3.2. Variáveis aleatórias

3.2.1. Modelos *determinísticos* X *probabilísticos*

Um *modelo* é uma representação simplificada da realidade. Uma miniatura de um avião, feita por um engenheiro para estudar a aerodinâmica de um projeto num túnel de vento, é um exemplo de *modelo físico* (o modelo é um objeto que existe fisicamente, no mundo real). Contudo, na maioria das vezes as ciências procuram construir modelos mais abstratos, principalmente modelo baseados em Matemática. Quando dizemos que a Terra tem a forma de uma esfera e que sua órbita é uma elipse, por exemplo, estamos usando dois modelos *geométricos* para descrever a realidade. Modelos também podem ser equações que mostrem com duas ou mais variáveis se relacionam.

Modelos matemáticos podem ser *determinísticos* ou *probabilísticos*. Quando Galileu estudou o problema de queda livre, por exemplo, verificou que o tempo de queda de um corpo não está relacionado à sua massa, mas sim à altura da queda h e à aceleração da gravidade g; podemos descrever esta relação através da equação $t = \sqrt{2h/g}$. Esta equação é um modelo matemático do tipo *determinístico*, que mostra como as variáveis *tempo* e *altura* estão associadas. Neste modelo, o valor da variável t é *determinado* pelos valores da variável h e da constante g: se conhecemos g e h poderemos, antes de fazer o experimento, prever qual será o tempo t de queda. A Física clássica trabalha sempre com modelos determinísticos, como por exemplo estes bem conhecidos:

$$V=RI \qquad\qquad F=ma \qquad\qquad e = e_0 + v_0 t + \frac{at^2}{2} \qquad\qquad E = mc^2$$

Por outro lado, suponha um problema de Biologia: um pesquisador estuda a reprodução de uma espécie animal e quer descobrir o número de filhotes nascidos em cada ninhada. Ou então, um problema de Engenharia: um engenheiro estuda o tráfego em um bairro e quer descobrir o número de carros que passam por minuto no cruzamento de duas ruas. Não poderemos escrever uma fórmula para *determinar* ou *prever* o número de filhotes, ou de carros; o que podemos fazer é considerar o números de filhotes ou de carros como *variáveis aleatórias*, e procurar calcular as probabilidades de seus valores por meio de modelos *probabilísticos*.

3.2.2. Variáveis aleatórias e seus modelos

"Variáveis aleatórias" (VAs) são aquelas que provém da realização de *experimentos aleatório*s (seção **3.1.4**). Estas variáveis podem ser *discretas* ou *contínuas*. Variáveis aleatórias discretas (VADs) resultam de experimentos que têm espaço amostral discreto, finito ou infinito; variáveis aleatórias contínuas (VACs), de experimentos que têm espaço amostral contínuo (a diferença entre espaços amostrais *discretos* e *contínuos* é mostrada na seção **3.1.4.1**). Variáveis aleatórias e seus modelos, é claro, são conceitos matemáticos - ou seja, são abstrações teóricas. Podem servir (ou não) para descrever os resultados de observações ou experimentos científicos; isto vai depender de vários fatores, mas principalmente do grau de aleatoriedade presente nestes resultados. Se uma professora *escolhe* um aluno na sala de aulas e pergunta a sua idade, esta idade pode ser considerada um exemplo de variável aleatória? Provavelmente não, porque faltou aí o elemento de "aleatoriedade". Vários fatores podem ter influenciado a escolha da professora, mas provavelmente não poderemos considerar que sejam aleatórios. Se por outro lado a professora *sorteia* um aluno, usando algum tipo de roleta, e pergunta a sua idade, este número provavelmente poderá ser tratado como uma VA.

O número de filhotes em cada ninhada, ou o número de carros em um cruzamento, nos problemas mencionados acima, frequentemente são tratados como variáveis aleatórias discretas; em geral, VADs são usadas como modelos em problemas nos quais os resultados provém de contagens (no exemplo, do número de filhotes ou de carros). As variáveis aleatórias contínuas, por outro lado, servem de modelos para variáveis provenientes de medições, como por exemplo o peso ou comprimento de cada filhote, ou o tempo decorrido entre a passagem de dois carros. A distinção entre VADs e VACs é bem clara em termos teóricos mas, na prática, às vezes é uma questão de mera conveniência (voltaremos a este assunto na seção **3.4.1**).

Nas seções a seguir, iremos com frequência usar exemplos nos quais as variáveis surgem de lançamentos de dados, moedas ou de outros jogos. Por que estes jogos são sempre usados como exemplos nos livros que tratam de probabilidades? Porque são problemas nos quais podemos garantir que existe aleatoriedade, e nos quais as probabilidades básicas (probabilidade de cada face da moeda ou do dado, por exemplo) são fáceis de estabelecer *a priori*. Estes jogos são versões simplificadas de vários problemas que ocorrem com grande frequência na realidade.

Variáveis aleatórias são sempre numéricas, mesmo que o espaço amostral original não seja. Por exemplo, se lançamos duas moedas, o resultados destes lançamentos podem ser (C: cara, K: coroa):

CC CK KC KK

Se contamos as caras, obtemos uma VAD, cujos valores são:

2 1 1 0

Um mesmo experimento pode dar origem a várias VAs diferentes. Se lanço três dados, por exemplo, algumas das VAs que podem ser geradas por este experimento são: o somatório dos números mostrados nos três dados; a quantidade de dados que mostram números pares; a quantidade de dados que mostram a face 4; a quantidade de dados que mostram faces maiores que 4, etc.

Variáveis aleatórias de qualquer tipo têm seu comportamento regido por leis probabilísticas, e podemos criar para ela modelos *probabilísticos*. Estes modelos nada mais são, em essência, do que fórmulas, tabelas, ou gráficos que nos permitem calcular as probabilidades de cada valor, ou de cada conjunto de valores, de uma VA; estes modelos não podem prever exatamente que valor a variável irá assumir, mas pelo menos poderá dizer quais são os valores ou conjuntos de valores mais prováveis e quais os menos prováveis.

Nas próximas seções, estudaremos alguns dos modelos mais úteis; começaremos pelos modelos para VADs, porque seus conceitos básicos são mais fáceis compreender do que os das VACs (que exigem algumas noções de Cálculo).