3.1. Probabilidades

3.1.1. Introdução e definições

3.1.1.1. Por que estudar Probabilidades?

Este capítulo trata de um assunto que, a rigor, não pertence à Estatística: a *Teoria das Probabilidades*, uma área da Matemática pura. Probabilidade e Estatística porém estão intimamente relacionadas; as probabilidades são a base das técnicas de *Inferência Estatística*, que nos permitem generalizar as conclusões obtidas a partir de amostras. Por exemplo, a Inferência responde à pergunta: se um medicamento foi testado numa grupo de pacientes e obteve bons resultados na maioria deles, qual é a probabilidade de que este bom resultado tenha acontecido por mero acaso, e de que o medicamento na verdade não funcione? Se esta probabilidade for muito pequena, consideraremos que é mais provável que o medicamento funcione; o resultado será considerado uma *evidência* a favor da eficácia do medicamento. Se porém a probabilidade de o resultado ter sido obtido por acaso for razoavelmente grande, consideraremos que este resultado não servirá como evidência. Note que as conclusões da Inferência nunca são *certas*; nunca apontam qual *é* a resposta correta de um problema, mas sim qual é a resposta que *tem maior probabilidade* de ser correta. Não é possível, por isso, fazer Inferência Estatística sem usar probabilidades.

A Teoria das Probabilidades pode ser estudada de forma abstrata, como qualquer outra área da Matemática. Contudo, se quisermos usar seus resultados em aplicações no mundo real, quase sempre iremos precisar do auxilio da Estatística. Uma agência de seguros, por exemplo, usa o cálculo de probabilidades para determinar o valor que um motorista deve pagar anualmente pelo seguro contra roubo de seu carro; para isto, porém, precisa de saber qual é a probabilidade daquele tipo de carro ser roubado, e esta probabilidade só pode ser estimada por meio de técnicas de Estatística.

A maioria dos cursos de Estatística no nível de graduação são na verdade cursos de *Probabilidade e Estatística*. Alguns livros-texto deixam isto bem claro no título (como por exemplo, *Noções de Probabilidade e Estatística* [[1]]); outros não mencionam as probabilidades no título, mas quase sempre incluem alguns capítulos sobre o assunto (por exemplo, *Estatística Básica* [[2]] tem uma seção de mais de 150 páginas sobre Probabilidades).

Na duas seções seguintes (**3.1.1.2** e **3.1.1.3**), discutiremos rapidamente o conceito de "probabilidade" (um daqueles conceitos na Matemática, como *ponto*, *reta* e *plano*, que parecem intuitivos, mas são extremamente difíceis, ou mesmo impossíveis, de definir rigorosamente). Da Seção **3.1.2** em diante mostraremos algumas técnicas básicas de cálculo de probabilidades.

3.1.1.2. A idéia de *acaso*

Nas ciências naturais, há eventos cuja ocorrência ou não-ocorrência são *previsíveis*; se as condições necessárias estão presentes, o evento *sempre* ocorre (os chamados eventos *certos*), ou então *nunca* ocorre (os eventos *impossíveis*). Por exemplo, sob pressão atmosférica normal e temperatura de 40 C, água se encontra sempre em estado líquido, nunca no estado sólido. Na Física e na Química clássicas, os fenômenos estudados são deste tipo; o que estas ciências procuram fazer é descrever o conjunto de condições necessárias para a ocorrência de cada evento, e criar modelos matemáticos que permitam prever esta ocorrência, se as condições estiverem presentes.

Existem contudo nas ciências e aplicações práticas eventos que, dado um conjunto de condições, podem ocorrer ou não, de forma imprevisível. Estes são os chamados *eventos aleatórios*, cujo resultado é determinado pelo acaso. As ciências da vida estão cheias de exemplos deste tipo de evento. Na Biologia, por exemplo: se cruzarmos variedades diferentes de uma espécie de plantas (como Gregor Mendel, que fez experimentos cruzando ervilhas verdes e ervilhas amarelas), não podemos prever qual será o resultado. (Os experimentos de Mendel, considerado o fundador da Genética moderna, serão vistos na seção 4.4.4). Na Medicina, é impossível prever com certeza se um paciente vai ou não reagir a um tratamento, ou se irá sofrer de algum efeito colateral; além disso, é muitas vezes impossível determinar com certeza se o paciente tem ou não uma doença a partir dos testes diagnósticos existentes. A incerteza é uma constante na Medicina (que foi por isso chamada de "Ciência da Incerteza" pelo médico canadense William Osler). A incerteza também está presente em várias áreas da vida prática, como nas finanças e nos seguros: não é possível prever com certeza se um investimento será ou não rentável, ou se é necessário ou não fazer um seguro de algum bem.

A *Teoria das Probabilidades* é o ramo da Matemática que estuda os *experimentos aleatórios.* Um exemplo clássico destes experimentos é o lançamento de um dado. (A palavra "aleatório". aliás, vem da palavra latina *alea*, que significa "dado de jogar"). Se lançarmos um dado, obtemos como resultado um número inteiro entre 1 e 6; se repetirmos o experimento algumas vezes, veremos que os resultados variam de forma imprevisível de um lançamento para o outro, e que, aparentemente, esta variação não obedece a nenhum padrão.

A idéia de que a Matemática seja usada para estudar estes experimentos aleatórios pode parecer contraditória, á primeira vista. Primeiro, porque se dizemos que um experimento é "aleatório", aceitamos que seus resultados são imprevisíveis; se são imprevisíveis, que sentido tem estudá-los? Segundo, porque a Matemática com seu rigor lógico e sua certeza, parece ser em tudo a antítese da idéia de aleatoriedade.

Na verdade, a idéia de probabilidade é bem difícil de entender e de definir formalmente, e talvez por isso a Teoria das Probabilidades seja um ramo da Matemática de desenvolvimento relativamente recente, que não tem antecedentes na Antigüidade (como tem a Geometria) ou na Idade Média (como tem a Álgebra). Como outros ramos da Matemática, porém, também surgiu de especulações de pesquisadores sobre problemas práticos que costumavam ser resolvidos de maneira empírica.

O conceito de "probabilidade" nasceu da tentativa de analisar de forma numérica a experiência prática dos jogadores com os jogos de cartas e de dados, muito populares na Europa. Estes jogadores sabiam que não é possível prever, nem aproximadamente, os resultados de cada rodada de um jogo de azar; contudo, depois de passarem grande parte de suas vidas analisando os resultados de experimentos aleatórios (isto é, jogando dados e cartas), percebiam que existe um padrão (uma *regularidade estatística*) que só se revela a longo prazo, depois que um número muito grande de resultados for analisado. Até certo ponto, este conhecimento é intuitivo. Por exemplo, se os jogadores A e B disputam um jogo de *cara-ou-coroa*, lançando uma moeda e apostando quantidades iguais em suas faces opostas, é intuitivo prever que este jogo será "equilibrado", e que ambos os jogadores terão a mesma probabilidade de ganhar. As palavras "equilibrado" e "mesma probabilidade" expressam a nossa crença de que o jogo, depois de repetido por algum tempo, provavelmente terminará empatado, já que os ganhos de cada jogador serão mais ou menos equivalentes à suas perdas. Suponhamos agora que estes jogadores façam apostas iguais sobre lançamentos de um dado. O jogador A ganha se a face mostrada for "1", o jogador B ganha se a face mostrada for qualquer outra, diferente de "1". É fácil prever, neste caso, que o jogo será *desequilibrado*, e que o jogador A provavelmente perderá a maioria das suas apostas.

Na prática, porém, os jogos de azar feitos em cassinos são muito mais elaborados do que isto, e suas regras incluem complicações que tornam mais difícil a previsão dos resultados a longo prazo. Uma destas complicações é usar dois dados, ao invés de um. Por exemplo, suponha que dois jogadores lançam um par de dados. Se a soma dos números mostrados pelos dados for 7, ganha o jogador A; se for 4 ou 11, ganha o jogador B; se for qualquer outra, os dados serão lançados de novo. Se os jogadores fazem apostas iguais a cada rodada, qual será o resultado provável deste jogo, a longo prazo? Aqui, a intuição não é suficiente para nos orientar. Jogadores muito experientes provavelmente seriam capazes de prever que o jogador A tem a cada rodada uma probabilidade levemente maior de ganhar do que o jogador B, e que esta maior probabilidade se traduziria, a longo prazo, num maior lucro acumulado. Faltavam contudo meios de justificar logicamente estas observações empíricas, pois não havia ainda ferramentas matemáticas para analisar os eventos de forma quantitativa.

Um jogador, o Chevalier de Méré, pode ser considerado um dos responsáveis pelo nascimento da Teoria da Probabilidades, pois foi ele quem, em 1654, apresentou alguns problemas sobre apostas em jogos de azar a seu amigo Blaise Pascal. Pascal se interessou pelo assunto, e começou a discuti-lo, por cartas, com Pierre de Fermat (o maior matemático francês daquela época). A correspondência entre os dois lançou as bases do estudo formal de probabilidades, mas não foi publicada na época.

O primeiro livro sobre o assunto foi publicado por De Moivre, em 1711, mas o autor mais influente nos primórdios do estudo das probabilidades foi Laplace, que publicou seu primeiro artigo sobre isto em 1774. No início, foram estudados apenas problemas relacionados a jogos de azar, que são uma classe de experimentos onde a idéia de "probabilidade" e "acaso" podem ser aplicadas de forma mais evidente. Durante o Iluminismo, filósofos como d'Alembert e Euler começaram a ampliar o campo de aplicação desta teoria de forma a abranger problemas em áreas de maior interesse prático ou científico, como problemas sobre expectativa de vida, loterias e seguros.

De certa forma, estes problemas também são muito similares aos jogos de azar. Um seguro contra roubo de um carro, por exemplo, nada mais é que uma aposta feita contra a companhia seguradora: se uma pessoa faz o seguro, está apostando que seu carro *será* roubado, enquanto a seguradora assume a posição contrária, e aposta que o carro *não será* roubado. Parece portanto muito natural falarmos da "probabilidade" deste roubo ocorrer (da mesma forma que falamos da probabilidade de uma *cara* aparecer no lançamento de uma moeda), e aplicarmos as mesmas técnicas de cálculo desenvolvidas para os jogos de cassino. A diferença, em relação aos jogos de azar, é que as probabilidades de que o segurado ganhe ou perca sua aposta não são tão fáceis de se estimar (como estimar a probabilidade de que um carro seja roubado?).

Esta extensão das aplicações do cálculo de probabilidades criou muitas dificuldades teóricas, e exigiu modificações e refinamentos na própria definição do que seja "probabilidade"; discutiremos brevemente esta evolução na próxima seção.

3.1.1.3. A definição de probabilidade

(i) Definição clássica

A razão pela qual os livros-texto continuam a usar como exemplo os problemas relacionados a jogos de azar é bastante simples: neles, é fácil fazer pressuposições iniciais sobre as probabilidades. Por exemplo, podemos dizer que a probabilidade de se obter uma cara no lançamento de uma moeda é ½, porque a moeda é um objeto simétrico, e não há nada que nos leve a crer que uma face deve ser mais provável do que a outra (Fig. **1A**).

Da mesma forma, no lançamento de um dado (Fig. **1B**), podemos supor que qualquer de suas faces tem a mesma probabilidade das outras, e portanto cada face tem probabilidade igual a $^1/_6$; na retirada aleatória de uma carta de um baralho completo (52 cartas), supomos que todas as cartas são igualmente prováveis, e atribuímos probabilidade $^1/_{52}$ para cada uma.

A. Moeda inglesa de uma libra esterlina B. Par de dados

Figura 1. Cara ou coroa, dados de jogar

Estes lançamentos de dados e moedas, ou retiradas de cartas de baralho, são chamados de *experimentos aleatórios.* Um *experimento*, neste contexto, é qualquer ação que dê um resultado observável. Um *experimento aleatório* é aquele que, se repetido várias vezes sob as mesmas condições, dá a cada vez resultados definidos mas imprevisíveis.

Esta pressuposição de que entre os diversos resultados possíveis de um experimento aleatório nenhum deles deve ser mais provável que os outros (são portanto *equiprováveis* ou *igualmente possíveis*), está na origem da definição "clássica" de probabilidade, que Pascal e Fermat aceitaram tacitamente, e que foi formulada explicitamente por Laplace [3]:

> **Definição 1:**
> A "probabilidade" é a razão entre o número de resultados favoráveis e o número total de resultados equiprováveis possíveis em um experimento aleatório.

A probabilidade de obtermos uma face com número *par* no lançamento de um dado, por exemplo, é dada por

P(face par) = número de faces com números pares / número total de faces = 3 / 6 = 1 / 2

A probabilidade de conseguirmos um ás, ao retirarmos uma carta aleatoriamente de um baralho completo, é dada por

P(um ás) = número de ases no baralho / número total de cartas no baralho = 4 / 52 = 1 / 13

Em lançamentos de moedas, dados ou roletas, a equiprobabilidade é uma suposição que parece lógica, já que estes objetos são simétricos. (Nos dados e roletas, aliás, a equiprobabilidade é exigida por lei. Como estes objetos são usados por cassinos em jogos que envolvem dinheiro, há fiscais que os testam regularmente, para garantir que sejam equilibrados.) A definição clássica é por isso freqüentemente usada na prática em problemas que envolvem jogos ou sorteios.

Em termos teóricos, porém, esta definição apresenta dois problemas. Primeiro, ela só é útil numa classe muito restrita de problemas. Quando o cálculo de probabilidades passou a ser aplicado a problemas fora da esfera dos jogos de azar, começou a ficar claro que uma definição baseada na contagem de resultados *equiprováveis* era limitada demais.

Há algumas poucas aplicações em que ainda podemos supor uma equiprobabilidade aproximada. Por exemplo, não é absurdo supor que, no nascimento de seres humanos, a probabilidade de nascer um menino seja *aproximadamente* igual a de nascer uma menina (mas não *exatamente* igual; veremos isto na seção **4.4.3**). Na maior parte das outras aplicações úteis, contudo, os resultados possíveis de um experimento não serão nem aproximadamente equiprováveis. Por exemplo, um problema importante na área de saúde é o de testar a qualidade de um exame de laboratório como ferramenta para apoiar o diagnóstico de uma doença. Para isto, é preciso estimar a probabilidade de que ele dê resultado *positivo* quando testado em um paciente sabidamente doente (esta probabilidade é chamada de *sensibilidade* do teste). Só há dois resultados possíveis neste teste: ou ele dá positivo (o que é um resultado *correto*, já que o paciente está doente), ou dá negativo (um resultado *incorreto*). É claro que a probabilidade de o teste dar resultado dar *positivo* deve ser muito maior do que a de ele dar *negativo*; quanto mais próxima de 1 for ser esta probabilidade, melhor será o teste. Se os resultados positivo e negativo fossem *equiprováveis*, usar o teste como auxílio no diagnóstico seria tão útil quando lançar uma moeda e decidir com base no *cara-ou-coroa*.

O segundo problema da definição clássica é um problema técnico: do ponto de vista da Lógica, esta definição é *tautológica*. Uma definição é "tautológica" se o conceito a ser definido aparece em ambos os membros da definição; no caso, o conceito de "probabilidade" está explícito no membro esquerdo, mas também está implícito no membro direito:

Probabilidade = n^o. de resultados favoráveis / n^o. de resultados equiprováveis

pois a palavra "equiprovável" evidentemente quer dizer "de mesma probabilidade". Se ainda não sabemos o que é "probabilidade" (é o que estamos tentando definir!), não podemos usar o conceito de "mesma probabilidade" do lado direito. Fazer isto seria o mesmo que definir "quilograma" como "*quilograma* é peso de um objeto que pesa um quilograma"; ou definir "sal" como "*sal* é um tempero que tem gosto salgado".

Várias outras tentativas de definir probabilidade começaram a ser feitas a partir do final do século XIX; veremos a seguir as três que serão mais importantes neste curso.

(ii) Definição "freqüencista" da probabilidade

A definição clássica nos fornece um método para a atribuição de valores numéricos às probabilidade em alguns problemas, mas nada nos diz sobre a interpretação empírica ou a utilidade prática destes valores. Ela nos diz, por exemplo, que a probabilidade de um jogador conseguir a soma 5 no lançamento de um par de dados é 4/36 = 0,111 e a de conseguir a soma 8 é de 5/36 = 0,139 (veremos como calcular isto na seção **3.1.3**). Que significam estes números, na prática, para um jogador? A definição não diz, mas tanto jogadores quanto matemáticos sabiam, por experiência ou intuição, que estas probabilidades representam um valor limite para o qual tende a *frequência relativa* destas somas; ou seja, que se os jogadores lançassem o par de dados um grande número de vezes, veriam a soma 5 aparecer em aproximadamente 11,1 % dos lançamentos, e a soma 8 aparecer em 13,9 % deles. Portanto, o jogador que apostasse na soma 8 teria uma pequena vantagem (13,9 − 11,1 = 2,8 %) sobre o que apostasse na soma 5, *se o jogo fosse repetido muitas vezes*. É importante frisar este ponto: ninguém pode prever o que acontecerá no lançamento de um par de dados, ou prever se o resultado será a soma 5 ou a soma 8; contudo, é possível prever que, a longo prazo, se o jogo for repetido muitas vezes, a soma 8 aparecerá mais freqüentemente do que a soma 5.

Esta é a idéia básica da Teoria de Probabilidades : o resultado de cada repetição de um experimento aleatório é imprevisível, mas o resultado combinado de um grande número de repetições do experimento tende a seguir um padrão (uma *regularidade estatística*) que é determinado por sua *probabilidade*. É isto que nos permite fazer previsões para grandes números de repetições; podemos prever, com razoável precisão, qual a proporção de meninas, entre as crianças que nascerão numa cidade no ano que vem; ou qual a proporção de pacientes que serão curados quando um novo tratamento for aplicado. Não podemos porém fazer previsões para casos individuais – prever o sexo de uma determinada criança que vai nascer, ou prever se um determinado paciente terá sucesso com o novo tratamento. (Os médicos sempre podem errar, pois a Medicina não é uma das Ciências Exatas; também erram os investidores, e as companhias de seguros - por isto existem as companhias de *re-seguro*, que fazem o seguro das companhias de seguro -, e todos aqueles que trabalham em áreas que sejam caracterizadas por uma grande incerteza,)

Alguns matemáticos do século XIX sugeriram, a partir destas idéias, uma definição *frequencista* de "probabilidade": a probabilidade de um resultado é o valor teórico para o qual tende a freqüência relativa deste resultado, a medida que aumenta o número de repetições do experimento:

Definição 2:
Se após n repetições de um experimento (n suficientemente grande), se observam h repetições de um determinado evento, então a "probabilidade" do evento é h/n.

Esta definição reflete o que fazemos na prática. Se queremos determinar a probabilidade de nascimento de *meninas*, por exemplo, examinarmos registros de nascimentos de um grande número de crianças; se numa amostra de 10.000 nascimentos ocorridos durante um ano encontramos 4.900 meninas, podemos usar a frequência relativa = 4.900/10.000 = 0,49 como uma estimativa da probabilidades de nascimentos de meninas (Esta estimativa será aliás bastante precisa, já que o número de repetições do "experimento" é grande; veremos na seção **4.9** como calcular também margem de erro desta estimativa). A frequência relativa é considerada neste caso "como uma medição experimental da probabilidade" [[4]].

Como definição, contudo, a idéia ainda não funcionou perfeitamente; que a freqüência relativa tende para um valor constante, que pode ser chamado de "probabilidade", é um fato empírico, observável por experimentação, mas não é possível usá-lo como base para uma teoria matemática; esta definição não gera teoremas que permitam (ou justifiquem os métodos existentes) para o cálculo de probabilidades.

(iii) Definição "subjetiva" da probabilidade

As duas definições anteriores referem-se a experimentos que podem ser repetidos (pelo menos teoricamente) um grande número de vezes, e servem para evidenciar a *regularidade estatística* presente nos resultado. Uma definição inteiramente diferente é a definição *subjetiva,* segundo a qual uma probabilidade não é mais que uma indicação da intensidade da crença, ou da evidência, em favor de uma afirmação:

Definição 3:
A "probabilidade" é um número entre 0 e 1 que expressa minha crença na ocorrência de um evento.

Por exemplo, se um astrônomo diz que há apenas 0,01 de probabilidade de haver vida em Marte, isto quer dizer que, segundo ele, é muito pouco provável que exista vida naquele planeta. Se digo que meu time tem probabilidade 0,90 de ganhar o Campeonato este ano, ou que um paciente tem 0,90 de probabilidade de responder bem ao tratamento, isto quer dizer que, na minha opinião, é quase certo que estes eventos vão ocorrer. Observe que, nestes exemplos, não existem indefinidas "repetições" do experimento aleatório: só existe um planeta Marte, um Campeonato este ano, e um paciente sendo tratado, não um conjunto infinito de Martes, Campeonatos e pacientes possíveis.

Esta definição é a base de uma subdivisão da Estatística, chamada de *bayesiana*, por se basear no *Teorema de Bayes*. Esta subdivisão tem aplicações importantes (por exemplo, em Medicina; veremos o teorema, e algumas aplicações, na Seção 5.8).

(iv) Definição "axiomática" da probabilidade

A definição mais comumente aceita hoje foi criada no século XX, a chamada definição *axiomática*. Basicamente, diz apenas que a probabilidade é um número, que tem algumas propriedades, definidas como axiomas; a partir destas propriedades, os matemáticos podem demonstrar os teoremas que serão usados na solução de problemas. A definição não dá nenhuma indicação da ligação que este número possa ter com os problemas no mundo real (ou seja, não diz para que pode servir este número, na prática); é uma definição inteiramente abstrata e formal, seguindo a tendência em que a Matemática moderna se desenvolveu a partir do século passado.

Definição 4:
Seja um E um experimento, e S o espaço amostral a ele associado. A cada evento A associaremos um número real representado por $P(A)$ e denominado "probabilidade de A", que satisfaça às seguintes propriedades:

1) $0 \leq P(A) \leq 1$
2) $P(S) = 1$

3) Se A e B forem eventos mutuamente excludentes, $P(A \cup B) = P(A) + P(B)$
4) Se $A_1, A_2, A_3, ..., A_n$ forem eventos mutuamente excludentes dois a dois, então
$$P(A_1 \cup A_2 \cup A_3 \cup ... \cup A_n) = P(A_1) + P(A_2) + P(A_3) + ... + P(A_n)$$

Esta definição e seus axiomas são hoje a base lógica para o cálculo das probabilidades. Nela são usados conceitos ("evento", "espaço amostral", "mutuamente excludentes") que ainda não foram definidos; veremos estes conceitos na seção **3.1.5**.

3.1.1.4. Conclusão

Esta breve discussão das mudanças nas definições serve para ilustrar quão complexos são as idéias de *aleatoriedade* e de *probabilidade*, e quantas maneiras diferentes podem ser usadas para abordá-las. Atualmente, os livros de Probabilidade podem ser separados quanto à maneira que abordam as definições: a maioria se baseia na definição axiomática (abordagem "clássica"), enquanto os outros se baseiam no definição subjetiva (abordagem "bayesiana"). Esta separação dá origem a duas correntes da Estatística que utilizam métodos bastante distintos entre si (clássica x bayesiana). Este livro, apesar de discutir brevemente o teorema de Bayes, se insere na corrente clássica; a corrente bayesiana não será estudada.

As definições *clássica* ou a *freqüencista* têm hoje, para os matemáticos, apenas valor histórico; na prática, contudo, quando ainda as usamos quando interpretamos probabilidades, em problemas no mundo real. Se por exemplo uma fábrica afirma que as peças que produz têm 0,01 de probabilidade de serem defeituosas, isto equivale a dizer que espera que em média 1 em cada 100 peças sejam defeituosas (o que é uma aplicação da definição *freqüencista*). Para estimar esta probabilidade de 0,01, porém, a fábrica provavelmente usou amostras aleatórias de sua linha de produção, retiradas de forma que cada peça produzida tenha a mesma probabilidade de ser escolhida do que qualquer outra peça; poderemos então calcular probabilidades supondo eqüiprobabilidade e usando a definição clássica.

Referências

[1] Magalhães, Marcos N. e Lima, Antônio C. P. de. *Noções de Probabilidade e Estatística*. São Paulo: EDUSP. 2015.
[2] Bussab, Wilton de O. e Morettin, Pedro A.. *Estatística básica*. São Paulo: Saraiva. 2013.
[3] Laplace. *Théorie Analytique des Probabilités*. Paris. 1812.
[4] Cramér, Harald. *Elementos da teoria das Probabilidades e algumas de suas aplicações*. São Paulo: Mestre Jou. 1973.