(y|x) = \frac{f(x,y)}{f_X(x)} . \tag{6.35}$$

Dessa igualdade obtém-se:

$$f(x,y) = f_X(x).f(y|x). \tag{6.36}$$

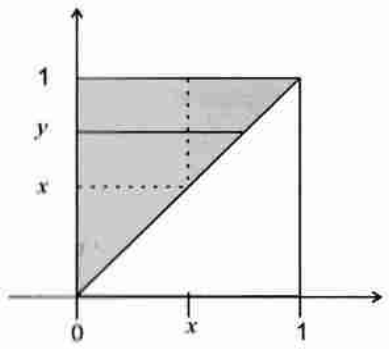
Vê-se assim que, conhecendo-se a densidade de probabilidade de X e a densidade condicional de Y dado X, obtém-se a densidade conjunta de X e Y.

EXEMPLO 6.5.1 Seja $f(x,y) = 21x^2y^3$ para $0 < x < y < 1$ e $f(x,y) = 0$ no complementar dessa região, a densidade conjunta das variáveis aleatórias X e Y. Determinar as densidades marginais $f_X(x)$, $f_Y(y)$ de X e Y, respectivamente, e a densidade condicional $f(y|x)$ de Y dado $X = x$.

A densidade é positiva na região hachurada.

Para $0 < x < 1$, tem-se de acordo com (6.4)

$$f_X(x) = \int_x^1 21x^2y^3dy = \frac{21x^2(1-x^4)}{4}$$



e $f_X(x) = 0$ fora do intervalo $(0,1)$. Para $0 < y < 1$

$$f_Y(y) = \int_0^y 21x^2y^3dx = 7y^6$$

e $f_Y(y) = 0$ fora do intervalo $(0,1)$.

De acordo com (6.35), tem-se para $0 < x < 1$ e para x fixado nesse intervalo e y tal que $x < y < 1$

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{21x^2y^3}{\frac{21x^2(1-x^4)}{4}} = \frac{4y^3}{1-x^4}.$$

EXEMPLO 6.5.2 No exemplo 6.3.1 verificamos que as variáveis aleatórias X e Y cuja densidade conjunta é $f(x,y) = xe^{-x(y+1)}$, para $0 < x < \infty$ e $0 < y < \infty$ e igual a zero no complementar, são dependentes. Determinar as densidades condicionais de Y dado X e de X dado Y.

No exemplo 6.3.1 determinamos as densidades marginais $f_X(x)$ e $f_Y(y)$ de X e Y dadas por:

$f_X(x) = e^{-x}$ para $x > 0$ e igual a zero no complementar;

$f_Y(y) = \frac{1}{(y+1)^2}$ para $y > 0$ e igual a zero no complementar.

Para determinar as densidades condicionais de Y dado X e de X dado Y vamos usar as expressões (6.35) e (6.36). Tem-se:

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{xe^{-x(y+1)}}{e^{-x}} = xe^{-xy} \text{ para } y > 0$$

e $f(y|x) = 0$ caso contrário.

Como x é constante, a distribuição condicional de Y dado $X = x$ é exponencial com parâmetro x.

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{xe^{-x(y+1)}}{(y+1)^2} \text{ para } x > 0$$

e igual a zero para $x < 0$ para cada y fixado.

EXEMPLO 6.5.3 Considere uma variável aleatória (X, Y) com distribuição normal bivariada cuja densidade conjunta $f(x, y)$ foi definida no exemplo 6.1.3. Determinar a densidade condicional de Y dado $X = x$.

Note que, para determinarmos a densidade marginal de probabilidade de X, decompusemos $f(x, y)$ da seguinte maneira:

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_1}\exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \\ &\quad \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2\right\}. \end{aligned}$$

De acordo com a expressão (6.36), como a distribuição marginal de X é $N(\mu_1, \sigma_1^2)$, vê-se que

$$\begin{aligned} f(y|x) &= \frac{1}{\sqrt{2\pi\sigma_2^2}\sqrt{1-\rho^2}} \\ &\quad \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left[y-\left(\mu_2\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\right]^2\right\} \end{aligned}$$

para x fixado e $-\infty < y < \infty$, que é uma distribuição normal com média $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$ e variância $(1 - \rho^2)\sigma_2^2$.

6.5.1 *Esperança Condicional*

Sejam X e Y variáveis contínuas com densidade conjunta $f(x, y)$. Seja $f_X(x)$ a densidade marginal de X e $f(y|x)$ a distribuição condicional de Y dado $X = x$, isto é,

$$f(y|x) = \frac{f(x,y)}{f_X(x)}.$$

DEFINIÇÃO 6.5.2 *A esperança condicional de Y dado $X = x$ é definida por*

$$E[Y|X = x) = \int_{-\infty}^{\infty} yf(y|x)dy. \tag{6.37}$$

Note que a esperança condicional de Y dado $X = x$ é a esperança de Y com relação à distribuição condicional de Y dado $X = x$. Em geral, quando falamos da esperança condicional de Y dado X, nos referimos à variável aleatória $E(Y|X)$ que para $X = x$ tem o valor dado pela expressão $E(Y|X = x)$.

EXEMPLO 6.5.4 Para as variáveis aleatórias X e Y definidas no exemplo 6.5.1, determinar a esperança condicional de Y dado X.

No exemplo 6.5.1 determinamos a densidade condicional de Y dado X, cujo valor para $0 < x < 1$ é:

$$f(y|x) = \frac{4y^3}{1-x^4} \text{ para } x < y < 1.$$

Usando a expressão (6.37) tem-se:

$$E[Y|X = x] = \int_x^1 y\frac{4y^3}{1-x^4}dy = \frac{4}{5}\frac{1-x^5}{1-x^4}.$$

EXEMPLO 6.5.5 No exemplo 6.5.3 mostramos que a densidade condicional $f(y|x)$ de Y dado $X = x$ quando (X, Y) tem distribuição normal bivariada cuja densidade foi dada no exemplo 6.5.1 é:

$$\begin{aligned} f(y|x) &= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)\sigma_2^2}\Big(y+\right. \\ &\quad \left.-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))\Big)^2\right\} \end{aligned}$$

para x fixado $-\infty < x < \infty$ e todo $-\infty < y < \infty$. Esta é a densidade de uma distribuição normal com média $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$ e variância

$\sigma_2^2(1-\rho^2)$. Desta maneira vê-se que esta é a esperança da distribuição condicional, isto é,

$$E[Y|X=x] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),$$

reescrevendo, tem-se:

$$E[Y|X=x] = \mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1 + \rho\frac{\sigma_2}{\sigma_1}x\,.$$

que é a equação de uma reta com coeficiente angular $\rho\frac{\sigma_2}{\sigma_1}$ e intercepto $\mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1$.

Esta reta é denominada regressão linear de Y em X.

6.6 VARIÁVEIS N-DIMENSIONAIS

Os conceitos introduzidos neste capítulo para um par de variáveis aleatórias se estendem de modo natural para um conjunto de n variáveis aleatórias.

Consideremos n variáveis aleatórias $X_1, X_2, \ldots, X_n$ definidas em um espaço amostral. Correspondendo à definição 6.1.1, tem-se:

DEFINIÇÃO 6.6.1 *A função $f(x_1, x_2, \ldots, x_n)$ definida para $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ não-negativa e tal que $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)dx_1dx_2\cdots dx_n$ é denominada densidade de probabilidade conjunta de $X_1, X_2, \ldots, X_n$ se para todo subconjunto $B \subset \mathbb{R}^n$ tiver-se:*

$$P[(X_1, X_2, \ldots, X_n) \in B] = \int_B \cdots \int f(x_1, x_2, \ldots, x_n)dx_1dx_2\ldots dx_n. \tag{6.38}$$

De modo análogo ao que foi feito em (6.4) e (6.5) para obter-se a densidade marginal de cada uma das variáveis aleatórias do par (X_1, X_2), pode-se obter a densidade marginal de cada uma das variáveis da n-upla $(X_1, X_2, \ldots, X_n)$

$$f_{X_i}(x_1) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)dx_2dx_3\ldots dx_n. \tag{6.39}$$

Para obter-se a densidade marginal de X_i para $l \leq i \leq n$, basta substituir X_1 por X_i e integrar-se em relação a todas as variáveis exceto x_i.

A densidade marginal do par de variáveis (X_1, X_2) é obtida integrando-se a densidade conjunta $f(x_1, x_2, \ldots, x_n)$ em relação às demais variáveis, isto é:

$$f_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)dx_3dx_4\cdots dx_n. \tag{6.40}$$

Se em (6.40) substituirmos (X_1, X_2) pelo par de variáveis (X_i, X_j) para inteiros i, j tais que $1 \leq i < j \leq n$ e integrarmos a densidade conjunta em relação a todas as variáveis, exceto x_i e x_j, obteremos a densidade marginal conjunta de X_i, X_j denotada por $f_{X_i,X_j}(x_i, x_j)$.

De modo análogo, pode-se obter a densidade marginal conjunta de quaisquer 3, 4 ou $n-1$ das variáveis.

Analogamente à definição 6.2.1 define-se a função de distribuição conjunta de n variáveis aleatórias $X_1, X_2, \ldots, X_n$.

DEFINIÇÃO 6.6.2 *A função de distribuição conjunta das variáveis aleatórias X_1, X_2, ..., X_n que denotaremos por*

$$F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n),$$

definida para $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, é dada por:

$$F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n]. \tag{6.41}$$

Para variáveis aleatórias n-dimensionais contínuas, com densidade conjunta $f(x_1, x_2, \ldots, x_n)$, a função de distribuição se exprime por uma expressão análoga a (6.8)

$$F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) =$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(u_1, u_2, \ldots, u_n)du_1du_2\cdots du_n. \tag{6.42}$$

Calculando-se sucessivamente as derivadas parciais com relação a $x_1, x_2, \ldots$ e x_n em (6.42), obtém-se a expressão correspondente a (6.9), ou seja,

$$f(x_1, x_2, \ldots, x_n) = \frac{\partial^n F(x_1, x_2, \ldots, x_n)}{\partial x_1\partial x_2\cdots\partial x_n}\,. \tag{6.43}$$

Pode-se obter também a extensão de (6.10) para n variáveis aleatórias. Sugere-se ao leitor que obtenha essa extensão para $n = 3$.

A definição 6.3.1 de independência de um par de variáveis aleatórias se estende para n variáveis aleatórias $X_1, X_2, \ldots, X_n$ da seguinte maneira:

DEFINIÇÃO 6.6.3 *As variáveis aleatórias $X_1, X_2, \ldots, X_n$ cuja densidade conjunta é $f(x_1, \ldots, x_n)$ e cujas densidades marginais são denotadas por $f_{X_1}(x_1)$, $f_{X_2}(x_2)$, $\ldots$, $f_{X_N}(x_n)$, são independentes se para todo $(x_1, x_2, \ldots, x_n)$ tal que $-\infty < x_1 < \infty$, $-\infty < x_2 < \infty$, $\ldots$, $-\infty < x_n < \infty$ tivermos:*

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n). \tag{6.44}$$

O lema 6.3.1 permanece válido, isto é, $X_1, X_2, \ldots, X_n$ são independentes se e somente se

$$F(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n). \tag{6.45}$$

Na seção 6.5 determinamos a distribuição conjunta de um par de variáveis aleatórias (Y_1, Y_2) obtidas por uma transformação de variáveis

$$\begin{aligned} Y_1 &= g_1(X_1, X_2) \\ Y_2 &= g_2(X_1, X_2) \end{aligned}$$

em que g_1 e g_2 são contínuas com derivadas parciais contínuas.

De modo análogo pode-se obter a distribuição conjunta de n variáveis aleatórias $(Y_1, Y_2, \ldots, Y_n)$ obtidas por meio da transformação

$$\begin{aligned} Y_1 &= g_1(X_1, X_2, \ldots, X_n) \\ Y_2 &= g_2(X_1, X_2, \ldots, X_n) \\ &\vdots \\ Y_n &= g_n(X_1, X_2, \ldots, X_n) \end{aligned} \tag{6.46}$$

onde supõe-se que $g_1, g_2, \ldots, g_n$ são contínuas com derivadas parciais contínuas e o jacobiano da transformação $g = (g_1 \cdots g_n)$ denotado por J_g, isto é,

$$J_g(x) = \det \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix} \tag{6.47}$$

é diferente de zero.

A densidade conjunta de $X_1, X_2, \ldots, X_n$, denotada $f_Y(y_1, \ldots, y_n)$, é de modo análogo a (6.28) fornecida por:

$$\begin{aligned} &f_Y(y_1, y_2, \ldots, y_n) \\ &= f_X\left(g_1^{-1}(y_1, y_2, \ldots, y_n), \ldots, g_n^{-1}(y_1, y_2, \ldots, y_n)\right) J_{g^{-1}}(y_1, \ldots, y_n). \end{aligned} \tag{6.48}$$

6.6.1 *Distribuições Condicionais*

Para as variáveis contínuas que são o objeto de estudo neste capítulo, a definição 6.5.1 de densidade condicional de uma variável aleatória Y dado o valor x da variável aleatória X pode ser estendida em várias direções. Pode-se considerar a variável n-dimensional cuja densidade conjunta denotamos $f(x_1, x_2, \ldots, x_n)$ e condicionar no valor de uma ou mais das variáveis.

A densidade de X_n dado $X_{n-1} = x_{n-1}, \ldots, X_2 = x_1$ é por definição dada por

$$f(x_n | x_{n-1}, \ldots, x_1) = \frac{f(x_1, x_2, \ldots, x_n)}{f_{X_1, \ldots, X_{n-1}}(x_1, \ldots, x_{n-1})} \tag{6.49}$$

onde no denominador $f_{X_1, \ldots, X_{n-1}}(x_1, \ldots, x_{n-1})$ é a densidade marginal conjunta de $X_1, \ldots, X_{n-1}$.

Podemos definir a densidade condicional de (X_{n-1}, X_n) dadas as demais, isto é, $X_1 = x_1$, $X_2 = x_2$, ..., $X_{n-2} = x_{n-2}$,

$$f(x_{n-1}, x_n | x_{n-2}, \ldots, x_1) = \frac{f(x_1, x_2, \ldots, x_n)}{f_{X_1,\ldots,X_{n-2}}(x_1, x_2, \ldots, x_{n-2})} . \quad (6.50)$$

No denominador $f_{X_1,\ldots,X_{n-2}}(x_1, x_2, \ldots, x_{n-2})$ denota a densidade marginal conjunta de $X_1, \ldots, X_{n-2}$.

Para obter a densidade condicional de um par qualquer das variáveis (X_i, X_j) dadas as demais, basta substituir x_{n-1}, x_n por (x_i, x_j), o que implica se i e j são diferentes de $n-1$ e n que x_{n-1} e x_n estarão entre as variáveis condicionantes e no lado direito estar-se-á calculando a densidade marginal conjunta das $n-2$ variáveis exceto X_i e X_j.

Estas considerações aplicam-se a um subconjunto de k variáveis das n para $3 \leq k \leq n-1$.

A esperança condicional de X_n dado $X_{n-1} = x_{n-1}$, $X_{n-2} = x_{n-2}$, ..., $X_1 = x_1$, analogamente à definição 6.5.2 é dada por

$$E(X_n | x_{n-1} = x_{n-1}, \ldots, X_1 = x_1) = \int x_n f(x_n | x_{n-1}, \ldots, x_1) dx_n .$$

6.7 EXERCÍCIOS

1. Sejam X e Y duas variáveis aleatórias contínuas cuja função densidade de probabilidade conjunta é definida como segue: $f(x,y) = cy^2$ para $0 \leq x \leq 2$ e $0 \leq y \leq 1$, e $f(x,y) = 0$ caso contrário. Determine: (a) O valor da constante c. (b) $P(X + Y > 2)$. (c) $P(X \leq 1)$. (d) $P(X \leq 1)$. (e) $P(X = 2Y)$.

2. A função densidade de probabilidade de X e Y é dada por: $f(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)$, $0 < x < 1$, $0 < y < 2$ e igual a zero caso contrário. (a) Verifique se esta é de fato uma função densidade conjunta. (b) Calcule a função densidade de X. (c) Determine $P(X > Y)$.

3. As variáveis aleatórias X e Y têm distribuição conjunta dada por: $f(x,y) = \frac{1}{8}x(x-y)$, $0 < x < 2$, $-x < y < x$, e igual a zero caso contrário. (a) Esboce o gráfico da região do plano em que $f(x,y) > 0$. (b) Verifique que $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$. (c) Encontre as densidades marginais de X e de Y.

4. A função densidade conjunta de X e Y é dada por: $f(x,y) = \frac{1}{y}e^{-(y+x/y)}$, $x > 0$, $y > 0$ e $f(x,y) = 0$, caso contrário. Determine $E(X)$, $E(Y)$ e $\mathrm{Cov}(X,Y)$.

5. Calcule a covariância entre as variáveis aleatórias X e Y se a função densidade conjunta de X e Y é dada por: $f(x,y) = \frac{2e^{-2x}}{x}$, $x \geq 0$, $0 \leq y \leq x$, e $f(x,y) = 0$ caso contrário.

6. Suponha que a função densidade de probabilidade conjunta de X e Y é especificada como segue: $f(x,y) = cx^2y$, $x^2 \leq y \leq 1$, e $f(x,y) = 0$ caso contrário. (a) Esboce o gráfico da região do plano em que $f(x,y) > 0$. (b) Determine o valor da constante c. (c) Determine $P(X \geq Y)$.

7. Paulo e Pedro decidem se encontrar em um certo local. Se cada um deles chega a esse local independentemente em um tempo uniformemente distribuído entre meio-dia e uma hora da tarde, qual a probabilidade de que o primeiro a chegar espere mais do que 10 minutos?

8. Se X é uniformemente distribuída no intervalo (0,1) e Y é exponencialmente distribuída com parâmetro $\lambda = 1$, determine a distribuição de: (a) $Z = X + Y$; (b) $W = X/Y$. Suponha que X e Y são variáveis independentes.

9. Sejam X e Y variáveis aleatórias independentes com distribuição comum exponencial de parâmetro λ. Calcule a função densidade de probabilidade de: (a) $Z = X + Y$; (b) $W = \frac{X}{X+Y}$.

10. Sejam X e Y variáveis aleatórias independentes com distribuição uniforme em $[\theta - \frac{1}{2}; \theta + \frac{1}{2}]$, onde $\theta \in \mathbb{R}$. Mostre que a distribuição de $Z = X - Y$ não depende de θ.

11. (a) Mostre que se X e Y são variáveis aleatórias independentes com distribuição gama com parâmetros (α, λ) e (β, λ), respectivamente, então $X + Y$ é uma variável aleatória gama com parâmetro $(\alpha + \beta, \lambda)$.

(b) Utilizando-se do resultado em (a) e o princípio da indução, mostre que se X_i, $i = 1, \ldots, n$, são variáveis aleatórias independentes com distribuição gama com parâmetros (α_i, λ), $i = 1, \ldots, n$, então $\sum_{i=1}^{n} X_i$ é uma variável aleatória gama com parâmetros $(\sum_{i=1}^{n} \alpha_i, \lambda)$.

(c) Se $X_1, \ldots, X_n$ são n variáveis aleatórias independentes com distribuição exponencial com parâmetro λ, qual a distribuição de $S = \sum_{i=1}^{n} X_i$?

12. Mostre que se X_i, $i = 1, \ldots, n$, são variáveis aleatórias independentes com distribuição normal com parâmetros μ_i e σ_i, $i = 1, \ldots, n$, respectivamente, então $\sum_{i=1}^{n} X_i$ é normalmente distribuída com parâmetros $\sum_{i=1}^{n} \mu_i$ e $\sum_{i=1}^{n} \sigma_i^2$.

13. Quando uma corrente elétrica I (medida em ampères) passa através de uma resistência R (medida em ohms), a potência gerada é dada por $W = RI^2$ (medida em watts). Suponha que I e R são variáveis aleatórias independentes com densidades $f_I(x) = 6x(1-x)$ e $f_R(x) = 2x$ para $0 \leq x \leq 1$ e zero no complementar. Determine: (a) A função densidade de W. (b) A esperança de W. (c) A probabilidade de que a potência gerada, sob estas condições, seja maior que 50%.

14. A intensidade luminosa em um dado ponto é dada pela expressão $I = \frac{C}{D^2}$, onde C é o poder luminoso da fonte de luz e D é a distância desse ponto à fonte de luz. Suponha que C seja uniformemente distribuída em (1, 2) e que D seja uma variável aleatória com densidade $f(d) = e^{-d}$ para $d \geq 0$. Determine a densidade de I (admita independência entre c e D).

15. Suponha que as dimensões A e B de uma chapa retangular de aço produzida por uma indústria sejam variáveis aleatórias independentes com as seguintes funções densidades de probabilidade:

$$f_A(x) = \begin{cases} x-1 & , 1 \leq x \leq 2\,, \\ 3-x & , 2 < x \leq 3\,, \\ 0 & , \text{caso contrário;} \end{cases}$$

$$f_B(y) = \begin{cases} \frac{1}{2} & , 2 \leq y \leq 4, \\ 0 & , \text{caso contrário.} \end{cases}$$

Determine: (a) A função densidade de probabilidade da área da chapa S. (b) A esperança de S. (c) A probabilidade de que a área de uma chapa produzida por esta indústria seja maior que 10 unidades de área.

16. Suponha que X_1, X_2, X_3 representam as componentes da velocidade de uma molécula de gás. Suponha que X_1, X_2 e X_3 são independentes, cada uma com densidade normal média 0 e variância σ^2, $\sigma > 0$. Diz-se em física que a magnitude da velocidade $Y = (X_1^2 + X_2^2 + X_3^2)^{1/2}$ tem distribuição de Maxwell. Determine a densidade da magnitude da velocidade.

17. Se X e Y são variáveis aleatórias independentes com funções densidade iguais a $f(x) = e^{-x}$ para $x > 0$, obtenha a densidade conjunta de $X + Y$ e $3X + 2Y$. Neste caso, $X + Y$ e $3X + 2Y$ são independentes?

18. Sejam X e Y independentes identicamente, distribuídas com distribuição normal padrão. Ache a distribuição de (U, V) onde $U = X^2 + Y^2$ e $V = \frac{X}{Y}$. Neste caso, U e V são independentes?

19. Seja (X, Y) um ponto aleatório no plano (no $\mathbb{R}^2$). Supondo que as coordenadas cartesianas do ponto X e Y são variáveis aleatórias independentes com distribuição normal padrão, determine a densidade conjunta das coordenadas polares deste ponto.

    Lembremos que as coordenadas polares de um ponto podem ser expressas em função de suas coordenadas cartesianas do seguinte modo: $R = \sqrt{X^2 + Y^2}$, $\theta = \text{arctg}\frac{Y}{X}$.

20. Se X e Y são variáveis aleatórias gama independentes com parâmetros (α, λ) e (β, λ), respectivamente. Calcule a densidade conjunta de $U = X + Y$ e $V = \frac{X}{X+Y}$. Neste caso, U e V são independentes?

21. Considere novamente os exercícios de 1 a 6. Verifique em quais casos X e Y são independentes.

22. Se U é uniforme no intervalo $(0, 2\pi)$ e Z, independente de U, é exponencial com parâmetro 1, mostre que X e Y definidos por $X = \sqrt{2Z}\cos U$ e $Y = \sqrt{2Z}\text{sen}U$ são variáveis aleatórias independentes, cada uma com distribuição normal padrão.

    (Este resultado é de grande aplicação na simulação de variáveis aleatórias normais e independentes, pois fornece um procedimento para gerar pares de observações da distribuição normal padrão a partir de números "pseudo-aleatórios".)

23. Verifique, em cada caso, se X e Y são independentes.

    (a) $f(x,y) = \begin{cases} 2xe^{-y} & , 0 \le x \le 1 \text{ e } y \ge 0, \\ 0 & , \text{caso contrário.} \end{cases}$

    (b) $f(x,y) = \begin{cases} 24xy & , x \ge 0, y \ge 0, x + y \le 1, \\ 0 & , \text{caso contrário.} \end{cases}$

    (c) $f(x,y) = \begin{cases} \frac{15}{4}x^2 & , 0 \le y \le 1 - x^2, \\ 0 & , \text{caso contrário.} \end{cases}$

24. Se X e Y são variáveis aleatórias independentes e uniformemente distribuídas no intervalo $(0,1)$, determine a densidade conjunta de U e V quando: (a) $U = X + Y$ e $V = \dfrac{X}{Y}$. (b) $U = X$, $V = \dfrac{X}{Y}$. (c) $U = X + Y$, $V = \dfrac{X}{X+Y}$.

25. Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias com densidade conjunta f e considere as variáveis $Y_1, \ldots, Y_n$ definidas por: $Y_i = \sum_{j=1}^{n} a_{ij}X_j$, $i = 1, \ldots, n$, onde a matriz $A = (a_{ij})$ tem determinante diferente de zero.

    (a) Verifique que $Y_1, \ldots, Y_n$ têm densidade conjunta $f_{Y_1,\ldots,Y_n}$ dada por:

    $$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = \frac{1}{|\det A|} f(x_1, \ldots, x_n).$$

    (b) Determine a distribuição conjunta de $X_1, X_2, \ldots, X_n$, sabendo que as variáveis $X_1, X_2 - X_1, \ldots, X_n - X_{n-1}$ são independentes e têm todas a mesma distribuição $N(0,1)$.

26. Sejam X e Y variáveis aleatórias com distribuição normal bivariada cuja função densidade de probabilidade conjunta de X e Y é dada por:

    $$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}} \exp\Big\{-\frac{1}{2(1-\rho^2)}\Big[\Big(\frac{X-\mu_x}{\sigma_x}\Big)^2 + \Big(\frac{y-\mu_y}{\sigma_y}\Big)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\Big]\Big\}.$$

    (a) Mostre que ρ é igual ao coeficiente de correlação de X e Y.
    (b) Para que valores de ρ, X e Y são independentes?

27. As idades de pais em um certo hospital podem ser aproximadas por uma distribuição normal bivariada com parâmetros $\mu_M = 28.4$, $\sigma_M = 6.8$, $\mu_P = 31,6$, $\sigma_P = 7.4$ e $\rho = 0,82$ (os parâmetros com índice M referem-se à idade das mães grávidas e os parâmetros com índice P referem-se à idade do pai). Determine: (a) A proporção de mães grávidas com mais de 30 anos de idade. (b) A proporção de casais em que ambos, o pai e a mãe, têm menos de 25 anos de idade.

28. (a) Mostre a desigualdade de Cauchy-Schwarz, dada por:

    $$(E[XY])^2 \le E[X^2]E[Y^2] .$$

    Dica: Observe que $E[(tX + Y)^2] \ge 0$, $\forall t \in \mathbb{R}$.

    (b) Se X e Y são variáveis aleatórias quaisquer:

    $$F_{X,Y}(x,y) \le \sqrt{F_X(x).F_Y(y)} .$$

    Dica: Defina variáveis indicadoras e utilize o resultado em (a).

29. Suponha que n variáveis $X_1, \ldots, X_n$ formam uma amostra aleatória de uma distribuição contínua para a qual a função densidade de probabilidade é f e a função de distribuição é F. Sejam $X_{(1)}, \ldots, X_{(n)}$ as estatísticas de ordem dessa amostra. Determine: (a) A densidade conjunta de $X_{(1)}$ e $X_{(n)}$. (b) A densidade de $A = X_{(n)} - X_{(1)}$ (a variável A definida desta forma é denominada a amplitude da amostra).

30. Considere novamente as condições do exercício 29.

(a) Determine $E[F(X_{(1)})]$ e $E[F(X_{(n)})]$.

Supondo agora que

$$f(x) = 1, \ 0 \leq x \leq 1, \qquad \text{e} \qquad F(x) = \begin{cases} 0 & , x < 0, \\ x & , 0 \leq x \leq 1, \\ 1 & , x > 1, \end{cases}$$

determine:

(b) A densidade de $A = X_{(n)} - X_{(1)}$.

(c) $E(X_{(1)})$, $E(X_{(n)})$ e a correlação entre $X_{(1)}$ e $X_{(n)}$.

31. Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes com distribuição exponencial de parâmetros $\lambda_1, \ldots, \lambda_n$, respectivamente:

(a) Determine a distribuição de $X_{(1)} = \min_{1 \leq i \leq n} X_i$.

(b) Mostre que $P(X_k = \min_{1 \leq i \leq n} X_i) = \dfrac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}$, para $k = 1, \ldots, n$.

Dica: Observe que X_k e $\min_{i \neq k} X_i$ são independentes e considere o evento $\{X_k \leq \min_{i \neq k} X_i\}$.

32. Mostre que a mediana de uma amostra aleatória de tamanho $2n+1$ de uma distribuição uniforme em $(0,1)$ tem distribuição beta com parâmetros $(n+1, n+1)$.

33. Suponha que $X_1, \ldots, X_n$ formam uma amostra aleatória de tamanho n de uma distribuição uniforme no intervalo $(0,1)$ e seja $X_{(n)} = \max_{1 \leq i \leq n} X_i$. Determine o menor valor de n tal que $P\{X_{(n)} \geq 0,99\} \geq 0,95$.

34. Considere novamente os exercícios de 1 a 6. Determine:
(a) $f(y|x)$ e $f(x|y)$.
(b) $E(X|Y = y)$ e $E(Y|X = x)$.

35. Suponha que a densidade conjunta de X e Y é dada por: $f(x,y) = \dfrac{e^{-y}}{y}$, $0 < x < y$ e $y \geq 0$, $f(x,y) = 0$ caso contrário.
(a) Calcule $E(X|Y = y)$. (b) Determine $\text{Cov}(X,Y)$.

36. Selecione um ponto X com distribuição uniforme no intervalo $(0,1)$. Após o valor $X = x$ ter sido observado, um ponto Y é escolhido uniformemente no intervalo $(x,1)$. Determine a densidade de probabilidade marginal de Y.

37. Seja (X,Y) um vetor aleatório bidimensional. Suponha que X tem distribuição exponencial com parâmetro $\frac{1}{2}$ e que a distribuição condicional de Y, dado que $X = x$, é uniforme em $[0, x^2]$. (a) Qual é a distribuição de $\frac{Y}{X^2}$? (b) Determine $E(X)$, $E(Y)$ e $\text{Cov}(X,Y)$.

38. Seja W uma variável aleatória com distribuição gama com parâmetros α e β. Admita que, condicional a $W = w$, $X_1, \ldots, X_n$ são variáveis aleatórias independentes com distribuição exponencial de parâmetro w. Mostre que a distribuição condicional de W dado que $X_1 = x_1, \ldots, X_n = x_n$ é gama com parâmetros $\alpha + n$ e $\beta + \sum_{i=1}^n x_i$.

39. Seja Y uma variável aleatória com distribuição beta com parâmetros α e β. Admita que, condicional a $Y = y$, $X_1, \ldots, X_n$ são variáveis aleatórias independentes com distribuição Bernoulli de parâmetro y. Verifique que a distribuição condicional de Y dado que $X_1 = x_1, \ldots, X_n = x_n$ é beta com parâmetros $\alpha + \sum_{i=1}^n x_i$ e $\beta + n$.

40. Seja X uma variável aleatória com densidade $f(x) = 2x$, $0 \leq x \leq 1$ e suponha que a distribuição condicional de uma variável aleatória Y, dado $X = x$, é exponencial com parâmetro λx. Calcule $E(Y)$ e $\text{Var}(Y)$.

41. Suponha que X, dado que $Y = y$, tem distribuição de Poisson com parâmetro y. Se Y é uma variável aleatória com distribuição exponencial de parâmetro 1, determine a distribuição de X.

42. Suponha que a função densidade de probabilidade da variável aleatória X é dada por: $f(x) = \dfrac{1}{n!} x^n e^{-x}$, $x > 0$, e $f(x,y) = 0$ caso contrário. Suponha também que, dado $X = x$ $(x > 0)$, as n variáveis aleatórias $Y_1, Y_2, \ldots, Y_n$ são independentes e a

função densidade condicional de cada uma delas é como segue: $g(y|x) = \frac{1}{x}$, $0 < y < x$, e $g(y|x) = 0$ caso contrário. Determine: (a) A função densidade conjunta de $Y_1, \ldots, Y_n$. (b) A função densidade condicional de X dado $Y_1 = y_1, \ldots, Y_n = y_n$.

7

Tipos de Convergência e Teoremas Limite

Ao longo do desenvolvimento deste texto, consideramos modelos probabilísticos cuja construção foi motivada pela tentativa de descrever fenômenos naturais ou o comportamento de sistemas complexos. Como exemplos, tivemos o comportamento de uma central telefônica, o processo de envelhecimento de materiais, progressões demográficas, atitudes coletivas eleitorais e outros.

Neste último exemplo, vem à discussão a pesquisa de opinião. Este tipo de processo vale-se de amostragem, geralmente estratificada, onde cada elemento, de cada extrato da população pesquisada, geralmente possui a mesma chance de ser entrevistado. Isto tudo é uma instância de um modelo recorrente em Inferência Estatística e Simulação de Sistemas: a experimentação por inúmeras repetições do mesmo modelo probabilístico. A descrição formal desta situação se dá, muitas vezes, pela consideração de seqüências de variáveis aleatórias independentes. Isto motiva o estudo do comportamento assintótico das seqüências de variáveis aleatórias ou de funções destas.

Um exemplo das questões de que trata este Capítulo já foi mencionado em termos breves no Capítulo 4. Lá introduzimos a distribuição de Poisson através da aproximação de distribuições binomiais, dependendo a qualidade da aproximação de quão "próximo" o produto np está de λ, o parâmetro da distribuição de Poisson. Ora, o que justifica esta aproximação é a convergência de seqüências de variáveis binomiais para uma variável aleatória com distribuição de Poisson. É este o principal assunto deste capítulo: modos de convergência de seqüências de variáveis aleatórias.

Certos procedimentos estatísticos são válidos quando o tamanho da amostra é grande. Deste modo, é importante o estudo das distribuições de variáveis aleatórias definidas para amostras grandes.

Introduziremos neste capítulo noções de convergência para seqüências de variáveis aleatórias e apresentaremos a lei dos grandes números e o teorema do limite central.

7.1 LEI DOS GRANDES NÚMEROS E TEOREMAS DO LIMITE CENTRAL

DEFINIÇÃO 7.1.1 *A seqüência de variáveis aleatórias X_n converge em probabilidade para a variável aleatória X se para todo $\epsilon > 0$ tem-se*

$$\lim_{n\to\infty} P[|X_n - X| \geq \epsilon] = 0. \tag{7.1}$$

Indicaremos que X_n converge em probabilidade para X com a notação $X_n \xrightarrow{p} X$.

Considere uma seqüência numérica $\{a_n\}$. Dizer-se que $a_n \to a$ é dizer que dado $\varepsilon > 0$ pode-se determinar n_0 tal que para todo $n \geq n_0$ tem-se: $|a_n - a| \leq \varepsilon$. Quando se realiza uma seqüência de experimentos aleatórios e se associa uma variável aleatória X_n, $n \geq 1$, a cada um deles, então para cada $\omega = (\omega_1, \omega_2, \ldots)$ que representa uma realização fica definida a seqüência $X_n(\omega)$, $n \geq 1$.

Aqui pode acontecer que para um resultado ω do experimento $|X_n(\omega) - X(\omega)| > \varepsilon$ para algum $n > n_0$. O que $X_n \xrightarrow{p} X$ garante é que a probabilidade de esse evento ocorrer é muito pequena e tende para zero. Assim, $X_n \xrightarrow{p} X$ significa que se n for suficientemente grande, então a probabilidade que X_n difira de X por mais que ϵ é muito pequena.

DESIGUALDADE DE TCHEBICHEV *Seja X uma variável aleatória com variância finita. Para todo $\epsilon > 0$ tem-se:*

$$P[|X - EX| \geq \epsilon] \leq \frac{\mathrm{Var}X}{\epsilon^2}. \tag{7.2}$$

DEMONSTRAÇÃO Vamos demonstrar a desigualdade supondo que X é uma variável contínua com densidade de probabilidade $f(x)$.

$$\begin{aligned}
\mathrm{Var}X &= \int_{-\infty}^{\infty} (X - EX)^2 f(x)dx = \\
&= \int_{\{x:|x-EX|\geq\epsilon\}} (X - EX)^2 f(x)dx + \\
&\quad + \int_{\{x:|x-EX|<\epsilon\}} (x - EX)^2 f(x)dx.
\end{aligned}$$

Como o integrando na segunda integral do lado direito é não-negativo e na primeira integral no lado direito $(x-EX)^2 \geq \epsilon^2$, substituindo-se temos:

$$\mathrm{Var}X \geq \epsilon^2 \int_{\{X:|x-EX|\geq\epsilon\}} f(x)dx = \epsilon^2 P[|X - EX| \geq \epsilon]$$

o que verifica a desigualdade.

LEI FRACA DOS GRANDES NÚMEROS *Seja $X_1, X_2, \ldots$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com média μ e variância σ^2. Seja $S_n = \sum_{i=1}^n X_i$. Tem-se*

$$\frac{S_n}{n} \xrightarrow{p} \mu.$$

DEMONSTRAÇÃO Aplicaremos a desigualdade de Tchebichev à variável aleatória $\frac{S_n}{n}$. Como $E(\frac{S_n}{n}) = \mu$ e $\mathrm{Var}(\frac{S_n}{n}) = \frac{\sigma^2}{n}$, a desigualdade (7.2) se transforma em

$$P\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] \leq \frac{\sigma^2}{n\epsilon^2}. \tag{7.3}$$

Daí segue tomando o limite para $n \to \infty$ que

$$\lim_{n\to\infty} P\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] = 0 .$$

Um caso particular importante é aquele em que $X_i = 1$ (sucesso) com probabilidade p e $X_i = 0$ (falha) com probabilidade $(1-p)$, para $i = 1, 2, \ldots$, que corresponde a uma seqüência de ensaios de Bernoulli. $\frac{S_n}{n}$ representa a freqüência relativa de sucessos em n ensaios. Neste caso, na fórmula (7.3) temos $\mu = p$ e $\sigma^2 = p(1-p)$. Assim,

$$P\left[\left|\frac{S_n}{n} - p\right| \geq \epsilon\right] \leq \frac{p(1-p)}{n\epsilon^2} . \tag{7.4}$$

A expressão do lado direito tende a zero para $n \to \infty$.

A lei fraca diz que $\frac{S_n}{n}$, a freqüência relativa de sucessos, converge em probabilidade para p.

EXEMPLO 7.1.1 Seja $0 < \delta < 1$ arbitrário. Podemos utilizar a fórmula (7.4) para determinar um n a partir do qual

$$P\left[\left|\frac{S_n}{n} - p\right| \geq \epsilon\right] < \delta.$$

Comparando-se com a fórmula (7.4), vemos que se $\frac{p(1-p)}{n\epsilon^2} < \delta$ esta condição será satisfeita. Resolvendo esta desigualdade em n temos

$$n \geq \frac{p(1-p)}{\delta\epsilon^2} . \tag{7.5}$$

Suponhamos que $p = 0,3$, $\delta = 0,01$ e $\epsilon = 0,05$. Substituindo esses valores temos $n \geq 8.400$.

Suponhamos agora que p fosse desconhecido e façamos a mesma pergunta com $\delta = 0,01$ e $\epsilon = 0,05$. Notemos que a função $p(1-p)$ está definida no intervalo $[0,1]$ e assume um valor máximo de $\frac{1}{4}$ para $p = \frac{1}{2}$. Logo, $p(1-p) \leq \frac{1}{4}$. Se substituirmos $p(1-p)$ por $\frac{1}{4}$ em (7.5), nós obtemos $n \geq 10.000$.

DEFINIÇÃO 7.1.2 *A seqüência de variáveis aleatórias X_n converge em distribuição para a variável aleatória X se $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$, para todo x tal que F_X seja contínua em x. Notação: $X_n \xrightarrow{d} X$.*

Sejam a e b dois números reais com $a < b$. Suponhamos que a e b são pontos onde F_X é contínua. Pela fórmula do Capítulo 3, temos:

$$P[a < X_n \leq b] = F_{X_n}(b) - F_{X_n}(a) \tag{7.6}$$

se $X_n \xrightarrow{d} X$, então $F_{X_n}(a) \to F_X(a)$ e $F_{X_n}(b) \to F_X(b)$. Passando ao limite em (7.6), temos:

$$\lim_{n\to\infty} P[a < X_n \leq b] = F_X(b) - F_X(a) = P[a < X \leq b] .$$

Enunciaremos, a seguir, um teorema cuja prova envolve conhecimentos que não são admitidos e que fogem aos objetivos deste livro, porém, cujo resultado será utilizado na demonstração do teorema do limite central.

TEOREMA 7.1.1 *Seja $X_1, X_2, \ldots$ uma seqüência de variáveis aleatórias cujas funções geradoras de momentos $\phi_{X_1}, \phi_{X_2}, \ldots$ existem. X_n converge em distribuição para uma variável aleatória X se, e somente se, $\lim_{n\to\infty} \phi_{X_n}(t) \to \phi_X(t)$, para todo $t \in \mathbb{R}$, onde ϕ_X é a função geradora de momentos de X.*

TEOREMA 7.1.2 (Teorema do Limite Central) *Seja $X_1, X_2, \ldots$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com média μ e variância σ^2. Seja $S_n = X_1 + X_2 + \cdots + X_n$ então $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ converge em distribuição para Z que tem distribuição normal com média zero e variância 1.*

DEMONSTRAÇÃO Vamos mostrar que a função geradora de momentos de $Z_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}$ converge para $e^{t^2/2}$ que é a função geradora de momentos da distribuição normal com média zero e variância 1. A demonstração estará concluída utilizando-se o teorema citado acima. Seja

$$\begin{aligned}\phi_{Z_n}(t) &= E(e^{tZ_n}) = E\left[\exp\left\{t\frac{\sum_{i=1}^n (X_i - \mu)}{\sigma\sqrt{n}}\right\}\right] = \\ &= E\left[\prod_{i=1}^n \exp\left\{\frac{t}{\sigma\sqrt{n}}(X_i - \mu)\right\}\right] =\end{aligned}$$

$$= \prod_{i=1}^{n} E\left[\exp\left\{\frac{t}{\sigma\sqrt{n}}(X_i - \mu)\right\}\right] .$$

Esta última igualdade decorre da independência das variáveis aleatórias $X_1, X_2, \ldots, X_n$.

Vamos reescrever esta última igualdade

$$\phi_{Z_n}(t) = \prod_{i=1}^{n} E\left[\exp\left\{\frac{t}{\sigma\sqrt{n}}(X_i - \mu)\right\}\right] . \tag{7.7}$$

Designamos por $\phi(t) = E\left[e^{t(X_1-\mu)}\right]$ a função geradora de momentos de $X_1 - \mu$. Como $X_1 - \mu$ tem média zero e o segundo momento de $X_1 - \mu$ é a variância de X_1, temos:

$$\phi'(0) = 0; \qquad \phi''(0) = \sigma^2 . \tag{7.8}$$

Como $X_1 - \mu$, $X_2 - \mu$, ..., $X_n - \mu$ são identicamente distribuídas, elas têm a mesma função geradora de momentos ϕ. No lado direito de (7.7) temos ϕ calculado no ponto $\frac{t}{\sigma\sqrt{n}}$. Fazendo essa substituição (7.7) se transforma em

$$\phi_{Z_n}(t) = \prod_{i=1}^{n} \phi\left(\frac{t}{\sigma\sqrt{n}}\right) = \left[\phi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n . \tag{7.9}$$

Pela fórmula de Taylor, existe s, $0 < s < t$ tal que

$$\phi(t) = \phi(0) + \phi'(0)t + \frac{\phi''(s)t^2}{2} = \frac{1 + \phi''(s)t^2}{2} .$$

Somando e subtraindo $\frac{\sigma^2 t^2}{2}$ na expressão acima obtemos

$$\phi(t) = 1 + \frac{\sigma^2 t^2}{2} + (\phi''(s) - \sigma^2)\frac{t^2}{2} . \tag{7.10}$$

Pondo $\frac{t\sigma}{\sqrt{n}}$ em (7.10) e substituindo em (7.9), vem

$$\phi_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \frac{(\phi''(s) - \sigma^2)t^2}{2\sigma^2 n}\right]^n , \tag{7.11}$$

onde $0 < s < \frac{t}{\sigma\sqrt{n}}$ e $-h\sigma\sqrt{n} < t < h\sigma\sqrt{n}$.

Para $n \to \infty$ segue-se que $s \to 0$ e como ϕ'' é contínua na origem, temos $\lim_{s\to 0}(\phi''(s) - \sigma^2) = 0$.

Usando o seguinte resultado do cálculo, $\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^a$ quando $a_n \to a$, concluímos a partir de (7.11) que

$$\lim_{n\to\infty} \phi_{Z_n}(t) = e^{t^2/2},$$

como queríamos demonstrar.

Observações: a) No caso particular em que cada $X_i = 1$ com probabilidade p e 0 com probabilidade $(1-p)$, $S_n = \sum_{i=1}^{n} X_i$ é o número de sucessos (ocorrência de 1) em n ensaios Bernoulli. Sabemos que S_n tem distribuição binomial com média np e variância $np(1-p)$. O teorema do limite central diz que $\frac{S_n - np}{\sqrt{np(1-p)}}$ tem, para n suficientemente grande, distribuição aproximadamente $N(0,1)$.

b) $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma/\sqrt{n}}$ pode ser reescrito, dividindo-se o denominador por n, da seguinte maneira:

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \tag{7.12}$$

e o teorema do limite central diz que para n suficientemente grande a média amostral tem distribuição aproximadamente $N(\mu, \frac{\sigma^2}{n})$ ou, o que é equivalente, (7.12) tem distribuição aproximadamente $N(0,1)$.

EXEMPLO 7.1.2 Sabe-se que 80% das peças produzidas por uma indústria passam por três testes de qualidade. Uma amostra de 200 peças é escolhida ao acaso da linha de produção. Qual é a probabilidade de que o número de peças na amostra que passam pelos três testes de qualidade esteja compreendido entre 154 e 170, inclusive os extremos?

Seja S o número de peças na amostra que passa pelos três testes. S tem distribuição binomial com $n = 200$ e $p = 0,8$. $ES = 200 \times 0,8 = 160$, $\mathrm{Var}S = 200 \times 0,8 \times 0,2 = 32$. O desvio padrão $\sigma_S = \sqrt{32} = 5,66$. Utilizando-se o teorema do limite central, temos:

$$\begin{aligned} P[154 \le S \le 170] &= P\left[\frac{154 - 160}{5,66} \le \frac{S - 160}{5,66} \le \frac{170 - 160}{5,66}\right] = \\ &= P[1,06 \le Z \le 1,33] = 0,76. \end{aligned}$$

EXEMPLO 7.1.3 Suponhamos que $X_1, X_2, \ldots$ são variáveis aleatórias independentes, com distribuição de Poisson com parâmetro λ. Seja $S_n = X_1 + X_2 + \cdots + X_n$. S_n tem distribuição de Poisson com parâmetro $n\lambda$. (Verifique usando o método da função geradora de momentos.) $ES_n = n\lambda$ e $\sigma_{S_n} = \sqrt{n}\lambda$. Aplicando-se o teorema do limite central tem-se:

$$\lim_{n\to\infty} P\left[\frac{S_n - n\lambda}{\sqrt{n}\lambda} \leq x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty} e^{-t^2/2} dt \ .$$

Isto é, para n suficientemente grande, a distribuição de Poisson com parâmetro $n\lambda$ é aproximadamente normal.

EXEMPLO 7.1.4 Um dado é lançado 2.500 vezes. Calcular a probabilidade de que a soma dos pontos obtidos seja menor que 8.850 usando a aproximação normal.

Os sucessivos lançamentos produzem variáveis aleatórias independentes com a mesma distribuição que X_1 tal que $P[X_1 = i] = \frac{1}{6}$, para $i = 1, 2, \ldots, 6$. Temos $EX_1 = 3,5$ e $\sigma(X_1) = 1,71$.

$$\begin{aligned} P[S_n \leq 8.850] &= P\left[\frac{S_n - 8.750}{85,5} \leq \frac{8.850 - 8.750}{85,5}\right] = \\ &= P[Z \leq 1,17] = 0,879, \end{aligned}$$

onde Z tem distribuição $N(0,1)$.

7.2 EXERCÍCIOS

1. Seja X uma variável aleatória tal que $P(X \geq 0) = 1$ e $P(X \geq 14) = \frac{1}{7}$. Mostre que $E(X) \geq 2$.

2. Suponha que X tem distribuição de Poisson de parâmetro λ. Use a desigualdade de Tchebyschev para verificar que:
   (a) $P(X \leq \frac{\lambda}{2}) \leq \frac{4}{\lambda}$. (b) $P(X \geq 2\lambda) \leq \frac{1}{\lambda}$.

3. Seja X uma variável aleatória inteira não-negativa com função geradora de momentos $\phi_X(t) = E(e^{tX})$, finita para todo t. Seja x_0 um número inteiro positivo. Mostre que: $P(X \geq x_0) \leq \frac{\phi_X(t)}{e^{tx_0}}$, $t > 0$ (desigualdade de Chernoff).

4. Considere X como no exercício 2. Mostre que $P(X \geq 2\lambda) \leq (\frac{e}{4})^{\lambda}$.

   Dica: Utilize o resultado do exercício 3 tomando t^* tal que

   $$\frac{\phi_X(t^*)}{e^{t^* \cdot 2\lambda}} = \min_{t>0} \frac{\phi_X(t)}{e^{t.2\lambda}}.$$

5. Seja X uma variável aleatória não-negativa tal que $\mu = E(X)$. Mostre que $P(X \geq 2\mu) \leq \frac{1}{2}$ (interprete o resultado).

6. Se X é uma variável aleatória com média μ e variância σ^2, mostre que
   $P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$, $k \in \mathbb{N}$ (interprete o resultado).

7. Seja Y uma variável aleatória com densidade $f(y) = \frac{1}{m!} y^m e^{-y}$, $y \geq 0$, $m \in \mathbb{N}$. Prove que $P(0 \leq Y \leq 2(m+1)) \geq \frac{m}{m+1}$.

8. Sabe-se que o número de itens produzidos em uma fábrica durante uma semana (X) é uma variável aleatória com média 50. (a) O que se pode afirmar sobre a probabilidade de X exceder 100 unidades? (b) Se a variância de X é 50, o que se pode afirmar sobre a probabilidade de X exceder 100 unidades? Compare os resultados em (a) e (b).

9. Sejam $X_1, X_2, \ldots$ variáveis aleatórias independentes e identicamente distribuídas (v.a.i.i.d.) com distribuição uniforme no intervalo (0,1). Sejam $Y_n = \max\{X_1, \ldots, X_n\}$ e $Z_n = \min\{X_1, \ldots, X_n\}$, para $n = 1, 2, \ldots$ Prove que: (a) $Y_n \xrightarrow{P} 1$; (b) $Z_n \xrightarrow{P} 0$.

10. Sejam $X_1, X_2, \ldots$ variáveis aleatórias independentes e identicamente distribuídas (v.a.i.i.d.), $U(0,1)$, e seja $Y_n = \min\{X_1, \ldots, X_n\}$, $n = 1, 2, \ldots$ Mostre que $nY_n \xrightarrow{d} Y$, onde Y tem distribuição exponencial com parâmetro λ.

11. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. Poisson (λ). Calcule o limite em probabilidade da seqüência $\{Y_n, n \geq 1\}$ onde $Y_n = \frac{X_1^2 + \cdots + X_n^2}{n}$. Sugestão: utilize a lei fraca dos grandes números.

12. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. com função de distribuição F que possui uma densidade. Seja $M_n = \max\{X_1, X_2, \ldots, X_n\}$ e considere a seqüência $\{Y_n, n \geq 1\}$ onde $Y_n = n[1 - F(M_n)]$. Ache o limite em distribuição de Y_n.

13. Seja $\{X_n, n \geq 1\}$ uma seqüência de variáveis aleatórias com distribuições geométricas com parâmetros $\frac{\lambda}{n}$, $0 < \lambda < n$. Seja $Y_n = \frac{X_n}{n}$, $\forall n \geq 1$. Prove que $Z_n \xrightarrow{d} Z$, onde Z tem distribuição $\exp(\lambda)$.

14. Considere novamente a situação do exercício 13. Verifique que Z_n não converge em probabilidade para Z ($Z_n \xrightarrow{P} Z$).

15. Mostre que $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$.
Note que a recíproca não é verdadeira! (Veja exercícios 13 e 14.)

16. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. com função densidade de probabilidade dada por: $f(x) = e^{-x+\theta}$, $x \geq \theta$, $f(x) = 0$, $x < \theta$, $\theta \in \mathbb{R}$. Mostre que $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{P} 1 + \theta$.

17. Seja $\{X_n, n \geq 1\}$ uma seqüência de variáveis aleatórias onde X_n tem distribuição gama com parâmetros n e β, $\beta > 0$. Calcule o limite em distribuição de $Z_n = \frac{X_n}{n}$.

18. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. com funções de probabilidade dadas por:
$$P\left\{X_n = \frac{k}{n}\right\} = \frac{1}{n+1}, \quad k = 0, 1, \ldots, n.$$
Ache o limite em distribuição de X_n.

19. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. com distribuição comum $N(\mu, \sigma^2)$. Qual o limite em probabilidade da seqüência
$$Y_n = \frac{X_1^2 + \cdots + X_n^2}{n} \; ?$$

20. Seja $\{X_n, n \geq 1\}$ uma seqüência de variáveis aleatórias com $E(X_n^2) < \infty$, $\forall n \geq 1$. Prove que se $\lim_{n\to\infty} E(X_n) = \alpha$ e $\lim_{N\to\infty} \mathrm{Var}(X_n) = 0$, então $X_n \xrightarrow{P} \alpha$.

21. Sejam $X_1, X_2, \ldots$ variáveis aleatórias independentes tais que $P(X_n = 1) = \frac{1}{n}$ e $P(X_n = 0) = 1 - \frac{1}{n}$. Ache o limite em probabilidade de X_n.

22. Em 10.000 lançamentos independentes de uma moeda, foram obtidas 6.400 caras. Você acha razoável assumir que esta moeda é honesta? Justifique. (Sugestão: use o teorema do limite central.)

23. Suponha que a duração, em horas, de uma lâmpada especial tenha distribuição exponencial de parâmetro 1/3. Se uma amostra de 36 lâmpadas é observada, qual é a probabilidade de que a média amostral seja inferior a duas horas? Justifique sua resposta.

24. Suponha que durante a semana o número de carros que chegam ao cruzamento da Av. Rebouças com a Av. Brasil entre 10 e 12 horas tenha distribuição de Poisson com média 100. Determine a probabilidade de que o número de carros que chegam neste cruzamento nesse horário numa determinada semana esteja entre 80 e 120. Justifique sua resposta.

25. Seja X uma variável aleatória com distribuição gama de parâmetros n e λ, $n \in \mathbb{N}$, isto é, $X \sim \text{gama}(n, \lambda)$. Justifique o emprego da distribuição normal para efetuar cálculos de probabilidades envolvendo a variável aleatória X. Mostre como calcular, por exemplo, $P(X \leq x)$, $x \in \mathbb{R}^+$.

26. O Ministério da Saúde deseja estimar a proporção de vítimas de uma certa moléstia no país. Determine o menor tamanho de amostra necessário para que a estimativa difira da verdadeira proporção por menos de 1% com probabilidade de 95% quando: (a) Sabe-se que a verdadeira proporção de vítimas dessa moléstia no país é inferior a 10%. (b) Nada se sabe sobre tal proporção.

27. Lança-se uma moeda honesta até se observar 100 caras. Determine a probabilidade de que sejam necessários no mínimo 220

lançamentos para se observar as 100 caras. Sugestão: utilize o Teorema do Limite Central.

28. Se 65% da população em uma comunidade é favorável à proposta de aumento nas mensalidades escolares, dê uma aproximação para a probabilidade de que uma amostra aleatória de 100 pessoas desta comunidade irá conter: (a) Pelo menos 50 pessoas favoráveis à proposta. (b) No máximo 75 pessoas favoráveis à proposta. Justifique seu procedimento.

29. Calcule $\lim_{n\to\infty} \frac{1}{(n-1)!}\int_0^n t^{n-1}e^{-t}dt$.

30. Um sistema é constituído de 100 componentes que funcionam independentemente. A probabilidade de que qualquer um dos componentes venha a falhar durante o período de operação é 1/10. A fim de que o sistema completo funcione, ao menos 85 componentes devem trabalhar. Calcule a probabilidade de que o sistema completo funcione. Justifique sua resposta.

31. Suponha que 30 dispositivos eletrônicos sejam empregados da seguinte maneira: inicia-se a operação com o primeiro dispositivo. Tão logo falhe o primeiro dispositivo, o segundo entra em operação. Tão logo falhe o segundo dispositivo, o terceiro dispositivo entra em operação e, assim, sucessivamente. Suponha que o tempo de duração de cada um destes 30 dispositivos seja uma variável aleatória exponencialmente distribuída com parâmetro $\beta = \frac{1}{10}$. Seja T o tempo total de operação dos 30 dispositivos. Qual é a probabilidade de que T ultrapasse 350 horas? Justifique seu resultado.

32. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. $N(0,\sigma^2)$. (a) Calcule $E(X_1^2)$ e $\mathrm{Var}(X_1^2)$. (b) Como podemos aproximar $P(X_1^2+\cdots+X_n^2 \le x)$ em termos de Φ, a função de distribuição de uma variável aleatória normal (0,1)?

33. Um garoto vende jornais junto a um semáforo de uma movimentada avenida. Suponha que cada carro que passa pelo semáforo compra um jornal com probabilidade $\frac{1}{3}$. Se Y denota o número de carros que passam pelo jornaleiro até que ele venda 90 exemplares do jornal, determine a distribuição (assintótica) de Y, especificando seus parâmetros. Justifique sua resposta.

34. Voltagens de sinais independentes, digamos, V_i, $i = 1,\ldots,n$, são recebidas num "somador". Seja V a soma das voltagens recebidas pelo "somador", isto é, $V = \sum_{i=1}^n V_i$. Suponha que cada V_i seja uma variável aleatória uniformemente distribuída no intervalo (0,10). Determine a probabilidade de que V exceda 105 ohms, quando $n = 20$. Justifique seus resultados.

35. Decide-se tomar uma amostra de n eleitores para avaliar a porcentagem daqueles eleitores que tencionem votar em determinado candidato em uma eleição. Seja p a proporção de eleitores desse candidato. Definindo $S_n = X_1 + \cdots + X_n$, onde $X_i = 1$ se eleitor tenciona votar no candidato e $X_i = 0$ caso contrário. Determine:
(a) $P\left(\left|\frac{S_n}{n} - p\right| \ge 0,025\right)$, quando $n = 900$.
(b) O tamanho da amostra, n tal que

$$P\left(\left|\frac{S_n}{n} - p\right| \ge 0,025\right) = 0,01.$$

(Interprete os resultados.)

36. Mostre que: $\lim_{n\to\infty} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} = \frac{1}{2}$. (Sugestão: considere uma seqüência de v.a.i.i.d. Poisson e utilize o Teorema do Limite Central.)

37. Calcule $\lim_{n\to\infty} e^{-n\lambda}\sum_{k=0}^{n}\frac{(n\lambda)^k}{k!}$, $0 < \lambda \ne 1$. (Sugestão: proceda como no exercício anterior.)

38. Um astrônomo está interessado em medir, em anos-luz, a distância de seu observatório para uma estrela distante. Apesar do astrônomo ter equipamentos de medição bastante sofisticados, ele sabe que, devido a mudanças na condição atmosférica e a erros de medição, cada medição feita por ele não dará a

distância exata, mas apenas uma estimativa. Desta forma, o astrônomo planeja realizar uma série de medições e então usar a média desses valores como estimativa da distância verdadeira. Se o astrônomo julga que as medições são v.a.i.i.d., tendo média μ (a verdadeira distância do seu observatório para a estrela) e variância comum 4 (anos-luz)^2, quantas medições ele precisa fazer para estar razoavelmente certo de que a distância estimada não difere da distância verdadeira por mais de 0,5 ano-luz?

39. Sejam $X_1, X_2, \ldots$ v.a.i.i.d. Poisson (1). Seja $S_n = X_1 + \cdots + X_n$, $n \geq 1$.

    (a) Calcule o valor exato de $P\{S_n = n\}$.

    (b) Utilizando o fato de que $\displaystyle\int_{-\frac{1}{\sqrt{n}}}^{0} e^{-\frac{x^2}{2}}\,dx \approx \frac{1}{\sqrt{n}}$ e o teorema do limite central, mostre que $P\{S_n = n\} \approx \frac{1}{\sqrt{2\pi n}}$.

    (c) Conclua de (a) e (b) que $n! \approx n^{n+\frac{1}{2}}e^{-n}\sqrt{2\pi}$, para n grande (esta é a famosa aproximação de Stirling para $n!$).

APÊNDICE

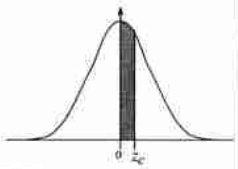
TABELAS

Resumos de tabelas extraídas de J. S. C. Pereira e W. O. Bussab, *Tábuas de Estatísticas*, São Paulo, Harbra & Row do Brasil, 1985.

DISTRIBUIÇÃO NORMAL REDUZIDA: $N(0;1)$

Probabilidades p tais que $p = P(0 < Z < z_c)$

SEGUNDA DECIMAL DE z_c



Parte inteira e 1ª decimal de z_c	0	1	2	3	4	5	6	7	8	9	Parte inteira e 1ª decimal de z_c
	$p = 0$										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

DISTRIBUIÇÃO NORMAL REDUZIDA: $N(0;1)$

Probabilidades p tais que $p = P(0 < Z < z_c)$

SEGUNDA E TERCEIRA DECIMAL DE z_c

Parte inteira e 1ª decimal de z_c	05	15	25	35	45	55	65	75	85	95	Parte inteira e 1ª decimal de z_c
	$p = 0$										
0,0	00199	00598	00997	01396	01795	02193	02591	02989	03387	03784	0,0
0,1	04181	04578	04974	05369	05764	06159	06553	06946	07339	07730	0,1
0,2	08121	08512	08901	09290	09677	10064	10450	10834	11218	11600	0,2
0,3	11982	12362	12741	13119	13495	13871	14244	14617	14988	15358	0,3
0,4	15726	16093	16458	16822	17184	17545	17903	18261	18500	18970	0,4
0,5	19322	19672	20021	20368	20712	21055	21396	21735	22073	22408	0,5
0,6	22741	23072	23401	23729	24054	24377	24697	25016	25333	25647	0,6
0,7	25959	26270	26577	26883	27186	27488	27786	28083	28377	28669	0,7
0,8	28959	29246	29531	29814	30094	30372	30648	30921	31192	31461	0,8
0,9	31727	31990	32252	32511	32767	33021	33273	33522	33769	34013	0,9
	05	15	25	35	45	55	65	75	85	95	

DISTRIBUIÇÃO DE POISSON: $P(\alpha)$

Probabilidades $P(X = r)$ Para $r = 0, 1, 2, \ldots, n$ e $\alpha = 0{,}001, \ldots, 10$

r \ α	0,001	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040	0,045	0,050	0,055	0,060	0,065	0,070	0,075	α / r
0	9990	9950	9900	9851	9802	9753	9704	9656	9608	9560	9512	9465	9418	9371	9324	9277	0
1	0010	0050	0099	0148	0196	0244	0291	0338	0384	0430	0476	0521	0565	0609	0653	0696	1
2	0^+	0^+	0^+	0001	0002	0003	0004	0006	0008	0010	0012	0014	0017	0020	0022	0026	2
3	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0001	3
≥ 4	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	≥ 4

r \ α	0,080	0,085	0,090	0,095	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000	1,200	1,400	α / r
0	9231	9185	9139	9094	9048	8187	7408	6703	6065	5488	4966	4493	4066	3679	3012	2466	0
1	0739	0781	0823	0864	0905	1637	2222	2681	3033	3293	3476	3595	3659	3679	3614	3452	1
2	0030	0033	0037	0041	0045	0164	0333	0536	0758	0988	1217	1438	1647	1839	2169	2417	2
3	0001	0001	0001	0001	0002	0011	0033	0071	0126	0198	0284	0383	0494	0613	0867	1128	3
4	0^+	0^+	0^+	0^+	0^+	0001	0003	0007	0016	0030	0050	0077	0111	0153	0260	0395	4
5	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0002	0004	0007	0012	0020	0031	0062	0111	5
6	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0002	0003	0005	0012	0026	6
7	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0002	0005	7
8	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	8
≥ 9	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	≥ 9

DISTRIBUIÇÃO DE POISSON: $P(\alpha)$ (Continuação)

Probabilidades $P(X = r)$ Para $r = 0, 1, 2, \ldots, n$ e $\alpha = 0{,}001, \ldots, 10$

r \ α	1,600	1,800	2,000	2,500	3,000	3,500	4,000	4,500	5,000	5,500	6,000	6,500	7,000	8,000	9,000	10,00	α / r
0	2019	1653	1353	0821	0498	0302	0183	0111	0067	0041	0025	0015	0009	0003	0001	0^+	0
1	3230	2975	2707	2052	1494	1057	0733	0500	0337	0225	0149	0098	0064	0027	0011	0005	1
2	2584	2678	2707	2565	2240	1850	1465	1125	0842	0618	0446	0318	0223	0107	0050	0023	2
3	1378	1607	1804	2138	2240	2158	1954	1687	1404	1133	0892	0688	0521	0286	0150	0076	3
4	0551	0723	0902	1336	1680	1888	1954	1898	1755	1558	1339	1118	0912	0573	0337	0189	4
5	0176	0260	0361	0668	1008	1322	1563	1708	1755	1714	1606	1454	1277	0916	0607	0378	5
6	0047	0078	0120	0278	0504	0771	1042	1281	1462	1571	1606	1575	1490	1221	0911	0631	6
7	0011	0020	0034	0100	0216	0385	0595	0824	1044	1234	1377	1462	1490	1396	1171	0901	7
8	0002	0005	0009	0031	0081	0169	0298	0463	0653	0849	1033	1188	1304	1396	1318	1126	8
9	0^+	0001	0002	0009	0027	0066	0132	0232	0363	0519	0688	0858	1014	1241	1318	1251	9
10	0^+	0^+	0^+	0002	0008	0023	0053	0104	0181	0285	0413	0558	0710	0993	1186	1251	10
11	0^+	0^+	0^+	0^+	0002	0007	0019	0043	0082	0143	0225	0330	0452	0722	0970	1137	11
12	0^+	0^+	0^+	0^+	0^+	0002	0006	0016	0034	0065	0113	0179	0264	0481	0728	0948	12
13	0^+	0^+	0^+	0^+	0^+	0^+	0002	0006	0013	0028	0052	0089	0142	0296	0504	0729	13
14	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0002	0005	0011	0022	0041	0071	0169	0324	0521	14
15	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0002	0004	0009	0018	0033	0090	0194	0347	15
16	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0003	0007	0014	0045	0109	0217	16
17	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0003	0006	0021	0058	0128	17
18	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0002	0009	0029	0071	18
19	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0004	0014	0037	19
20	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0002	0006	0019	20
21	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0003	0009	21
22	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	0004	22
23	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0002	23
24	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0001	24
≥ 25	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	0^+	≥ 25

DISTRIBUIÇÃO BINOMIAL: $B(n;p)$

Valores das distribuições binomiais para os seguintes valores de n : 5, 6, 10, 12, 15 e 20. Apresentação em forma de matriz onde na coluna da matriz estão apresentados os valores da variável binomial X. Na linha da matriz são apresentados os valores de p para cada distribuição binomial. No corpo da tabela são apresentadas as probabilidades. O valor da $P[X = 3]$ na distribuição binomial com $n = 5$ e $p = 0,25$ é encontrado no cruzamento da linha $n = 3$ com a coluna $p = 0,25$ na tabela com $n = 5$.

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 5$
$x = 0$	774	590	328	237	168	078	031	5
1	204	328	410	396	360	259	156	4
2	021	073	205	264	309	346	312	3
3	001	008	051	088	132	230	312	2
4	0^+	0^+	006	015	028	077	156	1
5	0^+	0^+	0^+	001	002	010	031	$0 = x$
$n = 5$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 6$
$x = 0$	735	531	262	178	118	047	016	6
1	232	354	393	356	303	187	094	5
2	031	098	246	297	324	311	234	4
3	002	015	082	132	185	276	312	3
4	0^+	001	015	033	060	138	234	2
5	0^+	0^+	002	004	010	037	094	1
6	0^+	0^+	0^+	0^+	001	004	016	$0 = x$
$n = 6$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 10$
$x = 0$	599	349	107	056	028	006	001	10
1	315	387	268	188	121	040	010	9
2	075	194	302	282	233	121	044	8
3	010	057	201	250	267	215	117	7
4	001	011	088	146	200	251	205	6
5	0^+	001	026	058	103	201	246	5
6	0^+	0^+	006	016	037	111	205	4
7	0^+	0^+	001	003	009	042	117	3
8	0^+	0^+	0^+	0^+	001	011	044	2
9	0^+	0^+	0^+	0^+	0^+	002	010	1
10	0^+	0^+	0^+	0^+	0^+	0^+	001	$0 = x$
$n = 10$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 12$
$x = 0$	540	282	069	032	014	002	0^+	12
1	341	377	206	127	071	017	003	11
2	099	230	283	232	168	064	016	10
3	017	085	236	258	240	142	054	9
4	002	021	133	194	231	213	121	8
5	0^+	004	053	103	158	227	193	7
6	0^+	0^+	016	040	079	177	226	6
7	0^+	0^+	003	011	029	101	193	5
8	0^+	0^+	001	002	008	042	121	4
9	0^+	0^+	0^+	0^+	001	012	054	3
10	0^+	0^+	0^+	0^+	0^+	002	016	2
11	0^+	0^+	0^+	0^+	0^+	0^+	003	1
12	0^+	0^+	0^+	0^+	0^+	0^+	0^+	$0 = x$
$n = 12$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 15$
$x = 0$	463	206	035	013	005	0^+	0^+	15
1	366	343	132	067	031	005	0^+	14
2	135	267	231	156	092	022	003	13
3	031	129	250	225	170	063	014	12
4	005	043	188	225	219	127	042	11
5	001	010	103	165	206	186	092	10
6	0^+	002	043	092	147	207	153	9
7	0^+	0^+	014	039	081	177	196	8
8	0^+	0^+	003	013	035	118	196	7
9	0^+	0^+	001	003	012	061	153	6
10	0^+	0^+	0^+	001	003	024	092	5
11	0^+	0^+	0^+	0^+	001	007	042	4
12	0^+	0^+	0^+	0^+	0^+	002	014	3
13	0^+	0^+	0^+	0^+	0^+	0^+	003	2
14	0^+	0^+	0^+	0^+	0^+	0^+	0^+	1
15	0^+	0^+	0^+	0^+	0^+	0^+	0^+	$0 = x$
$n = 15$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

p→	0,05	0,10	0,20	0,25	0,30	0,40	0,50	$n = 20$
$x = 0$	358	122	012	003	001	0^+	0^+	20
1	377	270	058	021	007	0^+	0^+	19
2	189	285	137	067	028	003	0^+	18
3	060	190	205	134	072	012	001	17
4	013	090	218	190	130	035	005	16
5	002	032	175	202	179	075	015	15
6	0^+	009	109	169	192	124	037	14
7	0^+	002	055	112	164	166	074	13
8	0^+	0^+	022	061	114	180	120	12
9	0^+	0^+	007	027	065	160	160	11
10	0^+	0^+	002	010	031	117	176	10
11	0^+	0^+	0^+	003	012	071	160	9
12	0^+	0^+	0^+	001	004	035	120	8
13	0^+	0^+	0^+	0^+	001	015	074	7
14	0^+	0^+	0^+	0^+	0^+	005	037	6
15	0^+	0^+	0^+	0^+	0^+	001	015	5
16	0^+	0^+	0^+	0^+	0^+	0^+	005	4
17	0^+	0^+	0^+	0^+	0^+	0^+	001	3
18	0^+	0^+	0^+	0^+	0^+	0^+	0^+	2
19	0^+	0^+	0^+	0^+	0^+	0^+	0^+	1
20	0^+	0^+	0^+	0^+	0^+	0^+	0^+	$0 = x$
$n = 20$	0,95	0,90	0,80	0,75	0,70	0,60	0,50	$\leftarrow p$

BIBLIOGRAFIA

BERRY, D. A. *Statistics a Bayesian Perspective.* Belmont, Duxbury Press, 1996.

BILLINGSLEY, P. *Probability and Measure.* 2. ed. New York, J. Wiley, 1986.

BREIMAN, L. *Probability.* Reading, Addison-Wesley, 1968.

CHUNG, K. L. *Elementary Probability Theory with Stochastic Processes.* New York/Heidelberg/Berlin, Springer-Verlag, 1974.

————. *A Course in Probability Theory.* 2. ed. New York, Academic Press, 1974.

CRAMÉR, H. *Elementos da Teoria da Probabilidade e Algumas de Suas Aplicações.* São Paulo, Mestre Jou, 1955.

DE GROOT, M. H. *Probability and Statistics.* 2. ed. Reading, Addison-Wesley, 1986.

FELLER, W. *An Introduction to Probability Theory and Its Application.* 3. ed. New York, John Wiley & Sons, 1967, vol. 1.

HOEL, P. G.; PORT, S. C. & STONE, C. J. *Introdução à Teoria da Probabilidade.* Rio de Janeiro, Interciência, 1978.

JAMES, B. R. *Probabilidade: Um Curso em Nível Intermediário.* Livros Técnicos e Científicos, 1981.

LINDLEY, D.V. *Making Decisions.* 2. ed. London, J. Wiley, 1985.

LOÈVE, M. *Probability Theory.* 3. ed. New York, D. Van Nostrand, 1963.