Capítulo 2

Teoria das Probabilidades

2.1 Introdução

No capítulo anterior, foram mostrados alguns conceitos relacionados à estatística descritiva. Neste capítulo apresentamos a base teórica para o desenvolvimento de técnicas estatísticas a serem utilizadas nos capítulos posteriores.

Vamos considerar as seguintes questões: Como saber se um determinado produto está sendo produzido dentro dos padrões de qualidade? Como avaliar a capacidade de um determinado exame acertar o verdadeiro diagnóstico? Questões como estas envolvem algum tipo de variabilidade ou incerteza, e as decisões podem ser tomadas por meio da teoria de probabilidades que permite a quantificação da incerteza.

A seguir, veremos alguns conceitos básicos de probabilidade.

2.2 Conceitos Básicos de Probabilidade

- Fenômeno Aleatório: É um processo de coleta de dados em que os resultados possíveis são conhecidos mas não se sabe qual deles ocorrerá. Assim, um fenômeno aleatório pode ser a contagem de ausências de um funcionário em um determinado mês, o resultado do lançamento de uma moeda, verificar o resultado de um exame de sangue, entre outros.

- Espaço Amostral: O conjunto de todos os resultados possíveis do fenômeno aleatório é chamado de espaço amostral. Vamos representá-lo por Ω.

Exemplo 2.1. Lançamento de uma moeda. $\Omega = \{$cara, coroa$\}$.

Exemplo 2.2. Lançamento de um dado. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Exemplo 2.3. Número de chips defeituosos em uma linha de produção durante 24 horas. $\Omega = \{0, 1, 2, 3, \ldots, n\}$, sendo n o número máximo de itens defeituosos.

Exemplo 2.4. Tempo de reação de uma pomada anestésica aplicada em queimados. $\Omega = \{t \in \Re \mid t \geq 0\}$.

- Evento: Qualquer subconjunto do espaço amostral Ω é chamado de evento. Serão representados por letras maiúsculas A, B, ... Dentre os eventos podemos considerar o evento união de A e B, denotado por $A \cup B$, que, equivale à ocorrência de A, ou de B, ou de ambos. A ocorrência simultânea dos eventos A e B, denotada por $A \cap B$ é chamada de evento interseção. Dois eventos A e B dizem-se mutuamente exclusivos ou disjuntos, quando a ocorrência de um deles impossibilita a ocorrência do outro. Os dois eventos não têm nenhum elemento em comum, isto é, $A \cap B = \emptyset$ (conjunto vazio).

Exemplo 2.5. Suponha um fenômeno aleatório conduzido com a finalidade de se conhecer a eficiência de uma terapia na cura de uma síndrome. Para tanto, dois pacientes foram tratados com a referida terapia. Vamos representar C e $\overline{C}$, como curado e não curado, respectivamente. O espaço amostral nesse caso é dado por:

$$\Omega = \{CC,\ C\overline{C},\ \overline{C}C,\ \overline{CC}\}.$$

Considere os seguintes eventos: A "obter uma cura" e B "obter quatro curas": Sendo assim, temos:

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

e

$$B = \emptyset.$$

2.2.1 Definição clássica de probabilidade

Em fenômenos aleatórios tais como lançamento de uma moeda, de um dado, extração de uma carta de um baralho entre outros, temos que todos os resultados possíveis tem a mesma chance de ocorrer. Assim, por exemplo no lançamento de uma moeda a probabilidade do evento cara ou coroa ocorrer são igualmente prováveis, ou seja, a probabilidade atribuída a cada um é 1/2.

A probabilidade de um evento A qualquer ocorrer pode ser definida por:

$$P(A) = \frac{\text{número de casos favoráveis ao evento A}}{\text{número de casos possíveis}}.$$

Exemplo 2.6. Considere o fenômeno aleatório lançamento de um dado e o evento A "sair número par". Qual a probabilidade deste evento ocorrer?

$$P(A) = \frac{3}{6} = 0,50.$$

Na maioria das situações práticas, os resultados não têm a mesma chance de ocorrer, deste modo, a probabilidade dos eventos deve ser calculada pela frequência relativa.

2.2.2 Aproximação da Probabilidade pela frequência relativa

Quando não se tem conhecimento sobre as probabilidades dos eventos, estas podem ser atribuídas após repetidas observações do fenômeno aleatório, ou seja, a proporção de vezes que um evento A qualquer ocorre pode ser estimada como segue:

$$P(A) = \frac{\text{número de ocorrências do evento A}}{\text{número de observações}}.$$

Exemplo 2.7. Uma escola particular pretende oferecer um treinamento esportista aos seus alunos. Dos 300 alunos entrevistados, 142 optaram pelo voleibol, 125 indicaram o basquete e 35 indicaram o futebol. Selecionado aleatoriamente um desses alunos, qual a probabilidade de obter alguém que prefere o voleibol?

$$P(V) = \frac{142}{300} = 0,47333.$$

À medida que o número de observações aumenta, as aproximações tendem a ficar cada vez mais próximas da probabilidade efetiva.

2.2.3 Propriedades de probabilidades

É uma função P(.) que associa números reais aos elementos do espaço amostral e satisfaz as condições:

1. $0 \leq P(A) \leq 1$, para qualquer evento A;

2. P(Ω) = 1;

3. P($\emptyset$) = 0;

4. P($\cup_{j=1}^{n} A_j$) $= \sum_{j=1}^{n} P(A_j)$.

Se $\overline{A}$ for o evento complementar de A, então P(A) = 1 - P($\overline{A}$).

2.2.4 Teorema da soma

Dado dois eventos A e B, a probabilidade de pelo menos um deles ocorrer é igual a soma das probabilidades de cada um menos a probabilidade de ambos ocorrerem simultaneamente, ou seja:

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

Se A e B forem mutuamente exclusivos, teremos P(A$\cap$B) = 0. Assim,

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

Exemplo 2.8. Considere o experimento lançamento de um dado e os seguintes eventos:
A = {sair número 5},
B = {sair número par} e
C = {sair número ímpar}.
Determinar: Ω, P(A), P(B), P(C), P(A∪B), P(A∪C) e P($\overline{A}$).

$$\begin{aligned}
\Omega &= \{1,2,3,4,5,6\}. \\
P(A) &= \frac{1}{6}. \\
P(B) &= \frac{3}{6}. \\
P(C) &= \frac{3}{6}. \\
P(A \cup B) &= \frac{1}{6} + \frac{3}{6} = \frac{4}{6}. \\
P(A \cup C) &= \frac{1}{6} + \frac{3}{6} - \frac{1}{6} = \frac{3}{6}. \\
P(\overline{A}) &= 1 - \frac{1}{6} = \frac{5}{6}.
\end{aligned}$$

Exemplo 2.9. Um estudo realizado por uma empresa de recursos humanos mostrou que 45% dos funcionários de uma multinacional saíram da empresa porque estavam insatisfeitos com seus salários, 28% porque consideraram que a empresa não possibilitava o crescimento profissional e 8% indicaram insatisfação tanto com o salário como com sua impossibilidade de crescimento profissional. Considere o evento S: "o funcionário sai da empresa em razão do salário" e o evento I: "o funcionário sai da empresa em razão da impossibilidade de crescimento profissional". Qual é a probabilidade de um funcionário sair desta empresa devido a insatisfação com o salário ou insatisfação com sua impossibilidade de crescimento profissional?

$$\begin{aligned}
P(S \cup I) &= P(S) + P(I) - P(S \cap I) \\
P(S \cup I) &= 0,45 + 0,28 - 0,08 = 0,65.
\end{aligned}$$

2.2.5 Probabilidade condicional

Existem situações em que a chance de um particular evento acontecer depende do resultado de outro evento. A probabilidade condicional de A dado que ocorreu B pode ser determinada dividindo-se a probabilidade de ocorrência de ambos os eventos A e B pela probabilidade do evento B; como se mostra a seguir:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \; P(B) > 0.$$

Exemplo 2.10. Em uma universidade foi selecionada uma amostra de 500 alunos que cursaram a disciplina de Estatística. Entre as questões levantadas estava: Você gostou da disciplina de Estatística? De 240 homens, 140 responderam que sim. De 260 mulheres, 200 responderam que sim. Para avaliar as probabilidades podemos organizar as informações em uma tabela. maneira:

Tabela 2.1: Gosto pela disciplina de estatística segundo sexo.

	Gostou		
Sexo	Sim	Não	Total
Homem	140	100	240
Mulher	200	60	260
Total	340	160	500

Qual é a probabilidade de que um aluno escolhido aleatoriamente:

(a) H = Seja um homem?

$$P(H) = \frac{240}{500} = 0,48.$$

(b) G = Gostou da disciplina de Estatística?

$$P(G) = \frac{360}{500} = 0,72.$$

(c) M = Seja uma mulher?

$$P(M) = \frac{260}{500} = 0,52.$$

(d) NG = Não gostou da disciplina de Estatística?

$$P(NG) = \frac{160}{500} = 0,32.$$

(e) Seja uma mulher ou gostou da disciplina de Estatística.

$$P(M \cup G) = \frac{260}{500} + \frac{340}{500} - \frac{200}{500} = 0,80.$$

(f) Seja uma mulher e gostou da disciplina de Estatística.

$$P(M \cap G) = \frac{200}{500} = 0,40.$$

(g) Dado que o aluno escolhido gostou da disciplina de Estatística. Qual a probabilidade de que o aluno seja um homem?

$$P(H \mid G) = \frac{P(H \cap G)}{P(G)} = \frac{140}{340} = 0,41176.$$

(h) Dado que o aluno escolhido é uma mulher. Qual a probabilidade de que ela não gostou da disciplina de Estatística?

$$P(NG \mid M) = \frac{P(NG \cap M)}{P(M)} = \frac{60}{260} = 0,23077.$$

2.2.6 Teorema do produto

Da definição de probabilidade condicional $P(A|B) = \frac{P(A \cap B)}{P(B)}$ podemos obter o teorema do produto, que nos permite calcular a probabilidade da ocorrência simultânea de dois eventos. Sejam A e B eventos de Ω, a probabilidade de A e B ocorrerem juntos é dada por:

$$P(A \cap B) = P(A)\ P(B|A), \text{ com P(A)} > 0$$

ou

$$P(A \cap B) = P(B)\ P(A|B), \text{ com P(B)} > 0.$$

Dois eventos A e B são independentes quando a ocorrência de um não altera a probabilidade de ocorrência do outro. Desse modo,

$$P(A \cap B) = P(A)\ P(B).$$

Exemplo 2.11. Uma empresária sabe por experiência, que 65% das mulheres que compram em sua loja preferem sandálias plataformas. Qual é a probabilidade de as duas próximas clientes comprarem cada uma delas, uma sandália plataforma? Vamos admitir que o evento A "a primeira cliente compra uma sandália plataforma" e o evento B "a segunda cliente compra uma sandália plataforma". Então,

$$P(A \cap B) = (0,65)(0,65) = 0,4225.$$

2.2.7 Teorema da probabilidade total

Suponha que os eventos C_1, C_2, ..., C_n formam uma partição do espaço amostral. Os eventos não têm interseções entre si e a união destes é igual ao espaço amostral. Seja A um evento qualquer desse espaço, então a probabilidade de ocorrência desse evento será dada por:

$$P(A) = P(A \cap C_1) + P(A \cap C_2) + \cdots + P(A \cap C_n)$$

e usando a definição de probabilidade condicional,

$$P(A) = P(C_1)\ P(A|C_1) + P(C_2)\ P(A|C_2) + \cdots + P(C_n)\ P(A|C_n).$$

Exemplo 2.12. Uma caixa I contém 2 fichas verdes e 3 vermelhas. Uma segunda caixa II contém 4 fichas verdes e 3 vermelhas. Escolhe-se, ao acaso, uma caixa e dela retira-se, também ao acaso uma ficha. Qual a probabilidade de que a ficha retirada seja verde? Se denotarmos por I e II o evento caixa I e caixa II, respectivamente e V o evento a ficha é verde, temos: $P(I) = \frac{1}{2}$, $P(V|I) = \frac{2}{5}$, $P(II) = \frac{1}{2}$ e $P(V|II) = \frac{4}{7}$.

Desta forma, o evento V ("ficha verde") pode ser escrito em termos de interseções do evento V com os eventos I e II,

$$V = (V \cap I) \cup (V \cap II)$$

$$\begin{aligned} P(V) &= (P(V \cap I)) + (P(V \cap II) \\ &= P(I)(P(V|I)) + (P(II)P(V|II) \\ &= \frac{1}{2}\frac{2}{5} + \frac{1}{2}\frac{4}{7} = 0,48571. \end{aligned}$$

Exemplo 2.13. Um estabilizador pode provir de três fabricantes I, II e III com probabilidades de 0,25, 0,35 e 0,40, respectivamente. As probabilidades de que durante determinado período de tempo, o estabilizador não funcione bem são, respectivamente, 0,10; 0,05 e 0,08 para cada um dos fabricantes. Qual é a probabilidade de que um estabilizador escolhido ao acaso não funcione bem durante o período de tempo especificado. Se denotarmos por A o evento "um estabilizador não funcione bem" e por C_1, C_2 e C_3 os eventos "um estabilizador vem do fabricante I, II e III", respectivamente. A probabilidade de que um estabilizador escolhido ao acaso não funcione bem durante o período de tempo especificado é:

$$\begin{aligned} P(A) &= P(C_1)P(A|C_1) + P(C_2)P(A|C_2) + P(C_3)P(A|C_3) \\ &= (0,25)(0,10) + (0,35)(0,05) + (0,40)(0,08) = 0,07450. \end{aligned}$$

2.2.8 Teorema de Bayes

Considere $C_1, C_2, ..., C_n$ eventos que formam uma partição do espaço amostral Ω, cujas probabilidades são conhecidas. Considere que para um evento A se conheçam as probabilidades condicionais, desta forma:

$$P(C_j|A) = \frac{P(C_j)\ P(A|C_j)}{P(C_1)\ P(A|C_1) + P(C_2)\ P(A|C_2) + \cdots + P(C_n).\ P(A|C_n)},\ j = 1, 2, \ldots, n.$$

Exemplo 2.14. Considere o exemplo anterior para o desenvolvimento do Teorema de Bayes. Dado que o estabilizador escolhido ao acaso não funciona bem durante o período de tempo especificado, qual a probabilidade de que tenha sido produzido pelo fabricante I, isto é, P(C_1|A).

$$\begin{aligned} P(C_1|A) &= \frac{P(A)\ P(A|C_1)}{P(C_1)P(A|C_1) + P(C_2)P(A|C_2) + P(C_3)P(A|C_3)} \\ &= \frac{(0,25)(0,10)}{0,07450} = 0,33557. \end{aligned}$$