UNIDADE

5

Conceitos básicos de Probabilidade

Objetivo

Nesta Unidade, você vai compreender os conceitos de Probabilidade e a importância do uso do raciocínio probabilístico para auxiliar o administrador na tomada de decisões em ambiente de incerteza.

Probabilidade: conceitos gerais

Caro estudante!

Nesta Unidade, vamos estudar os conceitos básicos de Probabilidade, tais como experimento aleatório, espaço amostral e eventos, axiomas e propriedades, probabilidade condicional e independência estatística. Nos EUA, há uma anedota que diz: "as únicas coisas que são certas são a morte e os impostos". Em outras palavras, estamos imersos na incerteza, e os administradores diariamente precisam tomar decisões, muitas delas extremamente sérias, em um cenário de grande incerteza:

- lançamos ou não um novo modelo de automóvel?
- convertemos nossos fornos de óleo combustível para gás natural?
- qual será a reação do nosso público às mudanças na grade de programação?

Por que há incerteza? Porque a **variabilidade** inerente à natureza impede a compreensão completa dos fenômenos naturais e humanos. Mas os seres humanos precisam tomar decisões. Assim, é necessário levar a incerteza em conta no processo: alguns apelam para a sabedoria popular, outros para o místico. Os administradores precisam tomar decisões de forma objetiva, e surge, então, a Probabilidade como uma das abordagens de tratamento da incerteza.

A utilização de métodos probabilísticos proporciona um grande auxílio na tomada de decisões, pois permite avaliar riscos e otimizar recursos (sempre escassos) para as situações mais prováveis. Você está convidado a conhecer mais sobre esse tema nesta Unidade.

Nas Unidades 3 e 4, foi utilizado um raciocínio predominantemente indutivo. Os dados foram coletados, e, através da sua organização em distribuições de freqüências e medidas de síntese, foi possível caracterizar a **variabilidade*** do fenômeno observado, e elaborar hipóteses ou conjecturas a respeito.

Suponha que estejamos estudando o percentual de meninos e meninas nascidos em um Estado brasileiro. Consultando dados do IBGE, provenientes de censos e levantamentos anteriores (portanto, distribuições de freqüências da variável qualitativa sexo dos recém-nascidos), há interesse em prever qual será o percentual de nascimentos no ano de 2009: em suma, será usado um raciocínio dedutivo; a partir de algumas suposições sobre o problema (a definição dos resultados possíveis, os percentuais registrados em anos anteriores), tenta-se obter novos valores.

Se o percentual de meninos no passado foi de 49%, a pergunta é: qual será o percentual de meninos nascidos no ano de 2009? É possível que seja um valor próximo de 49%, talvez um pouco acima ou um pouco abaixo, mas não há como responder com certeza absoluta, pela simples razão que o fenômeno ainda não ocorreu e que sua natureza é aleatória, ou seja, é possível identificar quais serão os resultados possíveis (menino ou menina), e há uma certa regularidade nos percentuais de nascimentos (verificados anteriormente), mas não é possível responder qual será o resultado exato antes de o fenômeno ocorrer.

A regularidade citada (que foi observada para um grande número de nascimentos) permite que seja calculado o grau de certeza ou confiabilidade da previsão feita, que recebe o nome de **Probabilidade***. Haverá uma grande probabilidade de que realmente o percentual de meninos nascidos em 2009 seja de 49%, mas nada impede que um valor diferente venha a ocorrer.

Sem saber, montamos um **modelo probabilístico*** para o problema em questão:

- foram definidos todos os resultados possíveis para o fenômeno (experimento);
- definiu-se uma **regra** que permite dizer quão provável será cada resultado ou grupo de resultados.

GLOSSÁRIO

*Variabilidade – diferenças encontradas por sucessivas medições realizadas em pessoas, animais ou objetos, em tempos ou situações diferentes. Fonte: Montgomery (2002).

*Probabilidade – descrição quantitativa da certeza de ocorrência de um evento, geralmente representada por um número real entre 0 e 1 (0% e 100%). Fonte: elaborado pelo autor.

*Modelo probabilístico – modelo matemático para descrever a certeza de ocorrência de eventos, no qual são definidos os eventos possíveis e uma regra de ocorrência para calcular quão provável é cada evento ou conjunto de eventos. Fonte: Barbetta (2006).

O modelo probabilístico permite expressar o grau de incerteza através de probabilidades.

A regra citada foi definida a partir de observações anteriores do fenômeno, mas também poderia ser formulada com base em considerações teóricas. Por exemplo, se há interesse em estudar as proporções de ocorrências das faces de um dado, e se este dado não é viciado, espera-se que cada face ocorra em 1/6 do total de lançamentos: se o dado for lançado um grande número de vezes, isso provavelmente ocorrerá, mas um resultado diferente poderia ser obtido sem significar que o dado está viciado, principalmente se forem feitos poucos lançamentos.

Neste ponto, é importante ressaltar que os modelos probabilísticos não têm razão de ser para fenômenos (experimentos) não aleatórios (**determinísticos**): aqueles em que, usando teorias e fórmulas apropriadas, se pode prever exatamente qual será o seu resultado antes de o fenômeno ocorrer, por exemplo, o lançamento de uma pedra de 5 kg de uma altura de dez metros, havendo interesse em cronometrar o tempo para que ela atinja o chão. Conhecendo o peso da pedra, a altura do lançamento, a aceleração da gravidade e as leis da Física, é perfeitamente possível calcular o tempo de queda, não há necessidade sequer de realizar o experimento.

Para construir ou utilizar modelos probabilísticos, é necessário que haja um grande número de realizações do fenômeno (experimento) para que uma regularidade possa ser verificada: é a Lei dos Grandes Números. No início do século XX, o estatístico inglês Karl Pearson lançou uma moeda não viciada 24.000 vezes (!) para verificar a validade dessa lei: obteve 12.012 caras, praticamente o valor esperado (12.000, 50%).

Para prosseguirmos, precisamos de algumas definições importantes para estudar os modelos probabilísticos. Precisamos definir exatamente as condições em que devemos usar modelos probabilísticos, e isso exige saber o que são experimento aleatório, espaço amostral e eventos. Vamos ver?

Definições prévias

Experimento aleatório

Experimento aleatório é um processo de obtenção de um resultado ou uma medida que apresenta as seguintes características:

- não se pode afirmar, antes de fazer o experimento, qual será o resultado de uma realização, mas é possível determinar o conjunto de resultados possíveis; e
- quando é realizado um grande número de vezes (replicado), apresentará uma regularidade que permitirá construir um modelo probabilístico para analisar o experimento.

São experimentos aleatórios:

a) o lançamento de um dado e a observação da face voltada para cima; não se sabe exatamente qual face vai ocorrer, apenas que será uma das seis, e que, se o dado for não viciado, e o lançamento, imparcial, todas as faces têm a mesma chance de ocorrer;

b) a observação dos diâmetros, em mm, de eixos produzidos em uma metalúrgica; sabe-se que as medidas devem estar próximas de um valor nominal, mas não se sabe exatamente qual é o diâmetro de cada eixo antes de efetuar as mensurações; e

c) o número de mensagens que são transmitidas corretamente por dia em uma rede de computadores; sabe-se que o mínimo possível é zero, mas não se sabe nem sequer o número máximo de mensagens que serão transmitidas.

Todo experimento aleatório terá alguns resultados possíveis, que constituirão o espaço amostral.

Espaço amostral (S ou Ω)

Espaço amostral é o conjunto de **todos** os resultados possíveis de um experimento aleatório. Para cada experimento aleatório, haverá um espaço amostral único Ω associado a ele.

Neste primeiro exemplo, veremos alguns experimentos aleatórios com os respectivos espaços amostrais:

a) o lançamento de um dado e a observação da face voltada para cima: $\Omega = \{1, 2, 3, 4, 5, 6\}$;

b) a retirada de uma carta de um baralho comum (52 cartas) e a observação do naipe: Ω = {copas, espadas, ouros, paus};

c) o número de mensagens que são transmitidas corretamente por dia em uma rede de computadores: $\Omega = \{0, 1, 2, 3, ...\}$;

d) a observação do diâmetro, em mm, de um eixo produzido em uma metalúrgica: Ω = {D, tal que D > 0}; e

e) as vendas mensais, em unidades, de determinado modelo de veículo: $\Omega = \{0, 1, ...\}$

Note que não há um limite superior conhecido, mas somente é possível a ocorrência de valores inteiros.

Não há um limite superior, e, teoricamente, pode haver uma infinidade de valores.

O espaço amostral pode ser:

- **finito**, formado por um número limitado de resultados possíveis, como nos casos a e b;
- **infinito numerável**, formado por um número infinito de resultados, mas que podem ser listados, como nos casos c ou e; ou
- **infinito**, formado por intervalos de números reais, como no caso d.

Um espaço amostral é dito **discreto** quando ele for finito ou infinito enumerável; é dito **contínuo** quando for infinito, formado por intervalos de números reais.

A construção do modelo probabilístico dependerá do tipo de espaço amostral, como será visto mais adiante.

Eventos

Eventos são quaisquer subconjuntos do espaço amostral. Um evento pode conter um ou mais resultados; se pelo menos um dos resultados ocorrer, o evento ocorre! Geralmente, há interesse em calcular a probabilidade de que um determinado evento venha a ocorrer, e este evento pode ser definido de forma verbal, precisando ser "traduzido" para as definições da Teoria de Conjuntos, que veremos a seguir.

Embora nem todos os autores concordem com esta abordagem, ela auxilia bastante na compreensão dos conceitos.

Sejam o experimento aleatório lançamento de um dado não viciado e observação da face voltada para cima: o seu espaço amostral será $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Definindo três eventos:

$E_1 = \{2, 4, 6\}$,

$E_2 = \{3, 4, 5, 6\}$ e

$E_3 = \{1, 3\}$

serão apresentadas as definições de **evento união**, **evento intersecção**, **eventos mutuamente exclusivos** e **evento complementar**.

Evento **união** de E_1 com E_2 ($E_1 \cup E_2$): evento que ocorre se E_1 ou E_2 ou ambos ocorrem.

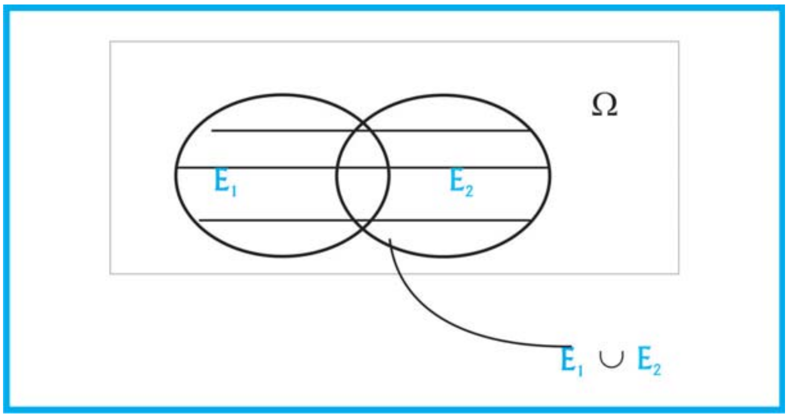

Figura 41: Evento união
Fonte: elaborada pelo autor

$E_1 \cup E_2 = \{2, 3, 4, 5, 6\}$

Composto por todos os resultados que pertencem a um **ou** ao outro, **ou** a ambos.

Evento **intersecção** de E_1 com E_2 ($E_1 \cap E_2$): evento que ocorre se E_1 **E** E_2 ocorrem simultaneamente.

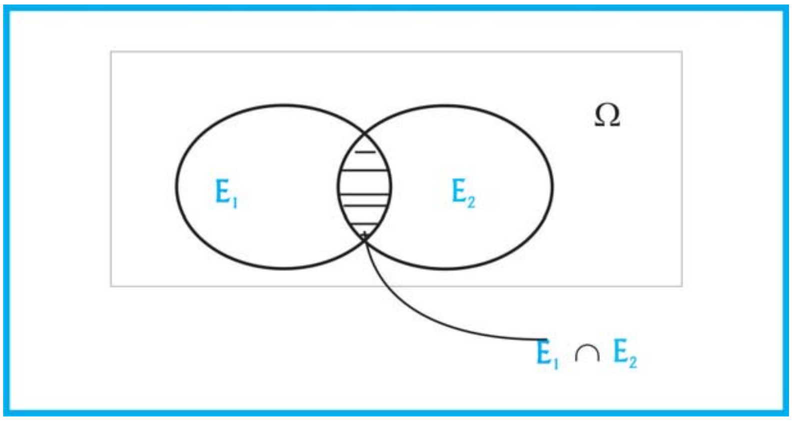

Figura 42: Evento intersecção
Fonte: elaborada pelo autor

Composto por todos os resultados que pertencem a ambos: $E_1 \cap E_2 = \{4, 6\}$

Eventos **mutuamente exclusivos** (M.E.): são eventos que não podem ocorrer simultaneamente, não apresentando elementos em comum (sua intersecção é o conjunto vazio).

Entre os três eventos definidos acima, observamos que os eventos E_1 e E_3 não têm elementos em comum:

> $E_3 = \{1, 3\}$ $E_1 = \{2, 4, 6\}$ $E_1 \cap E_3 = \varnothing \Rightarrow E_1$ e E_3 são mutuamente exclusivos.

Evento **complementar** de um evento qualquer é formado por todos os resultados do espaço amostral que não pertencem ao evento. A união de um evento e seu complementar formará o próprio espaço amostral, e a intersecção de um evento e seu complementar são o conjunto vazio.

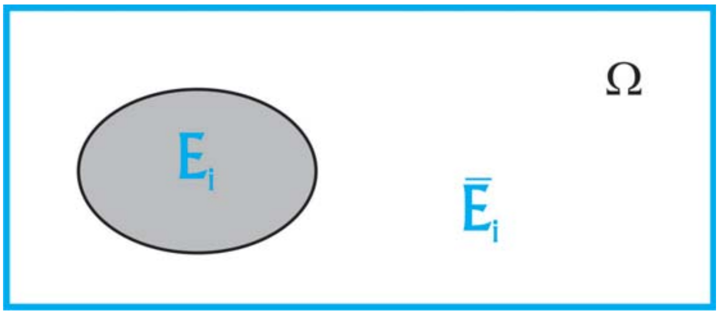

Figura 43: Evento complementar
Fonte: elaborada pelo autor

$E_i \cup \overline{E}_i = \Omega$ $\quad E_i \cap \overline{E}_i = \varnothing$

$E_1 = \{2, 4, 6\}$ $\quad \overline{E}_1 = \{1, 3, 5\}$

$E_2 = \{3, 4, 5, 6\}$ $\quad \overline{E}_2 = \{1, 2\}$

A compreensão das definições anteriores será extremamente útil quando calcularmos probabilidades, pois as expressões poderão ser deduzidas ou simplificadas se identificarmos que se trata de evento união, intersecção, ou se os eventos de interesse são mutuamente exclusivos ou complementares. Conhecido isso, podemos agora passar à definição de probabilidade ou, mais especificamente, às definições de probabilidade, que são complementares.

Definições de probabilidade

Por que usamos plural, “definições”, ao invés de “definição”? Porque ao longo dos séculos várias definições de probabilidade foram apresentadas, e elas se complementam.

A repetição de um experimento aleatório, mesmo sob condições semelhantes, poderá levar a resultados (eventos) diferentes. Mas se o experimento for repetido um número “suficientemente grande” de vezes, haverá uma regularidade nestes resultados que permitirá calcular a sua probabilidade de ocorrência. Essa é a base para as definições que veremos a seguir.

Definição clássica de probabilidade

Intuitivamente, as pessoas sabem como calcular algumas probabilidades para tomar decisões. Observe os seguintes exemplos.

Exemplo 1: vamos supor que você fez uma aposta com um amigo. O vencedor será aquele que acertar a face que ficar para cima após o lançamento de uma *moeda honesta*. Qual é a chance de você ganhar?

Usaremos o termo "moeda honesta" para referenciar uma moeda perfeitamente equilibrada e lançamentos imparciais. De forma análoga, usaremos o adjetivo honesto para dado, baralho, entre outros.

Intuitivamente, você responderia que há 50% (1/2) de chances de ganhar, uma vez que há apenas duas faces (resultados) possíveis. Mesmo sem saber o que é probabilidade, você pode calcular a chance de ocorrência de um evento de interesse, a face na qual você apostou.

Você continua apostando com o mesmo amigo. O vencedor agora será aquele que acertar o naipe de uma carta que será retirada ao acaso de um baralho comum de 52 cartas. Veremos neste segundo exemplo: qual é a chance de você ganhar?

Novamente, de forma intuitiva, você responderia que há 25% (1/4) de chance, uma vez que há apenas quatro naipes (resultados) possíveis.

O que há em comum entre as situações dos exemplos 1 e 2? Refletindo um pouco, você observará que em ambos temos experimentos aleatórios. Em cada realização do experimento, apenas um dos resultados possíveis pode ocorrer. Além disso, como se supõe que a moeda e o baralho são honestos, cada um dos resultados possíveis tem a mesma probabilidade de ocorrer: tanto cara quanto coroa tem 50% de chance de ocorrer, todos os quatro naipes (copas, espadas, ouros e paus) têm 25% de chance de ocorrer. Sem que você soubesse, você aplicou a **definição clássica de probabilidade** para obter as chances de ganhar.

Se um experimento aleatório puder resultar em $\mathbf{n}$ *diferentes e igualmente prováveis* resultados, e $\mathbf{n_{Ei}}$ destes resultados referem-se ao evento $\mathbf{E_i}$, então a probabilidade de o evento $\mathbf{E_i}$ ocorrer será:

$$P(Ei) = \frac{n_{Ei}}{n}$$

O problema reside em calcular o número total de resultados possíveis e o número de resultados associados ao evento de interesse. Isso pode ser feito usando técnicas de análise combinatória (que serão vistas posteriormente) ou por considerações teóricas (“bom senso”).

Seja o seguinte experimento aleatório: lançamento de um dado não viciado e observação da face voltada para cima. Neste Exemplo 3, vamos calcular as probabilidades de ocorrência dos seguintes eventos:

a) face 1;

b) face par; e

c) face menor ou igual a 2.

O espaço amostral deste experimento será: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Sendo assim, há um total de seis resultados possíveis, resultando em **n** = 6. Basta, então, definir quantos resultados estão associados a cada evento para que seja possível calcular suas probabilidades pela definição clássica.

O evento “face 1” tem apenas um resultado associado: { 1 }. Então, $\mathbf{n_{Ei}} = 1$, e a probabilidade de ocorrer a face 1 será:

$$P(Ei) = \frac{n_{Ei}}{n} = \frac{1}{6}$$

O evento “face par” tem três resultados associados: {2, 4, 6}. Então, $\mathbf{n_{Ei}} = 3$, e a probabilidade de ocorrer face par será:

$$P(Ei) = \frac{n_{Ei}}{n} = \frac{3}{6} = \frac{1}{2}$$

O evento “face menor ou igual a 2” tem dois resultados associados: {1, 2}. Então, $\mathbf{n_{Ei}} = 2$, e a probabilidade de ocorrência de face menor ou igual a 2 será:

$$P(Ei) = \frac{n_{Ei}}{n} = \frac{2}{6} = \frac{1}{3}$$

Como viu nos exemplos, a definição clássica, que foi desenvolvida a partir do século XVII, foi inicialmente aplicada para orientar apostas em jogos de azar. Surgiram dois problemas desta aplicação.

O primeiro é relativamente óbvio: muitos jogos de azar não eram “honestos”, os donos das casas inescrupulosamente “viciavam” dados e roletas, marcavam baralhos, de maneira a fazer com que os clientes perdessem sistematicamente, ou seja, o lançamento dos dados e a retirada da carta do baralho não eram mais experimentos aleatórios.

O segundo problema decorre da pergunta: será que em todos os experimentos aleatórios todos os eventos terão a mesma probabilidade de ocorrer? Será que a probabilidade de chover no mês de novembro na cidade de Brest (na França, que tem, em média, 225 dias nublados por ano) é a mesma na cidade de Sevilha (na Espanha, que tem, em média, 240 dias de sol por ano)? Precisamos partir para a **definição experimental de probabilidade**.

Definição experimental de probabilidade

Seja um experimento aleatório que é repetido **n** vezes, e $\mathbf{E_i}$, um evento associado.

A freqüência relativa do evento $\mathbf{E_i}$: $f_{REi} = \frac{n_{Ei}}{n} = \frac{\text{nº vezes que } E_i \text{ ocorreu}}{\text{total de tentativas}}$

Quando o número de repetições tende ao infinito (ou a um número suficientemente grande), $\mathbf{f_{REi}}$ tende a um limite: a probabilidade de ocorrência do evento $\mathbf{E_i}$. A probabilidade do evento pode ser estimada através da freqüência relativa. Lembre-se da Unidade 3, a descrição de um fenômeno pode ser feita por distribuição de freqüências.

Quando não há outra maneira de obter as probabilidades dos eventos, é necessário realizar o experimento (veja novamente a Unidade 1) várias vezes para que seja possível obter um número tal de tentativas que permita que as freqüências relativas estimem as probabilidades, para que se passa construir um modelo probabilístico para o experimento. Isso pode ser feito em laboratório, em condições controladas, por exemplo, a vida útil das lâmpadas vendidas no comércio é definida através de testes de sobrevivência realizados pelos fabricantes.

Mas, em alguns casos, não é possível realizar experimentos, a maioria dos fenômenos socioeconômicos e climáticos, por exemplo. Neste caso, precisamos estimar as probabilidades através das freqüências relativas históricas.

Independente de como obtemos as probabilidades, elas obedecem a alguns axiomas e propriedades que veremos a seguir.

Axiomas e propriedades da probabilidade

Alguns autores chamam estes axiomas e propriedades de definição axiomática da probabilidade.

Sejam um experimento aleatório e um espaço amostral associado a ele. A cada evento $\mathbf{E_i}$, associaremos um número real denominado $\mathbf{P(E_i)}$ que deve satisfazer os seguintes axiomas:

a) $0 \leq \mathbf{P(E_i)} \leq 1{,}0$

A probabilidade de ocorrência de um evento *sempre* é um número real entre 0 e 1 (0% e 100%)

b) $\mathbf{P\ (\Omega)} = 1{,}0$

A probabilidade de ocorrência do espaço amostral é igual a 1 (100%), pois pelo menos um dos resultados do espaço amostral ocorrerá. Por isso, o espaço amostral é chamado de **evento certo**.

c) Se E_1, E_2, ..., En são eventos mutuamente exclusivos, então $P(E_1 \cup E_2 \cup ... \cup E_n) = P(E_1) + P(E_2) + ... + P(E_n)$

Este axioma afirma que, ao unir resultados diferentes, devemos somar as probabilidades.

GLOSSÁRIO

*Evento impossível – evento com probabilidade de ocorrer igual a 0%, é o conjunto vazio. Fonte: Barbetta, Reis e Bornia (2004).

Além dos axiomas, há algumas propriedades básicas da probabilidade:

a) $P\ (\varnothing) = 0$

A probabilidade de ocorrência do conjunto vazio é **nula** (igual a zero), uma vez que não há resultados no conjunto vazio. Por isso, o conjunto vazio é chamado de **evento impossível***.

b) $\Sigma\ \mathbf{P(E_i)} = 1{,}0$

Se a probabilidade de ocorrência do espaço amostral é igual a 1 (100%), ao somar as probabilidades de todos os eventos que compõem o espaço amostral, o resultado deverá ser igual a 1 (100%).

c) $\mathbf{P(E_i)} = 1 - \mathbf{P(\overline{E_i})}$

A probabilidade de ocorrência de um evento qualquer será igual à probabilidade do espaço amostral (1 ou 100%) menos a probabilidade de seu evento complementar (a soma das probabilidades de todos os outros eventos do espaço amostral).

d) Sejam $\mathbf{E_i}$ e $\mathbf{E_j}$ dois eventos quaisquer: $P(\mathbf{E_i} \cup \mathbf{E_j}) = P(\mathbf{E_i}) + P(\mathbf{E_j}) - P(\mathbf{E_i} \cap \mathbf{E_j})$

A probabilidade de ocorrência do evento União de dois outros eventos será igual à soma das probabilidades de cada evento menos a probabilidade de ocorrência do evento intersecção dos mesmos dois eventos. Esta propriedade também é chamada de **regra da adição**.

Veja, neste quarto exemplo, que seja o experimento aleatório lançamento de um dado não viciado e observação da face voltada para cima definido no Exemplo 3: o seu espaço amostral será $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Definindo três eventos: $\mathbf{E_1}$ = face 1 = {1}, $\mathbf{E_2}$ = face par = {2, 4, 6} e $\mathbf{E_3}$ = face ≤ 2 {1, 2}, cujas probabilidades já foram calculadas.

Calcular a probabilidade de ocorrência dos seguintes eventos:

a) complementar de $\mathbf{E_1}$;

b) complementar de $\mathbf{E_2}$;

c) união de $\mathbf{E_2}$ e $\mathbf{E_3}$; e

d) união de $\mathbf{E_1}$ e $\mathbf{E_2}$.

No Exemplo 2, obtiveram-se $P(\mathbf{E_1}) = 1/6$, $P(\mathbf{E_2}) = 3/6$ e $P(\mathbf{E_3}) = 2/6$.
Usando as propriedades:

$\mathbf{P(E_1)} = 1 - \mathbf{P(\overline{E}_1)}$ então $\mathbf{P(\overline{E}_1)} = 1 - \mathbf{P(E_1)} = 1 - 1/6 = 5/6$
$\mathbf{\overline{E}_1} = \{2, 3, 4, 5, 6\}$

$\mathbf{P(E_2)} = 1 - \mathbf{P(\overline{E}_2)}$ então $\mathbf{P(\overline{E}_2)} = 1 - \mathbf{P(E_2)} = 1 - 3/6 = 3/6$
$\mathbf{\overline{E}_2} = \{1, 3, 5\}$

$\mathbf{P(E_2 \cup E_3)} = \mathbf{P(E_2)} + \mathbf{P(E_3)} - \mathbf{P(E_2 \cap E_3)}$ Observe que há apenas um elemento em comum entre os eventos $\mathbf{E_2}$ e $\mathbf{E_3}$: apenas um resultado associado => $\mathbf{P(E_2 \cap E_3)} = 1/6$

$\mathbf{P(E_2 \cup E_3)} = 3/6 + 2/6 - 1/6 = 4/6$

$\mathbf{P(E_1 \cup E_2)} = \mathbf{P(E_1)} + \mathbf{P(E_2)} - \mathbf{P(E_1 \cap E_2)}$ Não há elementos em comum entre os eventos $\mathbf{E_1}$ e $\mathbf{E_2}$: eles são mutuamente exclusivos, sua intersecção é o conjunto vazio, e a probabilidade de ocorrência do conjunto vazio é nula. **P($E_1 \cup E_2$)** = 1/6 + 3/6 – 0 = 4/6

Agora, vamos exercitar a mente! Imagine que você trabalha em uma corretora de ações e precisa aconselhar um cliente sobre investir ou não em ações da Petrobrás. Supõe-se que o preço do barril do petróleo subirá cerca de 10% nos próximos dias, há uma probabilidade estimada de isso acontecer. E, sabendo disso, você gostaria de saber qual é a probabilidade de que as ações da empresa subam também 10% na Bovespa. Este caso, em que queremos calcular a probabilidade de ocorrência de um evento condicionada à ocorrência de outro, somente poderá ser resolvido por **probabilidade condicional**, que veremos a seguir.

Probabilidade condicional

Muitas vezes, há interesse de calcular a probabilidade de ocorrência de um evento A qualquer, dada a ocorrência de um outro evento B. Por exemplo, qual é a probabilidade de chover amanhã em Florianópolis, sabendo-se que hoje choveu? Ou qual é a probabilidade de um dispositivo eletrônico funcionar sem problemas por 200 horas consecutivas, sabendo-se que ele já funcionou por 100 horas? Ou ainda, a situação levantada anteriormente: qual é a probabilidade de que as ações da Petrobrás aumentem 10%, se o preço do barril de petróleo subir 10% previamente?

Veja, queremos calcular a probabilidade de ocorrência de A **condicionada** à ocorrência prévia de B, simbolizada por P(A | B) – lê-se probabilidade de A dado B –, e a sua expressão será:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{para } P(B) > 0$$

A probabilidade de ocorrência de A condicionada à ocorrência de B será igual à probabilidade da intersecção entre A e B, dividida pela probabilidade de ocorrência de B (o evento que já ocorreu).

No denominador da expressão, é colocada **sempre** a probabilidade do evento que já ocorreu.

Se houvesse interesse no oposto, probabilidade de ocorrência de B condicionada à ocorrência prévia de A:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} \quad \text{para } P(A) > 0$$

Neste caso, o valor no denominador seria a probabilidade de A, uma vez que este evento ocorreu previamente, tal como B na outra expressão. É importante ressaltar que a operação de intersecção é **comutativa***, implicando:

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

GLOSSÁRIO

*** Operação comutativa** – operação em que a seqüência de realização não modifica o resultado, "a ordem dos fatores não altera o produto". Fonte: elaborado pelo autor.

Sejam o lançamento de dois dados não viciados, um após o outro, e a observação das faces voltadas para cima. Neste quinto exemplo, vamos calcular as probabilidades:

- de que as faces sejam iguais, supondo-se que sua soma é menor ou igual a 5; e
- de que a soma das faces seja menor ou igual a 5, supondo-se que as faces são iguais.

Observe que há interesse em calcular a probabilidade de eventos, supondo que outro evento ocorreu previamente.

Como todo problema de probabilidade, é preciso montar o espaço amostral. Neste caso, serão os pares de faces dos dados, e como os dados são lançados um após o outro, a ordem das faces é importante:

$$\Omega = \left\{ \begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix} \right\}$$

Figura 44: Espaço amostral do Exemplo 5
Fonte: elaborada pelo autor

Há um total de 36 resultados possíveis: **n** = 36. Agora, é preciso definir os eventos de interesse.

- "Faces iguais, sabendo-se que sua soma é menor ou igual a 5" significa dizer probabilidade de ocorrência de faces iguais supondo-se que **já ocorreram** faces cuja soma é menor ou igual a 5; chamando o evento faces iguais de $\mathbf{E_1}$ e o evento soma das faces menor ou igual a 5 de $\mathbf{E_2}$, estamos procurando P($\mathbf{E_1}$ | $\mathbf{E_2}$), probabilidade de ocorrência de $\mathbf{E_1}$ condicionada à ocorrência PRÉVIA de $\mathbf{E_2}$.

Usando a fórmula:

$P(E_1 | E_2) = \dfrac{P(E_1 \cap E_2)}{P(E_2)}$, é preciso encontrar os valores das probabilidades.

Primeiramente, definir o número de resultados do espaço amostral que pertencem aos eventos de interesse, para que seja possível calcular a sua probabilidade usando a definição clássica de probabilidade:

$\mathbf{E_1}$ = {(1,1) (2,2) (3,3) (4,4) (5,5) (6,6)} – faces iguais, 6 resultados, $\mathbf{nE_1}$ = 6.

$\mathbf{E_2}$ = {(1,1) (1,2) (1,3) (1,4) (2,1) (2,2) (2,3) (3,1) (3,2) (4,1)} – soma das faces ≤5, 10 resultados, $\mathbf{nE_2}$ = 10.

Os elementos em comum formarão o evento intersecção: $\mathbf{E_1} \cap \mathbf{E_2}$ = {(1,1) (2,2)} – faces iguais e soma das faces ≤5, 2 resultados, $\mathbf{nE_1 \cap E_2}$ = 2.

$\mathbf{P(E_2) = nE_2 / n = 10/36}$ $\mathbf{P(E_1 \cap E_2) = nE_1 \cap E_2 / n = 2/36}$

Tendo as probabilidades acima, é possível calcular a probabilidade condicional:

$$P(E_1 | E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)} = \frac{2/36}{10/36} = \frac{2}{10} = 0,2(20\%)$$

Então, a probabilidade de que as faces sejam iguais, sabendo-se que sua soma é menor ou igual a 5, é de 20%.

Este resultado poderia ser obtido de outra forma. Se a soma das faces é menor ou igual a 5, o evento $\mathbf{E_2}$ já ocorreu previamente, então o espaço amostral **modificou-se**, passando a ser o conjunto de resultados do evento $\mathbf{E_2}$:

novo Ω = {(1,1) (1,2) (1,3) (1,4) (2,1) (2,2) (2,3) (3,1) (3,2) (4,1)}

O novo espaço amostral tem dez resultados, novo **n** = 10.

O número de resultados do evento "faces iguais ($\mathbf{E_1}$) no novo espaço amostral" é igual a 2, novo n$\mathbf{E_1}$ = 2 (há apenas dois pares no novo espaço amostral, de soma das faces menor ou igual a 5, em que as faces são iguais).

Então, a probabilidade de ocorrer o evento $\mathbf{E_1}$ no novo espaço amostral, ou seja, a probabilidade de ocorrência do evento $\mathbf{E_1}$ **condicionada** à ocorrência prévia do evento $\mathbf{E_2}$, P($\mathbf{E_1}$ | $\mathbf{E_2}$), será:

> P($\mathbf{E_1}$ | $\mathbf{E_2}$) = novo **n**$\mathbf{E_1}$/ novo **n** = 2/10 = 0,2 (20%), o mesmo resultado obtido anteriormente.

b) "Soma das faces menor ou igual a 5, sabendo-se que as faces são iguais" significa dizer probabilidade de ocorrência de faces cuja soma é menor ou igual a 5, supondo-se que já ocorreram faces que são iguais; chamando o evento faces iguais de $\mathbf{E_1}$ e o evento soma das faces menor ou igual a 5 de $\mathbf{E_2}$, estamos procurando P($\mathbf{E_2}$ | $\mathbf{E_1}$), probabilidade de ocorrência de $\mathbf{E_2}$ condicionada à ocorrência PRÉVIA de $\mathbf{E_1}$.

Houve uma mudança no evento que ocorreu previamente.

Usando a fórmula: $P(E_2|\, E_1) = \dfrac{P(E_2 \cap E_1)}{P(E_1)}$, todos os valores já foram obtidos no item a.

$$P(E_2|\, E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{2/36}{6/36} = \frac{2}{6} = 0{,}33(33\%)$$

Então, a probabilidade de que as faces tenham soma menor ou igual a 5, sabendo-se que são iguais, é de 33%.

Da mesma forma que no item a, o resultado poderia ser obtido de outra forma. Se as faces são iguais, o evento $\mathbf{E_1}$ já ocorreu previamente, então o espaço amostral **modificou-se**, passando a ser o conjunto de resultados do evento $\mathbf{E_1}$:

novo Ω = { (1,1) (2,2) (3,3) (4,4) (5,5) (6,6)}

O novo espaço amostral tem 6 resultados, novo **n** = 6.

O número de resultados do evento "soma das faces menor ou igual a 5 ($\mathbf{E_2}$)" no novo espaço amostral é igual a 2, novo n$\mathbf{E_2}$ = 2 (há apenas dois pares no novo espaço amostral, de faces iguais, em que a soma das faces é menor ou igual a 5).

Então, a probabilidade de ocorrer o evento $\mathbf{E_2}$ no novo espaço amostral, ou seja, a probabilidade de ocorrência do evento $\mathbf{E_2}$ **condicionada** à ocorrência prévia do evento $\mathbf{E_1}$, $P(\mathbf{E_2} \mid \mathbf{E_1})$, será:

$P(\mathbf{E_2} \mid \mathbf{E_1})$ = novo $\mathbf{nE_2}$/ novo $\mathbf{n}$ = 2/6 = 0,33 (33%), o mesmo resultado obtido anteriormente.

É extremamente importante lembrar que, conceitualmente, $P(A|B) \neq P(B|A)$, pois os eventos que ocorreram previamente são diferentes.

No quinto exemplo, utilizamos a definição clássica para obter as probabilidades necessárias, mas poderíamos usar distribuições de freqüências de dados históricos ou experimentais para obtê-las.

Regra do produto

Uma das conseqüências da expressão da probabilidade condicional é a regra do produto, isolando a probabilidade da intersecção:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad => \quad P(A \cap B) = P(B) \times P(A \mid B)$$

Neste caso, o evento B ocorreu previamente, e o segundo valor é a probabilidade de ocorrência de A dado que B ocorreu.

$$P(A \mid B) = \frac{P(A \cap B)}{P(A)} \quad => \quad P(A \cap B) = P(A) \times P(B \mid A)$$

Neste caso, o evento A ocorreu previamente, e o segundo valor é a probabilidade de ocorrência de B dado que A ocorreu.

Não se esqueça que a intersecção é comutativa.

É importante que seja observada com cuidado a seqüência dos eventos para montar as expressões acima: analisar corretamente que evento já ocorreu.

No Exemplo 6, digamos que uma urna contém duas bolas brancas e três vermelhas. Retiram-se duas bolas ao acaso, uma após a outra. Veremos nos itens abaixo se a retirada foi feita **sem reposição**.

a) Qual é a probabilidade de que as duas bolas retiradas sejam da mesma cor?

b) Qual é a probabilidade de que as duas bolas retiradas sejam vermelhas, supondo-se que são da mesma cor?

Como em todos os problemas de probabilidade, primeiramente é preciso definir o espaço amostral. Há duas cores e duas retiradas, então podemos ter:

- a 1ª e a 2ª bolas brancas (duas bolas da mesma cor) – evento $E_1 = B_1 \cap B_2$;
- a 1ª bola branca e a 2ª bola vermelha – evento $E_2 = B_1 \cap V_2$;
- a 1ª bola vermelha e a 2ª bola branca – evento $E_3 = V_1 \cap B_2$;
- a 1ª bola vermelha e a 2ª bola vermelha (duas bolas da mesma cor) – evento $E_4 = V_1 \cap V_2$.

Então, o espaço amostral será:

$$\Omega = \{B_1 \cap B_2, B_1 \cap V_2, V_1 \cap B_2, V_1 \cap V_2\}$$

Todos os quatro eventos acima são mutuamente exclusivos: quando as bolas forem retiradas, apenas um, e somente um, dos eventos acima pode ocorrer.

As retiradas são feitas sem reposição: a segunda retirada depende do resultado da primeira. Se as retiradas forem feitas sem reposição, elas serão dependentes, pois o espaço amostral será modificado: em cada retirada, as probabilidades de ocorrência são modificadas, porque as bolas não são repostas.

- a probabilidade de retirar bola branca na 1ª retirada é de 2/5 (duas bolas brancas no total de cinco), $P(B_1) = 2/5$; e
- a probabilidade de retirar bola vermelha na 1ª retirada é de 3/5 (três bolas vermelhas em cinco), $P(V_1) = 3/5$.

Se a primeira bola retirada foi branca (o evento B_1 ocorreu previamente), restaram quatro bolas, uma branca e três vermelhas:

Repare que o número de bolas, número de resultados, diminuiu de cinco para quatro, porque as retiradas são feitas sem reposição.

- a probabilidade de retirar uma bola branca na 2ª retirada se na 1ª foi extraída uma branca, é de 1/4 (uma bola branca em quatro), $P(B_2 | B_1) = ¼$; e
- a probabilidade de retirar uma bola vermelha na 2ª retirada, se na 1ª foi extraída uma branca é de 3/4 (três bolas vermelhas em quatro), $P(V_2 | B_1) = 3/4$.

Se a primeira bola retirada foi vermelha (o evento V_1 ocorreu previamente), restaram quatro bolas, duas brancas e duas vermelhas:

- a probabilidade de retirar uma bola branca na 2ª retirada, se na 1ª foi extraída uma vermelha, é de 2/4 (duas bolas brancas em quatro), $P(B_2 | V_1) = 2/4$; e
- a probabilidade de retirar uma bola vermelha na 2ª retirada, se na 1ª foi extraída uma vermelha, é de 2/4 (duas bolas vermelhas em quatro), $P(V_2 | V_1) = 2/4$.

a) O evento que nos interessa: "bolas da mesma cor": brancas ou vermelhas, evento união brancas-vermelhas.

Chamando bolas da mesma cor de evento F:

$$F = [(B_1 \cap B_2) \cup (V_1 \cap V_2)]$$

Usando as propriedades da probabilidade:

$$P(F) = P\,[(B_1 \cap B_2) \cup (V_1 \cap V_2)] =$$
$$P(B_1 \cap B_2) + P(V_1 \cap V_2) - P\,(B_1 \cap B_2) \cup (V_1 \cap V_2)$$

Os eventos $(B_1 \cap B_2)$ e $(V_1 \cap V_2)$ são mutuamente exclusivos, se as bolas são da mesma cor, ou são brancas ou são vermelhas; então, a intersecção entre eles é o conjunto vazio, e a probabilidade de o conjunto vazio ocorrer é igual a zero (ver seção 5.3.3); então, simplesmente:

$$P(F) = P\,[(B_1 \cap B_2) \cup (V_1 \cap V_2)] = P(B_1 \cap B_2) + P(V_1 \cap V_2)$$

Usando a regra do produto:

$P(B_1 \cap B_2) = P(B_1) \text{ x } P(B_2 | B_1) = (2/5) \text{ x } (1/4) = 2/20 = 1/10$
$P(V_1 \cap V_2) = P(V_1) \text{ x } P(V_2 | V_1) = (3/5) \text{ x } (2/4) = 6/20 = 3/10$

Substituindo na expressão:

$P(F) = P\,[(B_1 \cap B_2) \cup (V_1 \cap V_2)] = P(B_1 \cap B_2) + P(V_1 \cap V_2) =$
$1/10 + 3/10 = 4/10 = 0{,}4\ (40\%)$

Então, se as retiradas forem feitas sem reposição, a probabilidade de que as duas bolas sejam da mesma cor será igual a 0,4 (40%).

b) Neste caso, sabe-se que as duas bolas são da mesma cor (o evento F acima JÁ OCORREU), e há interesse em saber a probabilidade de que as duas bolas sejam vermelhas:

$$\mathbf{P[(V_1 \cap V_2) \mid F] = P\{(V_1 \cap V_2) \mid [(B_1 \cap B_2) \cup (V_1 \cap V_2)]\}}$$

Usando a expressão de probabilidade condicional:

$$\mathbf{P\{(V_1 \cap V_2) \mid [(B_1 \cap B_2) \cup (V_1 \cap V_2)]\}} = \frac{P\{(V_1 \cap V_2) \cap [(B_1 \cap B_2) \cup (V_1 \cap V_2)]\}}{P[(B_1 \cap B_2) \cup (V_1 \cap V_2)]}$$

A probabilidade do denominador já é conhecida do item a. E a do numerador pode ser obtida facilmente.

Repare: o que há em comum entre o evento $(V_1 \cap V_2)$ e o evento $\mathbf{[(B_1 \cap B_2) \cup (V_1 \cap V_2)]}$, em suma, qual será o evento intersecção? O que há em comum entre duas bolas vermelhas e duas bolas da mesma cor? **O próprio evento duas bolas vermelhas** $(V_1 \cap V_2)$, então:

$(V_1 \cap V_2) \cap \mathbf{[(B_1 \cap B_2) \cup (V_1 \cap V_2)]} = (V_1 \cap V_2);$
$P\{(V_1 \cap V_2) \cap \mathbf{[(B_1 \cap B_2) \cup (V_1 \cap V_2)]}\} = P\,(V_1 \cap V_2) = 3/10.$

Sabendo que $P\{(V_1 \cap V_2) \mid \mathbf{[(B_1 \cap B_2) \cup (V_1 \cap V_2)]}\} = 4/10$ (do item a.1) e substituindo os valores na fórmula:

$$P\{(V_1 \cap V_2) \mid [\mathbf{(B_1 \cap B_2) \cup (V_1 \cap V_2)}]\} = \frac{P(V_1 \cap V_2)}{P[(B_1 \cap B_2) \cup (V_1 \cap V_2)]} = \frac{3/10}{4/10} = \frac{3}{4}$$

$$P\{(V_1 \cap V_2) \mid [\mathbf{(B_1 \cap B_2) \cup (V_1 \cap V_2)}]\} = 0{,}75\ (75\%)$$

Então, se as retiradas forem feitas sem reposição, e as duas bolas forem da mesma cor, a probabilidade de que sejam vermelhas será igual a 0,75 (75%).

As retiradas e as probabilidades podem ser representadas através de um diagrama chamado de "Árvore de probabilidades":

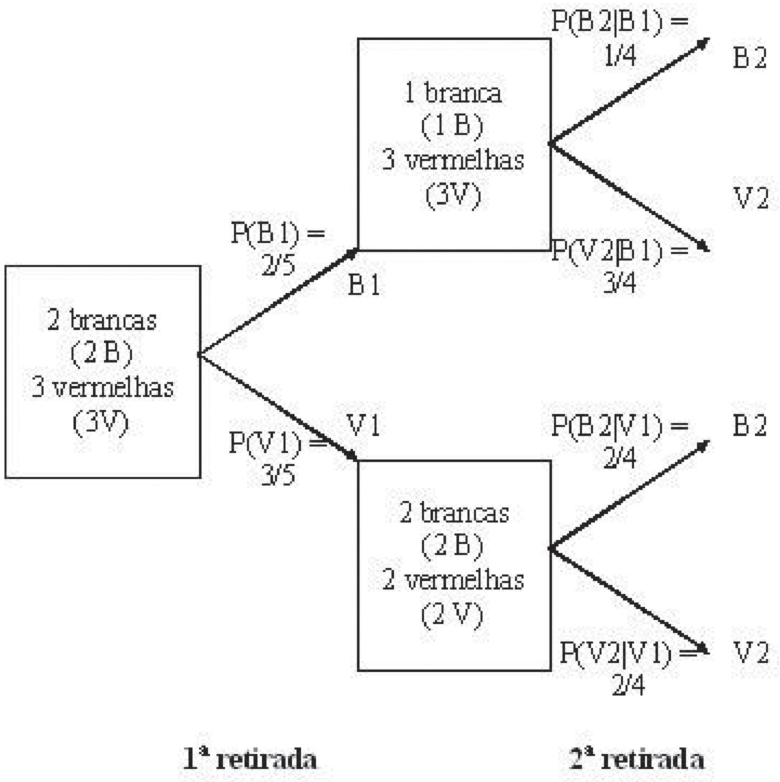

Figura 44: Árvore de probabilidades – Retiradas sem reposição
Fonte: elaborada pelo autor

Observe que, através da árvore de probabilidades, podemos chegar aos mesmos resultados obtidos anteriormente. Partindo do espaço amostral original, um dos ramos significa 1ª bola branca (B_1), e o outro, 1ª bola vermelha (V_1). Dependendo do resultado da primeira retirada, haverá um espaço amostral diferente: uma bola branca e três vermelhas, se na 1ª retirada obteve-se uma bola branca, ou duas bolas brancas e duas vermelhas, se na 1ª retirada obteve-se uma bola vermelha.

A partir dos novos espaços amostrais, é possível calcular as probabilidades condicionais para cada caso e depois substituí-las nas fórmulas adequadas. Contudo, a árvore será inútil se o evento para o qual se deseja calcular a probabilidade não for definido adequadamente: neste caso, no item a, bolas da mesma cor $[(B_1 \cap B_2) \cup (V_1 \cap V_2)]$, e no item b, bolas vermelhas sabendo que são da mesma cor $\{(V_1 \cap V_2) \mid [(B_1 \cap B_2) \cup (V_1 \cap V_2)]\}$.

A árvore será igualmente inútil se não forem usadas as definições de eventos dependentes (porque não há reposição) e de eventos mutuamente exclusivos (porque os eventos não podem ocorrer simultaneamente), e as expressões de probabilidade condicional e os axiomas de probabilidade.

O grande inconveniente da árvore de probabilidades surge quando o número de "retiradas" aumenta e/ou o número de resultados possíveis para cada retirada é considerável: torna-se impraticável desenhar a árvore, enumerando todos os resultados. Nestes casos, usa-se análise combinatória, que veremos adiante.

E se a ocorrência do evento A não modificasse a probabilidade de ocorrência de B? Os eventos A e B seriam chamados de independentes. Você pode imaginar situações práticas em que dois eventos sejam independentes?

Eventos independentes

Dois ou mais eventos são independentes quando a ocorrência de um dos eventos não influencia a probabilidade de ocorrência dos outros. Se dois eventos A e B são independentes, então a probabilidade de A ocorrer dado que B ocorreu é igual à própria probabilidade de ocorrência de A, e a probabilidade de B ocorrer dado que B ocorreu é igual à própria probabilidade de ocorrência de B.

Se A e B são independentes, então:

$$P(A \mid B) = P(A) \text{ e } P(B \mid A) = P(B)$$
$$P(A \cap B) = P(A) \times (B \mid A) = P(A) \times P(B)$$
$$P(A \cap B) = P(B) \times (A \mid B) = P(B) \times P(A)$$

*As expressões acima são válidas, se os eventos **A** e **B** forem independentes.*

Em situações práticas, dois eventos são independentes quando a ocorrência de um deles não modifica ou modifica muito pouco o espaço amostral do experimento aleatório. É o que ocorria na Unidade 2, quando fazíamos amostragem aleatória simples: naquele momento, não foi dito que a amostragem era com reposição, que dificilmente é feita na prática, mas admite-se que, sendo o tamanho da população muito grande, a retirada de uma pequena amostra não modificará muito as proporções dos eventos.

Exemplo 7: para a mesma situação do Exemplo 6. Uma urna contém duas bolas brancas e três vermelhas. Retiram-se duas bolas ao acaso, uma após a outra. Resolva os itens abaixo, se a retirada for feita **com reposição**.

a) Qual é a probabilidade de que as duas bolas retiradas sejam da mesma cor? R.: 0,52(52%).

b) Qual é a probabilidade de que as duas bolas retiradas sejam vermelhas, supondo-se que são da mesma cor? R.: 0,69 (69%).

A seguir, veremos como resolver um problema quando for inviável usar a árvore de probabilidades.

Probabilidade combinatória

Como já vimos, em muitos casos a resolução dos problemas de probabilidade enumerando todos os resultados possíveis torna-se extremamente difícil. Há uma forma mais rápida de enumerar os resultados: as técnicas de análise combinatória.

Relembremos a definição clássica de probabilidade, que consistia em calcular o quociente entre o número de resultados associados ao evento e o número total de resultados possíveis. O cálculo desses números de resultados pode ser feito utilizando análise combinatória, tanto para os casos em que os eventos são dependentes quanto quando há independência.

As técnicas de análise combinatória buscam basicamente calcular o número de maneiras de dispor um certo número de "objetos" em um número limitado de "espaços" distintos (menor do que o número de objetos), sendo um objeto em cada espaço. Se o número de "objetos" é, teoricamente, infinito (ou ilimitado), temos a análise combinatória com repetição ilimitada (situação de independência): é o que ocorre nos casos em que há reposição. Se, porém, o número de "objetos" é limitado, temos a análise combinatória sem repetição (situação de dependência): casos em que não há reposição.

Análise combinatória com reposição ilimitada

Há **n** objetos disponíveis em número ilimitado; em outras palavras, há reposição, de quantas maneiras diferentes é possível preencher **k** espaços distintos com os objetos, cada espaço com um objeto?

Tem-se um espaço e **n** objetos, há **n** maneiras de dispô-los no espaço. Tem-se dois espaços, e os mesmos **n** objetos disponíveis para cada um, haverá $\mathbf{n^2}$ maneiras: as **n** maneiras do primeiro espaço multiplicadas pelas **n** maneiras do segundo. Se houver três espaços, haverá $\mathbf{n^3}$ maneiras, e assim por diante.

Generalizando, se há **n** objetos disponíveis em número **ilimitado** para preencher **k** espaços distintos, cada espaço com um objeto, há $\mathbf{n^k}$ maneiras de fazê-lo, e cada preenchimento é independente dos outros.

Veremos neste oitavo exemplo quantas palavras de cinco letras podem ser escritas com as 26 letras do alfabeto, sem se preocupar com o significado.

Veja que, primeiramente, é preciso identificar os objetos e os espaços.

Os objetos, neste caso, são as letras do alfabeto, 26, então **n** = 26. Como não há preocupação com o significado das palavras, os objetos estão disponíveis em número ilimitado.

Os espaços são as letras da palavra: cada palavra deve ter cinco letras, então **k = 5**.

Usando a expressão de análise combinatória com repetição ilimitada, o número de palavras será: $\mathbf{n}^{\mathbf{k}} = 26^5 = 11.881.376$ palavras.

Uma urna contém duas bolas brancas e três vermelhas. Retiram-se ao acaso, uma após a outra, com reposição. Qual a probabilidade de que as duas bolas sejam da mesma cor? Neste nono exemplo, utilizaremos a análise combinatória.

Este exemplo é uma repetição do item a do Exemplo 7. Aqui chegaremos ao mesmo resultado usando análise combinatória.

O evento de nosso interesse: bolas da mesma cor = $F = [(B_1 \cap B_2) \cup (V_1 \cap V_2)]$. Vimos que os eventos $(B_1 \cap B_2)$ e $(V_1 \cap V_2)$ são mutuamente exclusivos, então:

$$P(F) = P[(B_1 \cap B_2) \cup (V_1 \cap V_2)] = P(B_1 \cap B_2) + P(V_1 \cap V_2)$$

Vamos calcular, então, as probabilidades necessárias.

$P(B_1 \cap B_2)$ = (Nº de resultados para duas bolas brancas) / (Nº total de resultados)

$P(V_1 \cap V_2)$ = (Nº de resultados para duas bolas vermelhas) / (Nº total de resultados)

Os denominadores serão os mesmos para os dois quocientes: há um total de cinco bolas (“objetos”) disponíveis em número ilimitado (porque há reposição) para extrair em duas retiradas (“espaços”), resultando **n** = 5 e **k** = 2, então:

$$\text{Nº total de resultados} = \mathbf{n}^{\mathbf{k}} = 5^2 = 25$$

Nº de resultados para duas bolas brancas: há um total de duas bolas brancas (“objetos”) disponíveis em número ilimitado (porque há reposição) para extrair em duas retiradas (“espaços”), resultando **n** = 2 e **k** = 2, então:

$$\text{Nº total de resultados} = \mathbf{n}^{\mathbf{k}} = 2^2 = 4$$

Nº de resultados para duas bolas vermelhas: há um total de três bolas vermelhas ("objetos") disponíveis em número ilimitado (porque há reposição) para extrair em duas retiradas ("espaços"), resultando **n** = 3 e **k** = 2, então:

Nº total de resultados = $\mathbf{n}^{\mathbf{k}} = 3^2 = 9$

Então:

$P(B_1 \cap B_2)$ = (Nº de resultados para duas bolas brancas) / (Nº total de resultados) = 4 /25

$P(V_1 \cap V_2)$ = (Nº de resultados para duas bolas vermelhas) / (Nº total de resultados) = 9/25

Substituindo na fórmula:

$$P(F) = P[(B_1 \cap B_2) \cup (V_1 \cap V_2)] = P(B_1 \cap B_2) + P(V_1 \cap V_2)$$
$$= 4/25 + 9/25 = 13/25 = 0{,}52\ (52\%)$$

Então, se as retiradas forem feitas com reposição, a probabilidade de que as duas bolas sejam da mesma cor será igual a 0,52 (52%). Observe que é exatamente o mesmo resultado obtido no Exemplo 7. Claro que, para este caso extremamente simples (apenas duas retiradas com dois resultados possíveis em cada uma), o uso de análise combinatória não é necessário, mas permite chegar aos mesmos resultados que seriam obtidos com as técnicas anteriores. Se, porém, houver muitas retiradas e/ou muitas opções, se tornará indispensável.

Análise combinatória sem reposição

Continua havendo **n** objetos para colocar em **k** espaços, mas os objetos não estão mais disponíveis em número ilimitado: não há repetição ou não há reposição. A seleção de um dos objetos modifica a probabilidade de seleção dos outros: há dependência. Para calcular o número de maneiras possíveis de preencher os espaços, é preciso relembrar os conceitos de arranjos e combinações.

Os arranjos são utilizados para calcular o número de maneiras de dispor os **n** objetos nos **k** espaços, quando a **ordem** e a **natureza** dos objetos são importantes para o problema. O número de arranjos de **n** objetos distintos tomados **k** a **k** será:

$A_{n,k} = \frac{n!}{(n-k)!}$ n! significa fatorial de n: n x (n-1) x (n-2) x....x 1; lembrando que 0! = 1.

Cinco carros, disputando os três primeiros lugares em uma corrida. Há quantas maneiras diferentes de classificá-los? Vejamos no Exemplo 10.

Observe que há cinco objetos a dispor em três espaços, então **n** = 5 e **k** = 3. Os objetos não estão disponíveis em número ilimitado: uma vez definido o primeiro colocado, ele não pode simultaneamente ocupar a terceira posição. Outro aspecto importante é que importam tanto a **ordem** quanto a **natureza** dos objetos: há diferença se o corredor A não chegar entre os três primeiros, mas também há diferença se o corredor chegar em primeiro ou segundo. Sendo assim, serão usados arranjos.

$$A_{n,k} = \frac{n!}{(n-k)!} = \frac{5!}{(5-3)!} = \frac{5 \times 4 \times 3 \times 2!}{2!} = 60 \text{ maneiras.}$$

Então, há 60 maneiras de classificar os cinco carros nos três primeiros lugares.

As combinações são utilizadas para calcular o número de maneiras de dispor os **n** objetos nos **k** espaços, quando apenas a natureza dos objetos é importante para o problema. O número de combinações de **n** objetos distintos tomados **k** a **k** será:

$$C_{n,k} = \frac{n!}{k! \times (n-k)!}$$

Vamos ver, no exemplo a seguir (Exemplo 11), de quantas maneiras diferentes podemos selecionar três dentre cinco pessoas para uma tarefa?

Observe que, novamente, há cinco objetos a dispor em três espaços, então **n** = 5 e **k** = 3. Os objetos não estão disponíveis em número ilimitado: uma vez que uma pessoa seja selecionada, não poderá novamente ser escolhida. Neste caso, importa apenas a natureza dos objetos, apenas definir as pessoas que serão selecionadas. Sendo assim, serão usadas combinações.

$$C_{n,k} = \frac{n!}{k! \times (n-k)!} = \frac{5!}{3! \times (5-3)!} = \frac{5 \times 4 \times 3!}{3! \times 2!} = 10 \text{ maneiras.}$$

Então, há dez maneiras de selecionar três dentre cinco pessoas.

Exemplo 12: uma urna contém 18 bolas brancas, 15 vermelhas e dez azuis. Serão retiradas X bolas, sem reposição, e observadas suas cores.

a) Seja X = 8 (oito bolas). Qual a probabilidade de que as bolas sejam da mesma cor?

b) Seja X = 6 (seis bolas). Qual é a probabilidade de que duas sejam brancas, duas sejam vermelhas e duas sejam azuis?

Este problema seria extremamente trabalhoso para resolver usando uma árvore de probabilidades, por possuir várias retiradas com três resultados cada. Observe que não há reposição, portanto deve-se usar análise combinatória sem repetição: repare que não há interesse na ordem das bolas retiradas (tanto no item a quanto no item b), mas apenas na cor das bolas (na sua "natureza"), sendo assim devem-se usar combinações para calcular o número de resultados necessários para calcular as probabilidades.

a) Há uma grande quantidade de resultados possíveis para este problema, deve-se identificar o evento de interesse: oito bolas da mesma cor. Neste caso, oito bolas brancas, ou oito bolas vermelhas ou oito bolas azuis, evento união oito brancas com oito vermelhas com oito azuis. Chamando o evento oito bolas da mesma cor de F:

F = (8 brancas $\cup$ 8 vermelhas $\cup$ 8 azuis).

Observe que os três eventos acima são mutuamente exclusivos: as oito bolas retiradas não podem ser brancas e azuis simultaneamente. Então:

P(F) = P (8 brancas $\cup$ 8 vermelhas $\cup$ 8 azuis) = P(8 brancas) + P(8 vermelhas) + P(8 azuis)

Para calcular as probabilidades dos eventos, pode-se usar a definição clássica de probabilidade:

P(8 brancas) = (Nº resultados para 8 brancas) / (Nº total de resultados)

P(8 vermelhas) = (Nº resultados para 8 vermelhas) / (Nº total de resultados)

P(8 azuis) = (Nº resultados para 8 azuis) / (Nº total de resultados)

O denominador será o mesmo para todas as expressões. Há um total de 43 bolas (43 objetos, **n** = 43) para colocar em oito espaços (8 retiradas, **k** = 8), usando combinações:

$$\text{N}^{\text{o}} \text{ total de resultados } = C_{n,k} = \frac{n!}{k!\times(n-k)!} = \frac{43!}{8!\times(43-8)!} = 145008513$$

Para as bolas brancas. Há 18 bolas brancas (18 objetos, **n** = 18) para colocar em oito espaços (8 retiradas, **k** = 8), usando combinações:

$$\text{N}^{\text{o}} \text{ de resultados para oito brancas } = C_{n,k} = \frac{n!}{k\times(n-k)!} = \frac{18!}{8!\times(18-8)!} = 43758$$

Para as bolas vermelhas. Há 15 bolas vermelhas (15 objetos, **n** = 15) para colocar em oito espaços (8 retiradas, **k** = 8), usando combinações:

$$\text{N}^{\text{o}} \text{ de resultados para oito vermelhas } = C_{n,k} = \frac{n!}{k!\times(n-k)!} = \frac{15!}{8!\times(15-8)!} = 6435$$

Para as bolas azuis. Há dez bolas azuis (10 objetos, **n** = 10) para colocar em oito espaços (8 retiradas, **k** = 8), usando combinações:

$$\text{N}^{\text{o}} \text{ de resultados para oito azuis } = C_{n,k} = \frac{n!}{k!(n-k)!} = \frac{10!}{8!\times(10-8)!} = 45$$

Substituindo os valores diretamente na fórmula geral:

P(F) = P(8 brancas) + P(8 vermelhas) + P(8 azuis)

$$P(F) = \frac{43758}{145008513} + \frac{6435}{145008513} + \frac{45}{145008513} = 0{,}000346$$

Este valor tão baixo era esperado, devido à quantidade de bolas e ao número total de combinações possíveis.

Arredondando, a probabilidade de que as oito bolas retiradas sejam da mesma cor é igual a 0,0003 (0,03%).

b) Neste caso, há interesse em calcular a probabilidade de que duas bolas sejam brancas, e duas sejam vermelhas e duas sejam azuis, evento intersecção duas brancas com duas vermelhas com duas azuis. Chamando este evento de G: G = (2 brancas $\cap$ 2 vermelhas $\cap$ 2 azuis).

Para os casos de intersecção, o cálculo do número de resultados associados precisa ser feito da seguinte forma: os números de resultados possíveis associados a cada "subevento" componente devem ser multiplicados para obter o número de resultados da intersecção.

Saiba que isso, porém, não significa que os eventos sejam independentes!

P(G) = (Nº res. 2 brancas x Nº res. 2 vermelhas x Nº res. 2 azuis)/ (Nº total de resultados)

Há um total de 43 bolas (43 objetos, **n** = 43) para colocar em seis espaços (6 retiradas, **k** = 6), usando combinações:

$$\text{Nº de resultados} = C_{n,k} = \frac{n!}{k! \times (n-k)!} = \frac{43!}{6! \times (43-6)!} = 6096454$$

Nº de res. duas brancas: há 18 bolas brancas (18 objetos, **n** = 18) para colocar em dois espaços (2 retiradas, **k** =2), usando combinações:

$$\text{Nº de resultados para duas brancas} = C_{n,k} = \frac{n!}{k! \times (n-k)!} = \frac{18!}{2! \times (18-2)!} = 153$$

Nº de res. duas vermelhas: há 15 bolas vermelhas (15 objetos, **n** = 15) para colocar em dois espaços (2 retiradas, **k** =2), usando combinações:

$$\text{Nº de resultados para duas vermelhas} = C_{n,k} = \frac{n!}{k! \times (n-k)!} = \frac{15!}{2! \times (15-2)!} = 105$$

Nº de res. dois azuis: há dez bolas azuis (10 objetos, **n** = 10) para colocar em dois espaços (2 retiradas, **k** =2), usando combinações:

$$\text{Nº de resultados para duas azuis} = C_{n,k} = \frac{n!}{k! \times (n-k)!} = \frac{10!}{2! \times (10-2)!} = 45$$

Substituindo na fórmula de P(G):

P(G) = (153 x 105 x 45)/ (6.096.454) = 0,11858

Arredondando, a probabilidade de que duas bolas sejam brancas, e duas vermelhas e duas azuis é igual a 0,12 (12%).

Saiba mais...

- Sobre conceitos básicos de Probabilidade, BARBETTA, P. A. *Estatística Aplicada às Ciências Sociais*. 6. ed. Florianópolis: Ed. da UFSC, 2006, capítulo 7.
- Também sobre conceitos básicos de Probabilidade STEVENSON, Willian J. *Estatística Aplicada à Administração*. São Paulo: Harbra, 2001, capítulo 3.
- LOPES, P. A. *Probabilidades e Estatística*. Rio de Janeiro: Reichmann e Affonso Editores, 1999, capítulo 3.

RESUMO

O resumo desta Unidade está mostrado na Figuras 45:

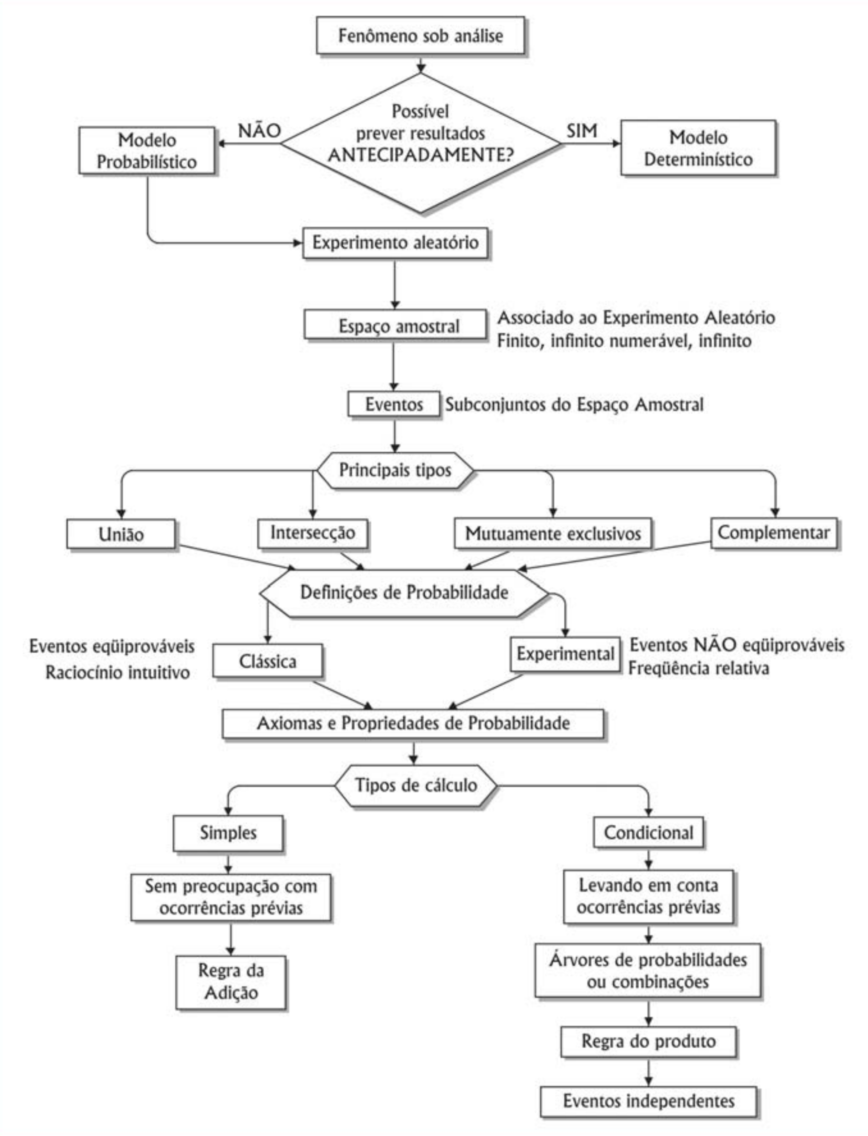

Figura 46: Resumo da Unidade 5
Fonte: elaborada pelo autor

Atividades de aprendizagem

As atividades de aprendizagem estão disponíveis no Ambiente Virtual de Ensino-Aprendizagem. Não deixe de respondê-las.

Chegamos ao final de Unidade 5. Esperamos que você tenha aprendido todos os conceitos trabalhados e, com os exemplo propostos, tenha colocado em prática as informações adquiridas. Neles propomos que você reconhecesse os modelos probabilísticos, modelos determinísticos, principais tipos de evento e os diferentes tipos de cálculo. Na Unidade 6, vamos prosseguir aprendendo o conceito de variável aleatória, que será indispensável para as Unidades 7, 8 e 9. Veremos, ainda, nas Unidades seguintes, a expansão do estudo para o conceito de variável aleatória e alguns dos modelos probabilísticos mais empregados. Tudo isso para chegarmos às Unidades 8 e 9, nas quais aplicaremos os conceitos de probabilidade no processo de inferência estatística, conforme já foi dito na Unidade 1.

Não desanime, caso tenha ficado alguma dúvida. Estamos com você sempre! Interaja, solicite auxílio e, caso necessário, releia o material. Realize a atividade de aprendizagem e entenda todo o processo amplamente.

Ótimos estudos!