é devida a von Mises. Esta noção tornou possível a construção de uma teoria das probabilidades estritamente matemática, com base na teoria da medida. Esse tipo de abordagem surgiu gradativamente na década de vinte sob a influência de vários autores. Um tratamento axiomático descrevendo a versão moderna foi dado por A. Kolmogorov.[3] Essa será a linha que seguiremos, embora a palavra axioma pareça demasiadamente formal, se considerarmos que o presente volume trata apenas do caso simples de probabilidades discretas.

[3]A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlim (Springer) 1933. Uma tradução inglesa (por N. Morrison) foi publicada em 1956: *Foundations of the theory of probability*, New York (Chelsea).

CAPÍTULO I

O espaço amostral

1. OS FUNDAMENTOS EMPÍRICOS

A teoria matemática das probabilidades ganha valor prático e significado intuitivo em conexão com experimentos reais ou imaginários tais como lançar uma moeda uma vez, lançar uma moeda cem vezes, jogar três dados, embaralhar cartas, emparelhar dois baralhos, jogar na roleta, observar o tempo de vida de um átomo radioativo ou de uma pessoa, selecionar uma amostra de indivíduos e observar na mesma o número de canhotos, cruzar duas espécies de plantas e observar o fenótipo da espécie resultante; ou com fenômenos, tais como o sexo de um recém-nascido, o número de troncos ocupados numa central telefônica, o número de chamadas para um telefone, o ruído casual num sistema de comunicação elétrica, controle de qualidade rotineiro de um processo de produção, freqüências de acidentes, o número de estrelas duplas numa região do céu, a posição de uma partícula sujeita a um processo de difusão. Todas essas descrições são um tanto vagas e a fim de que a teoria faça sentido é necessário que se estabeleça o que se entende por *resultados possíveis do experimento ou observação em questão.*

Quando se joga uma moeda para cima ela não cai necessariamente "cara" ou "coroa"; ela pode rolar para longe ou cair em pé. Não obstante consideraremos "cara" e "coroa" como os únicos resultados possíveis desse experimento. Essa convenção simplifica a teoria sem afetar a sua aplicabilidade. Idealizações desse tipo são comumente feitas. É impossível medir o tempo de vida de um átomo ou de uma pessoa sem erro algum, mas, para fins teóricos é conveniente imaginar que essas quantidades são números exatos. A questão que se coloca é sobre os números que poderiam na realidade representar o tempo de vida de uma pessoa. Existe uma idade máxima, além da qual a vida é impossível ou qualquer idade seria admissível? Relutamos em aceitar que um homem possa atingir mil anos de vida e entretanto a prática atuarial corrente não admite limites à duração possível da vida. De acordo com as fórmulas, nas quais são baseadas as tábuas modernas de mortalidade, a proporção de homens que atingem mil anos de vida é da ordem de grandeza de um em $10^{10^{36}}$ – um número com 10^{27} bilhões de zeros. Essa afirmação não faz sentido algum do ponto de vista biológico ou sociológico mas, considerada exclusivamente do ponto de vista estatístico, ela certamente não contradiz experiência alguma. Há menos do que 10^{10} pessoas nascidas num século. Para testar a controvérsia estatisticamente, mais de $10^{10^{35}}$ séculos seriam necessários, o que é consideravelmente superior a $10^{10^{34}}$ vezes a idade da Terra. Obviamente, probabilidades tão diminutas são compatíveis com a nossa noção de impossibilidade. Seu uso pode parecer extremamente absurdo, mas não traz inconveniente algum e permite a simplificação de muitas fórmulas. Além do mais, se estivéssemos seriamente dispostos a desprezar a possibilidade de viver mil anos, teríamos que aceitar a existência de uma idade máxima e a suposição de que seria

Bolas

b.1. *Aniversários.* As possíveis distribuições dos aniversários de r pessoas correspondem às diferentes disposições de r bolas em $n = 365$ compartimentos. (Supondo-se que o ano tenha 365 dias).
b.2. *Acidentes.* Classificar r acidentes, de acordo com o dia da semana em que eles ocorreram, equivale a colocar r bolas em $n = 7$ compartimentos.
b.3. *Tiro ao alvo.* Aos tiros associamos as bolas e aos alvos fazemos corresponder os compartimentos.
b.4. *Amostragem.* Consideremos um grupo de pessoas classificadas de acordo com a idade e a profissão. As classes desempenham o papel dos compartimentos e as pessoas o papel das bolas.
b.5. *Irradiação em biologia.* Quando as células da retina são expostas à luz, as partículas de luz fazem o papel das bolas e as células o dos compartimentos do nosso modelo. Do mesmo modo no estudo do efeito genético da irradiação, os cromosomas correspondem aos compartimentos do nosso modelo e às partículas alfa associamos as bolas do modelo.
b.6. *Experimentos com raios cósmicos.* As partículas que atingem os contadores Geiger representam as bolas e os contadores funcionam como compartimentos.
b.7. *O problema do elevador.* Um elevador começa a funcionar com r passageiros e pára em n andares. As diferentes distribuições da descida dos passageiros são réplicas das diferentes distribuições de r bolas em n compartimentos.
b.8. *Lançamento de dados.* Os possíveis resultados de um lançamento de r dados correspondem à colocação de r bolas em $n = 6$ compartimentos. No lançamento de uma moeda teremos $n = 2$ compartimentos.
b.9. *Dígitos aleatórios.* O conjunto de todas as permutações possíveis de uma seqüência de r dígitos pode ser associado à distribuição de r bolas (correspondentes aos lugares) pelos dez compartimentos denominados $0, 1, \ldots, 9$.
b.10. *Distribuição de r pessoas de acordo com o sexo.* Nesse caso teremos r bolas e $n = 2$ compartimentos.
b.11. *Coleções de cupons.* As diferentes espécies de cupons representam os compartimentos e os cupons colecionados representam as bolas.
b.12. *Distribuição de ases no bridge.* Os quatro jogadores representam os compartimentos e os quatro ases representam as bolas.
b.13. *Distribuição de genes.* Cada descendente de um indivíduo (pessoa, planta, animal) herda determinados genes de seu progenitor. Admitindo-se que um dado gene possa se apresentar em n formas distintas $A_1, \ldots, A_n$ então o descendente poderia ser classificado de acordo com a forma na qual esse gene particular se apresenta. Os descendentes fariam o papel das bolas e os genótipos $A_1, \ldots, A_n$ representariam os compartimentos.
b.14. *Química.* Suponha que um polímero de uma cadeia longa reaja com oxigênio. Uma cadeia individual poderá reagir com 0, 1, 2, ... moléculas de oxigênio. Aqui as moléculas de oxigênio reagentes fazem o papel de bolas e as cadeias do polímero o de compartimentos onde as bolas são colocadas.
b.15. *Teoria das emulsões fotográficas.* Uma lâmina fotográfica é coberta de granulações sensíveis aos quanta luminosos. Uma granulação reage se ela for atingida por um certo número r de quanta. Para a teoria do contraste branco-preto interessa saber o número de granulações suscetíveis de serem atingidas pelos r quanta. Temos aqui um problema de ocupação no qual as granulações fazem o papel dos compartimentos e os quanta de luz representam as bolas. (Na realidade a

situação é mais complicada pois, uma lâmina contém usualmente granulações de diferentes sensibilidades.)
b.16. *Erros de impressão.* As possíveis distribuições de r erros de impressão em n páginas de um livro correspondem a todas as diferentes distribuições de r bolas em n compartimentos desde que r seja menor do que o número de letras por página.

c) *O caso de bolas indistinguíveis.* Retornemos ao Exemplo a) e suponhamos que as três bolas sejam idênticas. Isso significa que não será possível diferenciar disposições como 4, 5, 6 e, portanto, a Tab. I.1 se reduzirá à Tab. I.2. Esta última *define* o espaço amostral do experimento ideal que denominaremos "colocação de três bolas indistinguíveis em três compartimentos" e um procedimento análogo se aplica ao caso de r bolas em n compartimentos.

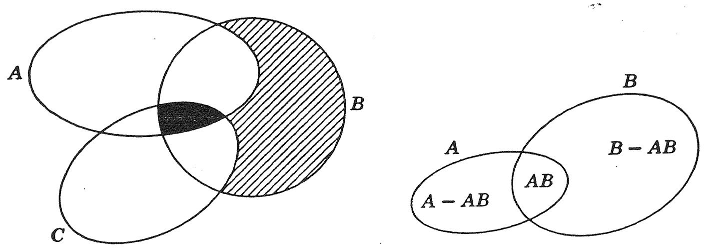
Tabela I.2

1.	{***	–	– }	6.	{ *	**	– }
2.	{ –	***	– }	7.	{ *	–	** }
3.	{ –	–	***}	8.	{ –	**	* }
4.	{**	*	– }	9.	{ –	*	** }
5.	{**	–	* }	10.	{ *	*	* }

O fato das bolas serem ou não realmente distinguíveis na prática é irrelevante para a nossa teoria. Mesmo que sejam distinguíveis é possível tratá-las teoricamente como indistinguíveis. Os ases no jogo de bridge (Exemplo b.12) ou as pessoas num elevador (Exemplo b.7) são certamente distinguíveis. Os dados do Exemplo b.8 podem ser coloridos para torná-los distinguíveis, mas, num dado problema, a escolha do modelo a ser usado é puramente uma questão de objetivo e conveniência. A natureza de um problema concreto pode determinar a escolha, mas, em qualquer circunstância, a nossa teoria só começa depois do modelo apropriado ter sido escolhido isto é, depois do espaço amostral ter sido definido.

No esquema considerado anteriormente tratamos do caso de bolas idênticas mas a Tab. I.2 ainda se refere a um primeiro, segundo e terceiro compartimentos e a ordem dos mesmos é essencial. Indo um pouco mais além podemos admitir que mesmo os compartimentos são indistinguíveis. (Por exemplo, o compartimento poderia ser escolhido ao acaso sem preocupação quanto ao seu conteúdo. Sendo ambos, bolas e compartimentos indistinguíveis, somente três disposições distintas são possíveis, a saber: {***|–|–}, {**|*|–}, {*|*|*}.

d) *Amostragem.* Suponha que uma amostra de cem pessoas seja selecionada com o objetivo de estimar o número de fumantes na população. Para esse fim a única propriedade de interesse na amostra é o número x de fumantes que poderá ser um inteiro qualquer entre 0 e 100. Neste caso o nosso espaço amostral terá 101 "pontos, $0, 1, \ldots, 100$. Cada amostra ou observação específica é completamente descrita pelo ponto x correspondente. Um exemplo de um evento composto é o resultado "a maioria das pessoas consultadas são fumantes". Isso significa que o experimento resultou em um dos cinqüenta eventos simples $51, 52, \ldots, 100$ mas, não é declarado qual deles foi o resultado. Da mesma maneira cada propriedade da amostra pode ser descrita enumerando-se os casos correspondentes, ou pontos amostrais. Para que haja uniformidade na terminologia, falaremos de



A Fig. I.1 ilustra relações entre eventos. À esquerda a região dentro do contorno é a união $A \cup B \cup C$. O domínio triangular (*sombreado escuro*) é a interseção ABC. O domínio em forma de lua (*sombreado claro*) é a interseção de B com o complementar de $A \cup C$

Podemos associar a dois eventos quaisquer A e B dois novos eventos, definidos pelas condições "ambos A e B ocorrem" e "ou A ou B ou ambos ocorrem". Esses eventos serão designados por AB e $A \cup B$, respectivamente. O evento AB contém todos os pontos amostrais que são comuns a A e a B. Se A e B se excluem mutuamente, então não há pontos comuns a A e a B e o evento AB é impossível; analiticamente esta situação é descrita pela relação

$$AB = 0, \tag{4.1}$$

que deve ser lida "A e B, são *mutuamente exclusivos*". O evento AB' deve ser interpretado como a ocorrência simultânea do evento A e do complementar de B ou, em outras palavras, que o evento A ocorre e o evento B não ocorre. Analogamente, $A'B'$ significa que nenhum dos eventos A e B ocorre. O evento $A \cup B$ significa que pelo menos um dos eventos A e B ocorre, ele contém todos os pontos amostrais com exceção daqueles que não pertencem nem a A nem a B.

Na teoria das probabilidades descrevemos o evento AB como a ocorrência simultânea de A e B. Na terminologia matemática usual AB se denomina a interseção (lógica) de A e B. Analogamente $A \cup B$ é a união dos conjuntos A e B. Essa noção se estende ao caso de mais de dois eventos: $A, B, C, D, \ldots$

Definição 3. *É possível associarmos dois novos eventos à toda coleção de eventos, $A, B, C, \ldots$ da seguinte maneira: o conjunto de todos os pontos amostrais que pertencem a todos os eventos dados será designado por $ABC \ldots$, e se denominará interseção*[2] *(ou realização simultânea) de $A, B, C, \ldots$ O conjunto dos pontos amostrais que pertencem a pelo menos um dos eventos dados, será denotado por $A \cup B \cup C \cup \ldots$ e se denominará união (ou realização de pelo menos um) dos eventos dados. Os eventos $A, B, C, \ldots$ são mutuamente exclusivos se dois quaisquer deles não têm pontos comuns, ou seja, se $AB = 0$, $AC = 0, \ldots$, $BC = 0, \ldots$*

Necessitamos, ainda, de um símbolo para expressar o fato de que a ocorrência de A implica na ocorrência de B. Isso significa que todo ponto de A é também um ponto de B. Uma analogia intuitiva seria dada pelo conjunto de todas as mães que é uma parte do conjunto de todas as mulheres: todas as mães são mulheres, mas nem todas as mulheres são mães.

Definição 4. *Os símbolos $A \subset B$ e $B \supset A$ são equivalentes e significam que todo ponto de A está contido em B; leia-se respectivamente, "A implica B" e "B é*

[2] A notação matemática usual para a interseção de dois ou mais conjuntos é $A \cap B$ ou $A \cap B \cap C$, etc. Essa notação é a mais adequada para certos propósitos específicos. Usaremos a notação AB, ABC, etc., por ser ela menos complicada para impressão.

implicado por A". Se este for o caso, escreveremos $B - A$ em vez de BA' para denotar o evento de que B ocorre e A não ocorre.

O evento $B - A$ contém todos os pontos que pertencem a B e não a A. Com essa notação podemos escrever: $A' = \mathfrak{S} - A$ e $A - A = 0$.

Exemplos. a) Se A e B são eventos mutuamente exclusivos a ocorrência do evento A implica que o evento B não ocorre e vice-versa. Portanto as relações, $AB = 0$, $A \subset B'$ e $B \subset A'$ são equivalentes.

b) A expressão $A - AB$ denota ocorrência do evento A, mas não a ocorrência simultânea de A e B. Portanto $A - AB = AB'$.

c) No Exemplo 2g, o evento AB significa que o marido tem mais de quarenta anos e é mais velho que a esposa; AB' significa que ele tem mais de quarenta anos mas não é mais velho que a esposa. AB é representado graficamente pela região trapezoidal infinita, determinada pelo eixo dos x e as retas $x = 40$ e $y = z$, e o evento AB' é representado pela região angular compreendida entre as retas $x = 40$ e $y = x$, sendo que a fronteira desta última está incluída na região. O evento AC significa que ambos, marido e mulher, têm mais de quarenta anos. O evento $A \cup C$ significa que, pelo menos um deles tem mais de quarenta anos e o evento $A \cup B$ significa que o marido tem mais de quarenta anos ou, se não for esse o caso, é pelo menos mais velho que a mulher (em outras palavras, a idade do marido supera o mínimo entre a idade da mulher e quarenta anos).

d) No Exemplo 2a vamos denotar por V_i $(i = 1, 2, 3)$ o evento correspondente a encontrarmos o i-ésimo compartimento vazio. De maneira análoga, U_i, D_i, T_i denotam, respectivamente, os eventos correspondentes a encontrarmos, no i-ésimo compartimento, uma, duas ou três bolas. Então, $V_1V_2 = T_3$, $U_1U_2 \subset U_3$ e $D_1D_2 = 0$. Observe também que $T_1 \subset V_2$, etc. O evento $D_1 \cup D_2 \cup D_3$ é definido pela condição de que existe pelo menos um compartimento duplamente ocupado.

e) *Bridge*. Denote por A, B, C, D respectivamente os eventos de que norte, sul, leste, oeste têm, pelo menos um ás. Desde que no mínimo um dos jogadores têm um ás, um ou mais desses quatro eventos ocorrem certamente. Portanto $A \cup B \cup C \cup D = \mathfrak{S}$ isto é, esse evento corresponde a todo o espaço amostral. O evento $ABCD$ ocorre se e só se cada jogador tiver um ás. O evento "oeste tem quatro ases" significa que nenhum dos três eventos A, B, C ocorreu; isto é equivalente à ocorrência simultânea de A', B', C' isto é, equivalente ao evento $A'B'C'$.

f) No Exemplo 2g tem-se $BC \subset A$; isto é "se o marido é mais velho que a mulher (evento B) e a mulher tem mais de quarenta anos (evento C) segue-se que o marido tem mais de quarenta anos (evento A). De que maneira descreveríamos, em palavras, o evento $A - BC$?

5. ESPAÇOS AMOSTRAIS DISCRETOS

Os espaços amostrais mais simples são os que contêm apenas um número finito, n, de pontos. Se n é razoavelmente pequeno (como, por exemplo, no caso do lançamento de algumas moedas) é fácil visualizar o espaço. O espaço correspondente às distribuições de cartas no jogo de bridge é mais complicado, mas é possível imaginar cada ponto amostral representado numa ficha e então considerar a coleção de fichas como representante do nosso espaço amostral. Um evento A (como, por exemplo, "norte tem dois ases") será representado por um certo conjunto a realidade. Mais importante ainda é o fato empírico de que os desvios que se verificam são quase sempre associados com fenômenos tais como a posição não-central do centro da gravidade da moeda. Dessa forma o nosso modelo ideal pode ser extremamente útil, mesmo que ele nunca se aplique de maneira exata. Por exemplo, na teoria moderna do controle estatístico de qualidade, baseada nos métodos de Shewhart, modelos probabilísticos ideais são usados para detectar as "causas possíveis" das discrepâncias flagrantes entre a realidade e os modelos, permitindo dessa forma a correção, ainda num estágio inicial, de defeitos iminentes de máquinas e de irregularidades no processo.

Observações semelhantes se aplicam a outras situações. Assim, por exemplo, o número total de possíveis distribuições de cartas no jogo de bridge alcança quase 10^{30} e usualmente consideramos essas distribuições como igualmente prováveis. Para verificarmos, na prática, o ajuste desse modelo seriam necessários mais de 10^{30} experimentos; se todos os habitantes do mundo jogassem uma partida por segundo, dia e noite, essa verificação levaria vários trilhões de anos. Entretanto, é possível verificarmos experimentalmente as conseqüências dessa pressuposição, por exemplo, através da observação do número de vezes em que um número múltiplo de ases aparece nas rodadas de bridge. Verifica-se então que o modelo, embora de forma grosseira, descreve a realidade suficientemente bem, principalmente se o embaralhamento das cartas for feito de maneira mais cuidadosa do que a usual. Mais importante do que a aplicação exata do modelo é a descoberta das causas possivelmente responsáveis pelas discrepâncias — no caso em questão essa descoberta nos leva a uma procura de melhores formas de embaralhamento. Esses exemplos são, é claro, de importância bastante restrita, mas eles servem para dar uma idéia da utilidade dos modelos teóricos. Casos mais interessantes aparecerão à medida que o nosso estudo prossegue.

Exemplos. a) *Bolas distintas.* No Exemplo a, Sec. 2 parece natural que se suponha serem todos os pontos amostrais *igualmente prováveis*, isto é, que *a cada ponto amostral se associe a probabilidade* $1/n^r$. Podemos tomar isso como uma *definição* e investigar suas conseqüências. Em algumas aplicações a pressuposição de que todos os pontos têm a mesma probabilidade é imposta por considerações de ordem física; em outras ela é introduzida por ser o modelo mais simples e com o objetivo de obtermos uma orientação geral, embora estejamos certos de que ela se constitui numa primeira aproximação bastante imperfeita. (Veja os Exemplos b1, Sec. 2 aniversários; b7, Sec. 2, o problema do elevador, ou b11, Sec. 2, colecionando cupons.)

b) *Bolas iguais. Estatísticas de Bose-Einstein.* Discutiremos agora o Exemplo c, Sec. 2 que trata da distribuição de 3 bolas iguais em 3 compartimentos. É possível argumentarmos que o experimento físico real não é afetado pela nossa incapacidade de distinguir entre as bolas. Neste caso existiriam ainda 27 possibilidades distintas, mesmo que somente 10 delas pudessem ser identificadas. Essas considerações nos levam a atribuir aos pontos da Tab. I.2 as probabilidades seguintes:

Número do ponto	1	2	3	4	5	6	7	8	9	10
Probabilidade	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$

Para a grande maioria das aplicações descritas no Exemplo b, Sec. 2 esse argumento parece válido e a atribuição das probabilidades parece ser bastante

razoável. Historicamente, esse argumento foi aceito, sem reservas, por um longo período de tempo e na estatística mecânica ele serviu de base para a derivação das estatísticas de Maxwell-Boltzmann para a distribuição de r bolas em n compartimentos. Grande foi, portanto, a surpresa quando Bose e Einstein mostraram que determinadas partículas estão sujeitas às estatísticas de Bose-Einstein (para maiores detalhes veja II.5). No caso, $r = n = 3$, o novo modelo atribui probabilidade 1/10 a *cada um dos dez pontos amostrais.*

Este exemplo mostra que é possível termos dois conjuntos distintos de valores de probabilidades associados ao mesmo espaço amostral e dessa forma ilustra a complicada inter-relação da teoria e experiência. Em particular ele nos ensina a não confiarmos demais em argumentos apriorísticos e a estarmos preparados para aceitar esquemas novos e não previstos.

c) *Lançamento de moedas.* Para que seja possível darmos uma interpretação freqüentista ao postulado das probabilidades iguais necessitamos ter à nossa disposição coleções de experimentos reais. Entretanto, dado que nenhuma moeda é perfeita, é possível conseguirmos um ajuste melhor ao modelo, através de experimentos simulados do que através do lançamento real de moedas. Para que se tenha uma idéia das flutuações que devem ser esperadas, apresentamos o resultado de um experimento simulado, que corresponde a 10 000 lançamentos de uma moeda[3]. A Tab. I.3 apresenta o número de caras resultante de uma seqüência de 100 experimentos, correspondendo cada experimento a 100 lançamentos de uma moeda. O total geral de caras é 4 979. Ao olhar para esses números o leitor, por certo, se sentirá tentado a perguntar: E daí? A verdade é que uma teoria bem mais avançada é necessária para que possamos julgar até que ponto esses dados empíricos estão de acordo com o nosso modelo abstrato. (Incidentalmente voltaremos a esse assunto em III.6).

Tabela I.3

Número do ensaio	Número de caras										Total
0– 1 000	54	46	53	55	46	54	41	48	51	53	501
– 2 000	48	46	40	53	49	49	48	54	53	45	485
– 3 000	43	52	58	51	51	50	52	50	53	49	509
– 4 000	58	60	54	55	50	48	47	57	52	55	536
– 5 000	48	51	51	49	44	52	50	46	53	41	485
– 6 000	49	50	45	52	52	48	47	47	47	51	488
– 7 000	45	47	41	51	49	59	50	55	53	50	500
– 8 000	53	52	46	52	44	51	48	51	46	54	497
– 9 000	45	47	46	52	47	48	59	57	45	48	494
–10 000	47	41	51	48	59	51	52	55	39	41	484

4979
510
4979/5021

7. DEFINIÇÕES BÁSICAS E REGRAS OPERACIONAIS

Convenção fundamental. Dado um espaço amostral discreto $\mathfrak{S}$ *cujos pontos denotaremos por* $E_1, E_2, \ldots$, *vamos supor que a cada ponto* E_j *esteja associado um*

[3]Na realidade, a tabela foi construída a partir das freqüências dos algarismos pares encontrados numa parte do trabalho, *A million random digits with* 100,000 *normal deviates*, publicado por The Rand Corporation, Glencoe Illinois (The Free Press), 1955.

6. Modifique o Exemplo b, Sec. 5 de tal forma que seja possível levarmos em conta a possibilidade de empates nos jogos individuais. Descreva o espaço amostral apropriado. De que maneira você definiria as probabilidades?

7. No Prob. 3 mostre que $A_1A_2A_3 \subset A_4$ e $A_1A_2A_3' \subset A_4'$.

8. Usando a notação do Exemplo d, Sec. 4 mostre que a) $U_1U_2D_3 = 0$; b) $U_1D_2 \subset V_3$; c) $V_3 - D_2U_1 \supset U_2D_1$.

9. No lançamento de dois dados seja A o evento de que a soma dos resultados seja ímpar e B o evento de que pelo menos um dos resultados seja igual a 1. Descreva os eventos AB, $A \cup B$, AB'. Determine as probabilidades desses eventos, admitindo que todos os 36 pontos amostrais têm a mesma probabilidade.

10. No Exemplo g, Sec. 2 discuta o significado dos eventos seguintes: a) ABC; b) $A - AB$; c) $AB'C$.

11. No Exemplo g, Sec. 2 mostre que: $AC' \subset B$.

12. *Bridge* (veja nota de rodapé na p. 8) Para $k = 1, 2, 3, 4$ denote por N_k o evento de que norte tem pelo menos k ases. Sejam S_k, L_k, O_k os eventos análogos para sul, leste e oeste. Se x designa o número de ases nas mãos do jogador oeste, determine as relações que x deve satisfazer nos eventos seguintes:

a) O_1'; b) N_2S_2; c) $N_1'S_1'L_1'$; d) $O_2 - O_3$; e) $N_1S_1L_1O_1$; f) N_3O_1; g) $(N_2 \cup S_2)L_2$.

13. No problema anterior, verifique: a) $S_3 \subset S_2$; b) $S_3O_2 = 0$; c) $N_2S_1L_1O_1 = 0$; d) $N_2S_2 \subset O_1'$; e) $(N_2 \cup S_2)O_3 = 0$; f) $O_4 = N_1'S_1'L_1'$.

14. Verifique as relações[4]

a) $(A \cup B)' = A'B'$;
b) $(A \cup B) - B = A - AB = AB'$;
c) $AA = A \cup A = A$;
d) $(A - AB) \cup B = A \cup B$;
e) $(A \cup B) - AB = AB' \cup A'B$;
f) $A' \cup B' = (AB)'$;
g) $(A \cup B)C = AC \cup BC$.

15. Simplifique as expressões a seguir.

a) $(A \cup B)(A \cup B')$, b) $(A \cup B)(A' \cup B)(A \cup B')$, c) $(A \cup B)(B \cup C)$.

16. Nas relações seguintes, determine quais são verdadeiras e quais são falsas:

a) $(A \cup B) - C = A \cup (B - C)$;
b) $ABC = AB(C \cup B)$;
c) $A \cup B \cup C = (A - AB) \cup B$; $= A \cup (B - AB) \cup (C - AC)$
d) $A \cup B = (A - AB) \cup B$;
e) $AB \cup BC \cup CA \supset ABC$;
f) $(AB \cup BC \cup CA) \subset (A \cup B \cup C)$;
g) $(A \cup B) - A = B$;
h) $AB'C \subset A \cup B$;
i) $(A \cup B \cup C)' = A'B'C'$;
j) $(A \cup B)'C = A'C \cup B'C$;
k) $(A \cup B)'C = A'B'C$;
l) $(A \cup B)'C = C - C(A \cup B)$.

17. Sejam A, B, C três eventos arbitrários. Escreva as expressões correspondentes aos eventos abaixo:

a) Somente A ocorre; b) A e B ocorrem mas, C não ocorre; c) Os três eventos ocorrem; d) Pelo menos um dos três eventos ocorre; e) Pelo menos dois dos eventos ocorrem; f) Um e somente um dos eventos ocorre; g) Dois e somente dois dos eventos ocorrem; h) Nenhum dos três eventos ocorre; i) No máximo dois eventos ocorrem.

18. A união $A \cup B$ de dois eventos pode ser expressa como uma união de dois eventos mutuamente exclusivos, a saber: $A \cup B = A \cup (B - AB)$. Expresse, de maneira similar, a união de três eventos A, B, C.

19. Usando o resultado do Prob. 18 prove que

$$\mathbf{P}\{A \cup B \cup C\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} + \mathbf{P}\{C\} - \mathbf{P}\{AB\} - \mathbf{P}\{AC\} - \mathbf{P}\{BC\} + \mathbf{P}\{ABC\}.$$

Essa igualdade é um caso particular de IV.(1.5).

[4]Observe que $(A \cup B)'$ denota o complementar de $A \cup B$, o que não é a mesma coisa que $A' \cup B'$. Analogamente, $(AB)'$ não é igual a $A'B'$.

CAPÍTULO II

Elementos de análise combinatória

Este capítulo explica as noções básicas da análise combinatória e desenvolve os fundamentos probabilísticos correspondentes; a última parte descreve algumas técnicas analíticas simples. Para a leitura deste livro não é necessário conhecer muita análise combinatória e portanto, os leitores, que não têm um interesse especial no assunto, deverão, logo que possível, passar para o Cap. V, onde a linha teórica principal, do Cap. I, é retomada. Talvez o melhor seja ler as seções individuais do presente capítulo conjuntamente com os tópicos de capítulos posteriores, que com elas se relacionam.

No estudo de jogos de azar simples, de procedimentos amostrais, de problemas de ordenação e ocupação, etc., estaremos, usualmente, lidando com espaços amostrais finitos, nos quais a mesma probabilidade é atribuída a todos os pontos. Para calcular a probabilidade de um evento A, temos então que dividir o número de pontos amostrais em A ("casos favoráveis") pelo número total de pontos amostrais ("casos possíveis"). Isto é facilitado pelo uso sistemático de algumas regras que iremos rever agora. Com o objetivo de simplificar e de economizar trabalho usaremos umas poucas técnicas já padronizadas, em lugar de descrevermos, para cada caso particular, o método computacional mais curto e mais adequado[1].

1. PRELIMINARES

Pares. Com m elementos $a_1, \ldots, a_m$ e n elementos $b_1, \ldots, b_n$ é possível formar mn pares (a_j, b_k) que contêm um elemento de cada grupo.

Demonstração. Disponha os pares numa tabela retangular, com a forma de uma tabela de multiplicação, tendo m linhas e n colunas de modo que (a_j, b_k) aparece na interseção da j-ésima linha com a k-ésima coluna. Nessas condições cada par irá aparecer uma e uma só vez e a afirmação torna-se óbvia. ►

Exemplos. a) *Cartas no jogo de Bridge* (veja nota de rodapé no Cap. I). Como conjuntos de elementos considere os quatro naipes e os treze valores das cartas respectivamente. Cada carta fica determinada pelo seu naipe e pelo seu valor e existem $4 \cdot 13 = 52$ combinações possíveis, isto é, 52 cartas.

b) "*Quebra-luzes com sete graduações*". Esse tipo de quebra-luz de chão, podendo ser operado de sete maneiras diferentes, contém três lâmpadas comuns e um acessório de luz indireta que pode operar em três níveis ou ser mantido desligado. Cada uma dessas quatro possibilidades pode ser combinada com o

[1]O leitor interessado encontrará muitos tópicos de análise combinatória elementar tratados no livro texto clássico *Choice and Chance*, de autoria de W. A. Whitworth, quinta edição, Londres (1901), reimpresso por G. E. Stechert, New York, 1942. O livro do mesmo autor, *DCC exercises*, reimpresso em New York, 1945, contém 700 problemas com soluções completas.

amostra da população de 26 letras. Desde que sejam permitidas repetições, existirão 26^{10} palavras desse tipo. Por outro lado, numa gráfica as letras existem não apenas teoricamente mas, também fisicamente sob a forma de tipos. Por simplicidade vamos admitir que a gráfica disponha exatamente de mil tipos para cada letra. Para compor uma palavra o técnico deverá escolher dez tipos e nessa situação são excluídas as repetições. Uma palavra pode portanto ser composta de $(26\,000)_{10}$ maneiras diferentes. Este número é praticamente igual a $(26\,000)^{10}$ e é superior a 10^{44}.

c) O Sr. e a Sra. Smith constituem uma amostra de tamanho dois, retirada da população humana; ao mesmo tempo eles formam uma amostra de tamanho um, retirada da população de todos os casais. Esse exemplo mostra que o tamanho da amostra é definido somente com relação a uma dada população. Efetuar r lançamentos de uma moeda é uma forma de obter uma amostra de tamanho r da população constituída pelas duas letras C e $\overline{C}$. Essa mesma disposição de r letras C e $\overline{C}$ é um único ponto amostral no espaço associado ao experimento que consiste no lançamento de uma moeda r vezes.

d) *Comentário sobre ordenação e amostragem na prática.* Quando estudamos o hábito de fumar numa população, por amostragem, sentimos intuitivamente que a ordem dentro da amostra não deverá ser relevante, o que leva o principiante a considerar amostras não-ordenadas. Entretanto só é possível tirarmos conclusões a partir de uma amostra, se nos basearmos em certas pressuposições probabilísticas e para estas, é necessário dispormos de um modelo adequado para o experimento teórico de obtenção de uma amostra. Um experimento desse tipo envolve obviamente escolhas que podem ser diferenciadas uma da outra, isto é, escolhas que são rotuladas de alguma forma. Do ponto de vista teórico o mais simples é usarmos os inteiros como índices e isso implica numa ordenação da amostra. Na prática, outros procedimentos podem ser preferíveis mas, mesmo a referência ao "terceiro indivíduo entrevistado pelo pesquisador na terça-feira" constitui uma forma de indexação. Em outras palavras, mesmo que no fim a ordem dentro das amostras seja abandonada, o experimento teórico envolve amostras ordenadas e veremos agora como isso irá afetar a atribuição adequada de probabilidades. ▶

Ao retirarmos sucessivamente r elementos de uma população de tamanho n estamos realizando um experimento cujos resultados possíveis são amostras de tamanho r. O número dessas amostras é n^r ou $(n)_r$ dependendo de estarmos ou não repondo os elementos retirados. Em qualquer caso, o nosso experimento teórico fica descrito por um espaço amostral no qual cada ponto individual representa uma amostra de tamanho r.

Até agora não falamos sobre a atribuição de probabilidades às nossas amostras. Usualmente atribuímos probabilidades iguais a todas elas dizendo então que estamos considerando amostras casuais. A palavra "casual" não está bem definida mas, quando aplicada a amostras ou seleções ela tem um significado único. O termo *escolha casual* é usado para implicar que todos os resultados são igualmente prováveis. Analogamente, sempre que falarmos de *amostras casuais de tamanho fixo r, o adjetivo casual implica que todas as possíveis amostras têm a mesma probabilidade*, a saber, n^{-r} em amostragem com reposição e $1/(n)_r$ em amostragem sem reposição, n denotando o tamanho da população de onde foi retirada a amostra. Se n é grande e r é relativamente pequeno o quociente $(n)_r/n^r$

está próximo de um. Isso nos leva a esperar que, para populações grandes e amostras relativamente pequenas, as duas formas de amostragem sejam praticamente equivalentes (veja os Probs. 11.1 e 11.2 e o Prob. 35 de VI, 10.

Introduzimos anteriormente uma terminologia prática, mas, ainda não discutimos a aplicabilidade do nosso modelo de amostragem casual à realidade. O lançamento de moedas, dados e outras atividades similares, podem ser interpretados como experimentos de amostragem casual com reposição e as nossas probabilidades se aproximam dos valores observados das freqüências, em experimentos de longa duração, apesar de não existirem dados e moedas perfeitamente balanceados. Uma amostragem casual sem reposição pode ser exemplificada através da retirada sucessiva de cartas de um baralho (desde que o ato de embaralhar seja realizado de maneira mais cuidadosa do que a usual). Na amostragem de populações humanas, o estatístico encontra freqüentemente dificuldades consideráveis e imprevisíveis e a amarga experiência tem mostrado que é difícil obtermos uma aproximação, mesmo grosseira, de casualidade.

Exercício. Na amostragem sem reposição a probabilidade de que um elemento fixado da população seja incluído numa amostra casual de tamanho r é

$$1-\frac{(n-1)_r}{(n)_r}=1-\frac{n-r}{n}=\frac{r}{n}.$$

Na amostragem com reposição a probabilidade de que um elemento seja incluído, pelo menos uma vez, é $1-(1-1/n)^r$.

3. EXEMPLOS

Os exemplos desta seção são casos especiais do problema seguinte. Uma amostra casual de tamanho r, com reposição, é retirada de uma população de n elementos. Desejamos obter a probabilidade de que nenhum elemento apareça duas vezes na amostra, isto é, de que a nossa amostra pudesse também ter sido obtida por amostragem sem reposição. O último teorema mostra que existem ao todo n^r amostras distintas, das quais $(n)_r$ satisfazem a condição estipulada. Admitindo-se que todos os arranjos tenham a mesma probabilidade, conclui-se que *a probabilidade de não existirem repetições na nossa amostra é*

$$p=\frac{(n)_r}{n^r}=\frac{n(n-1)\cdots(n-r+1)}{n^r}. \tag{3.1}$$

A seguir daremos algumas interpretações práticas dessa fórmula, que irão revelar aspectos surpreendentes.

a) *Números obtidos por amostragem casual.* Considere a população formada pelos dez algarismos 0,1, ..., 9. Toda seqüência de cinco algarismos representa uma amostra de tamanho $r=5$ e admitiremos que a cada seqüência desse tipo esteja associada a probabilidade 10^{-5}. Segue-se de (3.1) que *a probabilidade de que cinco algarismos aleatórios consecutivos sejam todos distintos é* $p=(10)_5\cdot 10^{-5}=0{,}3024$.

É de se esperar intuitivamente que, em tabelas matemáticas grandes, com um grande número de casas decimais, os últimos cinco algarismos sejam aproximadamente aleatórios. (Observe que nas tábuas comuns de logaritmos e em muitas outras, a diferença tabular é aproximadamente constante e, portanto, o último algarismo varia de forma regular). Num experimento foram selecionadas tabelas

r. O número de subpopulações de tamanho r é, portanto, dado por $(n)_r/r!$. Expressões desse tipo são conhecidas pelo nome de *coeficientes binomiais* e a notação usual para elas é

$$\binom{n}{r} = \frac{(n)_r}{r!} = \frac{n(n-1)\cdots(n-r+1)}{1\cdot 2\cdots(r-1)\cdot r}. \tag{4.1}$$

Acabamos de provar o

Teorema 1. *Uma população de n elementos possui* $\binom{n}{r}$ *subpopulações distintas de tamanho* $r \le n$.

Em outras palavras, um subconjunto de r elementos pode ser escolhido de $\binom{n}{r}$ maneiras diferentes. Um subconjunto como esse fica determinado pelos $n-r$ elementos que *não* pertencem a ele; esses elementos, por sua vez, formam uma subpopulação de tamanho $n-r$. Segue-se que o número de subpopulações de tamanho r é exatamente igual ao número de subpopulações de tamanho $n-r$ e, portanto, para $1 \le r \le n$ devemos ter

$$\binom{n}{r} = \binom{n}{n-r}. \tag{4.2}$$

(4.2) pode ser provada diretamente se observarmos que uma maneira alternativa de escrevermos o coeficiente (4.1) é

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}. \tag{4.3}$$

[Isto segue se multiplicarmos o numerador e o denominador de (4.1) por $(n-r)!$] Observe que o lado esquerdo de (4.2) não é definido para $r = 0$, mas, o lado direito é. Com o objetivo de tornar a Eq. (4.2) válida para todo r, $0 \le r \le n$, inteiro, *definiremos agora*

$$\binom{n}{0} = 1, \quad 0! = 1 \quad e \quad (n)_0 = 1. \tag{4.4}$$

Exemplos a) *Bridge e Pôquer* (conforme a primeira nota de rodapé do Cap. I). Por convenção, não levaremos em conta a ordem das cartas e, portanto, existem $\binom{52}{13} = 635\,013\,559\,600$ mãos distintas no bridge e $\binom{52}{5}$ = 2598960 mãos no pôquer. Vamos calcular a probabilidade, x, de que uma mão de pôquer tenha cinco cartas com valores diferentes. Os valores das cartas podem ser escolhidos de $\binom{13}{5}$ maneiras e correspondendo a cada carta nós podemos escolher um dos 4 naipes. Segue-se que $x = 4^5 \cdot \binom{13}{5}\Big/\binom{52}{5}$, o que dá aproximadamente 0,5071. Para o bridge, a probabilidade de termos 13 cartas com valores diferentes é $4^{13}\Big/\binom{52}{13}$ o que dá, aproximadamente, 0.0001057.

b) Nos Estados Unidos, cada um dos cinqüenta estados é representado por dois senadores. Suponhamos que uma comissão de cinqüenta senadores é escolhida ao acaso e consideremos os eventos: (1) um determinado estado está representado na comissão, (2) todos os estados estão representados.

No primeiro caso é melhor calcularmos a probabilidade, q, do evento complementar, isto é, de que o estado em questão *não* esteja representado. Existem cem senadores dos quais 98 não são do estado considerado. Portanto

$$q = \binom{98}{50}\Big/\binom{100}{50} = \frac{50\cdot 49}{100\cdot 99} = 0{,}24747\cdots$$

Em seguida, usando o teorema da Sec. 2, vemos que a comissão que inclui um senador de cada estado, pode ser escolhida de 2^{50} maneiras diferentes. Portanto a probabilidade de que *todos* os estados sejam incluídos na comissão é $p = 2^{50}\Big/\binom{100}{50}$. Usando a fórmula de Stirling (veja a Sec. 9), pode-se mostrar que $p \approx \sqrt{2\pi}\cdot 5\cdot 2^{-50} \approx 4{,}126\cdot 10^{-14}$.

c) *Um problema de ocupação.* Vamos considerar uma vez mais uma distribuição casual de r bolas em n compartimentos. (Isto é, cada um dos n^r arranjos possíveis tem probabilidade n^{-r}). Para determinar a *probabilidade*, p_k, *de que um compartimento especificado, contenha exatamente k bolas* ($k = 0, 1, 2, \ldots, r$) observar que as k bolas podem ser escolhidas de $\binom{r}{k}$ maneiras e que as restantes $r-k$ bolas, podem ser colocadas, nos $n-1$ compartimentos restantes, de $(n-1)^{r-k}$ maneiras. Segue-se que

$$p_k = \binom{r}{k}\cdot\frac{1}{n^r}(n-1)^{r-k} = \binom{r}{k}\cdot\frac{1}{n^k}\cdot\left(1-\frac{1}{n}\right)^{r-k}. \tag{4.5}$$

Este é um caso especial da chamada *distribuição binomial* que será estudada no Cap. VI. Valores numéricos podem ser encontrados na Tab. IV.3.

A distinção que se faz entre elementos distinguíveis e não-distinguíveis, tem similaridades com a relação que existe entre uma subpopulação e as amostras ordenadas correspondentes. Se abandonarmos todos os índices num arranjo (ou agrupamento) de r elementos, $a_1, \ldots, a_r$, obteremos um arranjo de r letras indistinguíveis. Reciprocamente, dado um agrupamento desse último tipo, uma enumeração arbitrária das r letras irá produzir um agrupamento das letras $a_1, \ldots, a_r$. Esse procedimento nos dará $r!$ arranjos diferentes desde que, é claro, toda troca de a_i com a_k seja contada como um novo arranjo. Os exemplos seguintes mostram como esse princípio pode ser aplicado a situações, nas quais os elementos a_k são identificados apenas parcialmente.

Exemplos. d) *Bandeiras de uma ou duas cores.* Vimos no Exemplo f, Sec. 1 que r bandeiras podem ser hasteadas em n mastros de $N = n(n+1)\cdots(n+r-1)$ maneiras diferentes. Vamos agora considerar o mesmo problema para bandeiras de uma mesma cor — impossíveis de serem diferenciadas. Se enumerarmos as bandeiras, cada exposição irá produzir $r!$ exposições de r bandeiras distinguíveis e, *portanto, r bandeiras de uma mesma cor podem ser expostas de N/r! maneiras.*

Suponha, a seguir, que p das bandeiras sejam vermelhas — e indistinguíveis — e q sejam azuis ($p + q = r$). É fácil ver que toda exposição de r bandeiras claro que o número de distribuições distinguíveis é igual ao número de maneiras segundo as quais podemos escolher r lugares, num total de $n + r - 1$, isto é, $A_{r,n}$.

A condição de que nenhum compartimento permaneça vazio impõe a restrição de que não existam dois traços adjacentes. Os r asteriscos determinam $r - 1$ espaços, dos quais $n - 1$ devem ser ocupados por traços; temos assim $\binom{r-1}{n-1}$ escolhas e a afirmação fica demonstrada. ▶

Exemplos. a) Existem $\binom{5+r}{5}$ resultados distinguíveis no lançamento de r dados indistinguíveis.

b) *Derivadas parciais.* As derivadas parciais de ordem r de uma função analítica $f(x_1, \ldots, x_n)$, de n variáveis, não dependem da ordem de diferenciação mas, sim do número de vezes que cada variável aparece. Dessa forma, a cada variável corresponde um compartimento e, portanto, existirão $\binom{n+r-1}{r}$ *derivadas parciais distintas de r-ésima ordem.* Uma função de três variáveis tem quinze derivadas de quarta ordem e 21 derivadas de quinta ordem. ▶

Consideremos agora n inteiros fixados que satisfazem (5.1). O número de distribuições de r bolas em n compartimentos, que resultam nos números de ocupação, $r_1, \ldots, r_n$, é dado pelo Teor. 4.2. Se admitirmos que todas as n^r possíveis distribuições são igualmente prováveis, segue-se que *a probabilidade de obtermos os números de ocupação, $r_1, \ldots, r_n$, dados é igual a*

$$\frac{r!}{r_1!r_2!\cdots r_n!}\, n^{-r}. \tag{5.3}$$

Essa atribuição de probabilidades foi usada em todas as aplicações consideradas até agora e era ponto passivo que ela era inerente à noção intuitiva de casualidade. Argumentos probabilísticos ou intuitivos não sugerem qualquer outra alternativa para essa atribuição de probabilidades. Do ponto de vista metodológico, é portanto interessante notar que a *experiência* obrigou os físicos a substituirem a distribuição (5.3) por outras que, originalmente, eram surpreendentes para a intuição. Esse assunto será discutido na próxima subseção. [Na física, (5.3) é conhecida pelo nome de distribuição de Maxwell-Boltzmann].

Em várias situações, é necessário irmos um pouco mais além e considerarmos os próprios compartimentos como indistinguíveis; isto implica em que não levemos em conta a ordem entre os números de ocupação. O exemplo seguinte serve para explicar um método rotineiro de solução para problemas que surgem dessa maneira.

Exemplo. c) Configurações de $r = 7$ bolas em $n = 7$ compartimentos. (Os compartimentos podem ser interpretados como dias da semana, as bolas como telefonemas, cartas, acidentes, etc.) Para sermos mais explícitos vamos considerar as distribuições com números de ocupação, 2, 2, 1, 1, 1, 0, 0, *aparecendo numa ordem arbitrária.* Esses sete números de ocupação induzem uma partição dos *sete compartimentos* em três subpopulações (categorias) que consistem, respectivamente, em dois compartimentos duplamente ocupados, três com uma única bola e dois vazios. Uma partição desse tipo em três grupos de tamanho 2, 3 e 2 pode ser efetuada de $7! \div (2! \cdot 3! \cdot 2!)$ maneiras. A cada associação particular entre os nossos números de ocupação e os sete compartimentos, correspondem $7! \div (2!2!1!1!1!0!0!) = 7! \div (2!2!)$ distribuições diferentes de $r = 7$ bolas pelos sete compartimentos. Segue-se que *o total das distribuições cujos números de ocupação são.* 2,2,1,1,1,0,0, *em alguma ordem é*

$$\frac{7!}{2!\,3!\,2!} \times \frac{7!}{2!\,2!}. \tag{5.4}$$

É fácil ver que este resultado foi deduzido através de *uma dupla aplicação* de (4.7), especificamente a bolas e a compartimentos. O mesmo resultado pode ser derivado e reescrito de várias maneiras, mas o presente método dá uma técnica rotineira bastante simples para uma grande variedade de problemas. (Veja os Probs. 43-45 da Sec. 10). A Tab. II. 1 nos dá os valores análogos a 5.4) e às probabilidades de todas as configurações possíveis de números de ocupação, no caso $r = n = 7$.

Tabela 1 - II

Distribuição ao acaso de 7 bolas em 7 compartimentos

Números de ocupação	Número de distribuições iguais a $7! \times 7!$ dividido por	Probabilidade (número de distribuições dividido por 7^7)
1, 1, 1, 1, 1, 1, 1	$7! \times 1!$	0,006 120
2, 1, 1, 1, 1, 1, 0	$5! \times 2!$	0,128 518
2, 2, 1, 1, 1, 0, 0	$2!\,3!\,2! \times 2!\,2!$	0,321 295
2, 2, 2, 1, 0, 0, 0	$3!\,3! \times 2!\,2!\,2!$	0,107 098
3, 1, 1, 1, 1, 0, 0	$4!\,2! \times 3!$	0,107 098
3, 2, 1, 1, 0, 0, 0	$2!\,3! \times 3!\,2!$	0,214 197
3, 2, 2, 0, 0, 0, 0	$2!\,4! \times 3!\,2!\,2!$	0,026 775
3, 3, 1, 0, 0, 0, 0	$2!\,4! \times 3!\,3!$	0,017 850
4, 1, 1, 1, 0, 0, 0	$3!\,3! \times 4!$	0,035 699
4, 2, 1, 0, 0, 0, 0	$4! \times 4!\,2!$	0,026 775
4, 3, 0, 0, 0, 0, 0	$5! \times 4!\,3!$	0,001 785
5, 1, 1, 0, 0, 0, 0	$2!\,4! \times 5!$	0,005 355
5, 2, 0, 0, 0, 0, 0	$5! \times 5!\,2!$	0,001 071
6, 1, 0, 0, 0, 0, 0	$5! \times 6!$	0,000 357
7, 0, 0, 0, 0, 0, 0	$6! \times 7!$	0,000 008

a) *Estatísticas de Fermi-Dirac e Bose-Einstein*

Vamos considerar um sistema mecânico constituído por r partículas indistingüíveis. Em mecânica-estatística, é usual subdividirmos o espaço de fase num grande número, n, de pequenas regiões ou compartimentos de tal forma que cada partícula esteja contida em um dos compartimentos. Dessa forma, o estado global do sistema é descrito em termos de uma distribuição casual de r partículas em n compartimentos. À primeira vista poderia parecer que, (pelo menos com uma definição adequada dos n compartimentos), todas as n^r distribuições deveriam ter a mesma probabilidade. Se isso ocorresse os físicos se refereriam *às estatísticas de Maxwell-Boltzmann* (o termo "estatística" é usado aqui num sentido peculiar à física). Numerosas tentativas foram feitas, no sentido de demonstrar que partículas físicas se comportam de acordo com as estatísticas de Maxwell-Boltzmann mas, a teoria moderna tem mostrado, sem dúvida alguma, que essas estatísticas *não se aplicam a nenhuma partícula conhecida*; em nenhum caso as n^r distribuições apresentaram, ainda que aproximadamente, a mesma probabilidade. Dois modelos probabilísticos distintos vêm sendo usados e cada um deles descreve, de maneira tude: $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_a$ e $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_b$. Em seguida juntamos os dois conjuntos numa única seqüência ordenada segundo a ordem crescente de magnitude. Um caso extremo é aquele no qual, todos os alfas precedem todos os betas e isso pode ser tomado como uma indicação da existência de uma diferença significante entre os dois tratamentos ou, entre as duas populações. Por outro lado, se os dois tratamentos forem idênticos os alfas e betas deverão aparecer numa ordem aproximadamente casual. Wald e Wolfwitz[11] mostraram que, em muitos casos, a teoria das seqüências pode ser aplicada, com vantagens, para detectar pequenas diferenças sistemáticas. (Um exemplo ilustrativo, que é porém analisado de forma diferente, será apresentado em III.1b).

Muitos problemas sobre seqüências podem ser resolvidos de maneira extremamente simples. Dados a alfas indistinguíveis e b betas também indistinguíveis, sabemos do Exemplo e, Sec. 4, que existem $\binom{a+b}{a}$ arranjos distintos desses elementos. Se existirem n_1 α-seqüências, o número de β-seqüências é necessariamente um dos números $n_1 - 1, n_1, n_1 + 1$. Distribuir os a alfas em n_1 seqüências é equivalente a distribuí-los em n_1 compartimentos, sem que existam compartimentos vazios. Pelo último lema isto pode ser feito de $\binom{a-1}{n_1-1}$ maneiras diferentes. Segue-se, por exemplo, que existem $\binom{a-1}{n_1-1}\binom{b-1}{n_1}$ arranjos com n_1 α-seqüências e $n_1 - 1$ β-seqüências. [Continua nos Probs. 20-25 da Sec. 11].

c) Na física a teoria das seqüências é utilizada no estudo de fenômenos cooperativos. Na teoria de Ising sobre reticulados a uma dimensão a energia depende do número de vizinhos diferentes, isto é, do número de seqüências.

6. DISTRIBUIÇÃO HIPERGEOMÉTRICA

Um grande número de problemas combinatórios poderá ser reduzido à forma seguinte. Numa população de n elementos, n_1 são vermelhos e $n_2 = n - n_1$ são pretos. Um grupo de r elementos é escolhido ao acaso. O que se procura é a probabilidade q_k, de que o grupo escolhido contenha exatamente k elementos vermelhos. Neste caso, k é um inteiro qualquer compreendido entre zero e o menor dos números n_1 e r.

Para determinar q_k, observe que o grupo escolhido deverá conter k elementos vermelhos e $r-k$ elementos pretos. Os vermelhos podem ser escolhidos de $\binom{n_1}{k}$ maneiras diferentes e os pretos de $\binom{n-n_1}{r-k}$ maneiras. Desde que todo grupo de k elementos vermelhos pode ser combinado com qualquer grupo de pretos, temos que

$$(6.1) \qquad q_k = U \cdot \frac{\binom{n_1}{k}\binom{n-n_1}{r-k}}{\binom{n}{k}}.$$

[11]A. Wald e J. Wolfowitz, *On a test whether two samples are from the same population*, Ann. Math. Statist., Vol 2 (1940) pp. 147-162.

A distribuição de probabilidades assim definida se denomina *distribuição hipergeométrica*[12]. Usando (4.3) é possível reescrevermos (6.1) na forma

$$(6.2) \qquad q_k = \frac{\binom{r}{k}\binom{n-r}{n_1-k}}{\binom{n}{n_1}}.$$

Observação. As probabilidades q_k são definidas somente para k não superior a r ou a n_1 mas, desde que $\binom{a}{b} = 0$ sempre que $b > a$, as fórmulas (6.1) e (6.2) nos dão $q_k = 0$ se $k > n_1$ ou $k > r$. Segue-se, portanto, que as definições (6.1) e (6.2) podem ser usadas para todo $k \geq 0$, desde que a relação $q_k = 0$ seja interpretada como uma impossibilidade.

Exemplos. a) *Inspeção de qualidade.* No controle de qualidade industrial, lotes de tamanho n são submetidos à inspeção para verificação de sua qualidade. As peças defeituosas desempenham o papel dos elementos "vermelhos". Seu número total, n_1, é, naturalmente, desconhecido. Uma amostra de tamanho r é retirada do lote determinando-se então o número k de peças defeituosas na amostra. A fórmula (6.1) permite então que se façam inferências sobre a provável magnitude de n_1; este é um problema típico de estimação estatística mas, está fora dos objetivos deste livro.

b) No Exemplo b, Sec. 4, a população consiste em $n = 100$ senadores dos quais $n_1 = 2$ representam o estado fixado (são "vermelhos"). Um grupo de $r = 50$ senadores é escolhido ao acaso. Este grupo poderá incluir $k = 0, 1$, ou 2 do estado considerado. Relembrando (4.4) e usando (6.2) temos

$$q_0 = q_2 = \frac{50 \cdot 49}{100 \cdot 99} = 0{,}24747\ldots, \qquad q_1 = \frac{50}{99} = 0{,}50505\ldots$$

No Exemplo 4.b o valor de q_0 foi obtido de forma diferente.

c) *Estimação do tamanho de uma população animal a partir de dados de recaptura.*[13] Suponha que mil peixes sejam capturados em um lago, marcados com pontos vermelhos e em seguida devolvidos ao lago. Depois de algum tempo um novo lote de mil peixes é capturado e constata-se que cem dentre eles têm pontos vermelhos. Que conclusões podem ser tiradas com relação ao número total de peixes do lago? Este é um problema típico de *estimação estatística.* Não caberia aqui uma descrição dos vários métodos que um estatístico moderno poderia usar mas, vamos mostrar de que maneira a distribuição hipergeométrica pode ajudar na solução do problema. Admitiremos naturalmente que os dois lotes possam ser considerados como amostras casuais da população de todos os peixes

[12]O nome é justificado pelo fato da função geradora (veja o Cap. IX) de $\{q_k\}$ poder ser expressa em termos de funções hipergeométricas.

[13]Quando esse exemplo foi apresentado na primeira edição, não sabíamos que esse método era grandemente utilizado na prática. As mais recentes contribuições à literatura, incluem N. T. J. Bailey, *On estimating the size of mobile populations from recapture data*, Biometrika, Vol 38 (1951), pp. 293-306 e D. G. Chapman, *Some properties of the hypergeometric distribution with applications to zoological sample censuses*, University of California Publications in Statistics, Vol 1 (1951), pp. 131-160.

Duas situações desse tipo serão discutidas explicitamente: (i) *A colocação de bolas continua até que, pela primeira vez, uma bola seja colocada num compartimento já ocupado.* O processo termina quando ocorre a primeira duplicação de bolas num compartimento. (ii) *Um compartimento é fixado (por exemplo, o primeiro compartimento) e o processo de colocação de bolas continua durante todo o tempo em que esse compartimento permanece vazio.* Nesse caso, o processo termina no instante em que uma bola é colocada no compartimento fixado.

Algumas interpretações possíveis para esse modelo servirão para elucidar o problema.

Exemplos. a) *Aniversários.* No exemplo do aniversário (Sec. 3d), aos 365 dias do ano correspondem os compartimentos e às pessoas correspondem as bolas. O nosso modelo (1) pode ser interpretado agora da seguinte forma: se selecionarmos pessoas ao acaso, uma a uma, quantas pessoas devemos selecionar para que existam duas pessoas com o mesmo dia de aniversário? O modelo (ii) corresponde a esperar até que o *meu* aniversário apareça na amostra.

b) *O problema das chaves.* Um indivíduo deseja abrir uma porta. Ele dispõe de n chaves das quais somente uma abre a porta. Devido a razões que não vêm ao caso, o homem experimenta as chaves aleatoriamente de tal forma que, em cada ensaio, cada chave tenha probabilidade $1/n$ de ser escolhida e todos os possíveis resultados, que envolvem o mesmo número de tentativas, são igualmente prováveis. Qual é a probabilidade de que o homem abra a porta exatamente na r-ésima tentativa? Esse é um caso especial do modelo (ii). É interessante comparar essa procura aleatória da chave com uma tentativa mais sistemática. (Prob. 11 da Sec. 10; veja também o Prob. 5 da Sec. 8, Cap. V).

c) No exemplo anterior, a amostragem das chaves pode ser substituída por uma amostragem de uma população arbitrária. Uma situação como essa ocorre com o *colecionador de figurinhas.* Nesse caso é também possível perguntarmos quando se espera a primeira duplicação ou, quando um elemento fixado irá aparecer pela primeira vez.

d) *Moedas e dados.* No exemplo do Cap. I, Sec. 5a, uma moeda é lançada até que apareça uma cara. Este é um caso particular do modelo (ii), no qual $n = 2$. Quando um dado é lançado até que o número 1 apareça pela primeira vez, temos novamente o modelo (ii) com $n = 6$. (Outros tempos de espera são discutidos nos Probs. 10.21, 10.22, 10.36 e 11.12).

Discutiremos inicialmente o modelo (i) que é mais fácil de ser visualizado. É conveniente usarmos símbolos da forma $(j_1, j_2, \ldots, j_r)$ para indicar que a primeira, a segunda, ..., a r-ésima bolas foram colocadas nos compartimentos de número $j_1, j_2, \ldots, j_r$ e que o processo terminou com a colocação da r-ésima bola. Isso significa que $j_1, j_2, \ldots, j_{r-1}$ são inteiros distintos, compreendidos entre 1 e n, e que j_r coincide com um deles. Cada arranjo desse tipo representa um ponto amostral. Os únicos valores possíveis para r são $2, 3, \ldots, n+1$, uma vez que a primeira duplicação não poderá ocorrer antes da colocação da segunda bola e nem após a colocação da $(n+1)$-ésima bola. A conexão existente entre o presente problema e o anterior, da colocação de um número fixo de bolas em n compartimentos nos leva a atribuir a cada ponto amostral, $(j_1, \ldots, j_r)$, que envolva exatamente r bolas, a probabilidade n^{-r}. A seguir mostraremos que essa convenção é permissível (isto é, que a soma dessas probabilidades é igual a 1) e que ela conduz a resultados razoáveis.

Para r fixado, o conjunto de todos os pontos amostrais $(j_1, \ldots, j_r)$ representa *o evento de que o processo termina com a colocação da r-ésima bola.* Segue-se, de (2.1), que os números $j_1, \ldots, j_{r-1}$ podem ser escolhidos de $(n)_{r-1}$ maneiras diferentes; para j_r existem $r-1$ escolhas possíveis pois, o seu valor é um dos números $j_1, \ldots, j_{r-1}$. Segue-se que *a probabilidade de que o processo termine com a colocação da r-ésima bola é*

$$q_r = \frac{(n)_{r-1} \cdot (r-1)}{n^r} = \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{r-2}{n}\right) \cdot \frac{r-1}{n}, \tag{7.1}$$

com $q_1 = 0$ e $q_2 = 1/n$. *A probabilidade de que o processo dure mais do que r passos é* $p_r = 1 - (q_1 + \cdots + q_r)$ ou $p_1 = 1$ e

$$p_r = \frac{(n)_r}{n^r} = \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{r-1}{n}\right), \tag{7.2}$$

como pode ser verificado por indução. Em particular, $p_{n+1} = 0$ e $q_1 + \cdots q_{n+1} = 1$, como era de se esperar. Além disso, para $n = 365$, (7.2) se reduz a (3.2) e, em geral, o novo modelo nos leva aos mesmos resultados quantitativos obtidos pelo modelo anterior, que envolvia um número fixo de bolas.

O modelo (ii) difere do (i) pelo fato de depender de *um espaço amostral infinito.* As seqüências $(j_1, \ldots, j_r)$ estão agora sujeitas à condição de que $j_1, \ldots, j_{r-1}$ são diferentes de um número $a \leq n$, prefixado, mas, j_r deve ser igual a a. Mais ainda, não há razão alguma, *a priori*, para que o processo deva terminar. Para r fixado, atribuiremos outra vez a cada ponto $(j_1, \ldots, j_r)$, a probabilidade n^{-r}. Para cada um dos números $j_1, \ldots, j_{r-1}$ temos $n-1$ escolhas possíveis e nenhuma para j_r uma vez que seu valor é prefixado. *A probabilidade de que o processo termine na r-ésima etapa*, será portanto dada por

$$q_r^* = \left(\frac{n-1}{n}\right)^{r-1} \cdot \frac{1}{n}, \qquad r = 1, 2, \ldots \tag{7.3}$$

Observando que essa é uma série geométrica, obtemos $q_1^* + q_2^* + \cdots = 1$. Dessa forma a soma das probabilidades é igual a 1 e não há, portanto, necessidade de introduzirmos um ponto amostral, para representar a possibilidade de que nenhuma bola seja colocada no compartimento fixado. *A probabilidade* $p_r^* = 1 - (q_1^* + q_2^* + \cdots + q_r^*)$, *de que o processo dure mais do que r etapas é dada por*

$$p_r^* = \left(1 - \frac{1}{n}\right)^r, \qquad r = 1, 2, \ldots \tag{7.4}$$

como era de se esperar.

A mediana da distribuição $\{p_r\}$ é definida como sendo o valor de r para o qual temos $p_1 + \cdots + p_{r-1} \leq 1/2$ mas, $p_1 + \cdots + p_r > 1/2$; a probabilidade de que o processo ultrapasse a mediana é mais ou menos a mesma de que ele pare antes de atingi-la. [No Exemplo d, Sec. 3, dos aniversários, a mediana é igual a 23]. Para calcular a mediana de $\{p_r\}$ tomamos logaritmos de ambos os lados, como fizemos em (3.4). Quando r é pequeno, em comparação com n, é fácil ver que $-\log p_r$ vale aproximadamente $r^2/2n$. Segue-se que *a mediana* de $\{p_r\}$ está próxima de $\sqrt{2n \log 2}$ ou aproximadamente $6/5\sqrt{n}$. É interessante observar que a mediana cresce com a raiz quadrada do tamanho da população. Contrastando com essa situação, *a mediana de* $\{p_r^*\}$ *vale aproximadamente* $n \log 2$ ou $0{,}7\,n$