

Princípios de Probabilidade

Termodinâmica Química

Prof. Guilherme Duarte, Ph. D.

Conforme vimos nas aulas anteriores, a entropia (S) é um dos conceitos centrais da Termodinâmica. Até este ponto, entretanto, apenas a observamos como uma quantidade macroscópica sem uma interpretação simples além de ser a função que dita a espontaneidade de processos em um sistema isolado. Conforme vimos em Química Geral, entretanto, a entropia está diretamente relacionada ao mundo microscópico e, para lidar dentro desse universo, precisamos entender um pouco sobre probabilidade e análise combinatória.

1 As regras da probabilidade são receitas para fazer inferências consistentes

Considere N como o número total de resultados de uma aposta. Se, desse total, n_A são resultados de uma categoria A , então dizemos que a **probabilidade** de ter um resultado A nesse jogo é:

$$P_A = \frac{n_A}{N} \quad (1)$$

Probabilidades necessariamente são números entre zero (resultado impossível) e um (resultado com 100% de chances de acontecimento) e podem ser computadas para combinações de eventos diferentes. Por exemplo, considere um dado de seis lados sem nenhum viés. A probabilidade de lançar o dado e obter o número 4 virado para cima é de $1/6$, pois cada face do dado tem $1/6$ de probabilidade de ficar voltada para cima após um lançamento. E se eu jogar dados três vezes seguidas, qual é a probabilidade de obter 4 voltado para cima três vezes seguidas? E qual é a probabilidade de dois lançamentos resultarem no número 3 e apenas o último resultar em 4? E se eu quiser saber a probabilidade de obter dois 2 e um 6 em qualquer ordem? Para determinar essas probabilidades, precisamos de algumas definições básicas.

Eventos mutuamente exclusivos: A_1, A_2, \dots, A_t são mutuamente exclusivos se a ocorrência de cada um impede o acontecimento de todos os outros. De forma simplificada, se A e B forem eventos mutuamente exclusivos, então se A ocorre, B não ocorre. Em lançamento de dados, se eu jogar um dado, somente uma face ficará para cima, logo se o lance resulta em 4 para cima, nenhum outro resultado deste dado é possível.

Eventos coletivamente exaustivos: A_1, A_2, \dots, A_t são coletivamente exaustivos se eles formam o conjunto completo de todos os resultados possíveis. Por exemplo, os números 1, 2, 3, 4, 5 e 6 são o conjunto coletivamente exaustivo de todos os resultados de um lançamento de um único dado. No caso de uma moeda, *cara* e *coroa* formam o conjunto coletivamente exaustivo de todos os resultados de um lançamento.

Eventos independentes: A_1, A_2, \dots, A_t são eventos independentes se cada um não tiver correlação com qualquer outro evento do conjunto. Por exemplo, cada resultado de lançamento de um

dado deve ser independente do outro, do contrário ou o dado é enviesado ou há um problema de desonestidade do jogador.

Multiplicidade: A multiplicidade de eventos é o número total de formas em que resultados diferentes podem acontecer. Se o número de resultados A for n_A , o número de resultados B for n_B e o número de resultados C for n_C , então o número total de combinações de resultados é dado pela multiplicidade W :

$$W = n_A \cdot n_B \cdot n_C \quad (2)$$

Se, por exemplo, eu tiver dois modelos de carro e três cores disponíveis, a multiplicidade de carros distintos que eu posso ter é $W = 2 \times 3 = 6$.

Para calcular as probabilidades de combinações de eventos, temos duas regras importantes, a regra da adição e a da multiplicação. Se os resultados A, B, \dots, E forem mutuamente exclusivos e ocorrerem com probabilidades $P_A = n_A/N, P_B = n_B/N, \dots, P_E = n_E/N$, onde N é o número total de observações, então a probabilidade de observar A ou B ou \dots ou E ($P(A \cup B \cup \dots \cup E)$) é igual a soma das probabilidades:

$$P(A \cup B \cup \dots \cup E) = \frac{n_A}{N} + \frac{n_B}{N} + \dots + \frac{n_E}{N} = \frac{n_A + n_B + \dots + n_E}{N}. \quad (3)$$

Observe que a notação acima remonta a diagramas de Venn: a probabilidade de observar os resultados A ou B ou \dots ou E é equivalente à união das áreas equivalentes a esses eventos em um espaço \mathcal{U} , como sugere o diagrama da Fig. (1). Se A, B, \dots, E forem resultados mutuamente exclusivos e coletivamente exaustivos, então:

$$n_A + n_B + \dots + n_E = N \quad (4)$$

e a Eq. (3) resulta em $P(A \cup B \cup \dots \cup E) = P_A + P_B + \dots + P_E = 1$.

Se os resultados A, B, \dots, E forem independentes, então a probabilidade de observar A e B e \dots e E é o produto das probabilidades:

$$P(A \cap B \cap \dots \cap E) = \frac{n_A}{N} \times \frac{n_B}{N} \times \dots \times \frac{n_E}{N} = P_A \cdot P_B \cdot \dots \cdot P_E. \quad (5)$$

Observe que ao lidar com as probabilidades de eventos independentes estamos lidando com a notação de intersecção (\cap) no diagrama de Venn. Assim, a probabilidade de observar dois resultados independentes A e B equivale a razão entre a área de $A \cap B$ e a área de \mathcal{U} .

Para entender a regra da adição, considere o lançamento de um dado. A probabilidade que o resultado do lançamento seja 1 é de $1/6$. A probabilidade que o resultado do lançamento seja 1 ou 6, é igual a $1/6 + 1/6 = 1/3$. Observe que os resultados são mutuamente exclusivos: no mesmo lance, o dado não pode dar 1 e 6. No caso da regra da multiplicação, considere o lançamento de dois dados (ou dois lançamentos de um dado). Note que o resultado de um dado é independente do outro. Dizemos que os dados não são correlacionados. Assim, a probabilidade que um lance resulte em 1 e o outro resulte em 6 é igual a $(1/6)(1/6) = 1/36$. Esse raciocínio se aplica apenas a eventos independentes ou mutuamente exclusivos. Quando há correlação entre os resultados obtidos, precisamos considerar *probabilidades condicionais* que serão objeto de estudo em outro curso.

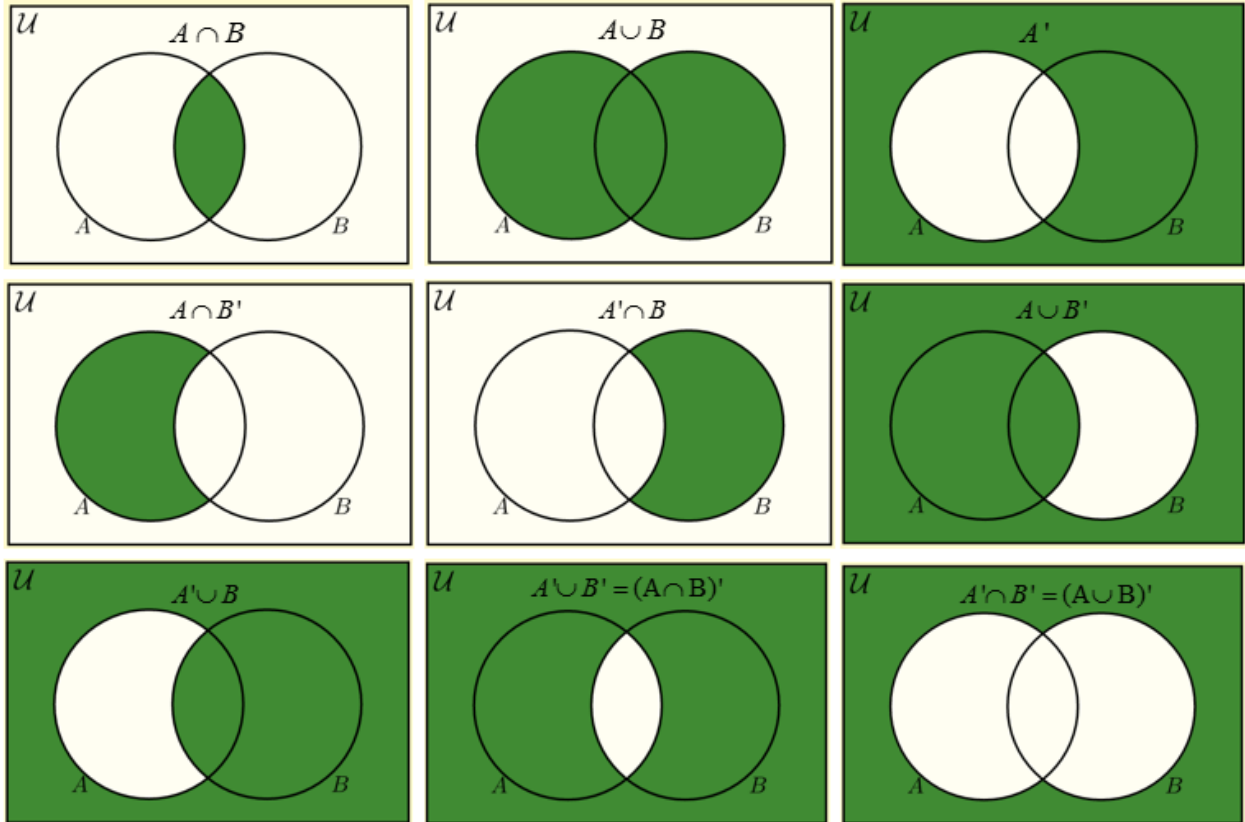


Figura 1: Diagramas de Venn podem ser usados para entender as regras básicas da probabilidade. Se A e B forem mutuamente exclusivos e quisermos saber a probabilidade de A ou B , nos referimos à proporção da área $A \cup B$ com relação à área total de \mathcal{U} . Se A e B forem independentes e quisermos saber a probabilidade de ocorrência de um evento A e B , então nos referimos à proporção da área $A \cap B$ com relação à área total de \mathcal{U} . O apóstrofe ' significa "não", assim $A \cup B'$ se refere a toda a área de \mathcal{U} , exceto a área pertencente **exclusivamente** a B .

2 Contar eventos é uma atribuição da análise combinatória.

Uma das preocupações principais da termodinâmica estatística é a contagem da quantidade de microestados disponíveis em um sistema. Para isso, precisamos entender melhor como fazer a contagem e o conceito de multiplicidade, que é central para o entendimento do significado físico da entropia. O conceito de "*desordem*", por exemplo, é associado ao número de maneiras que um sistema pode ser configurado. A análise combinatória se preocupa justamente com a composição dos eventos e não com a sequência deles. Se antes perguntávamos qual era a probabilidade de observar determinado evento, agora nos perguntamos de quantas formas podemos ordenar uma série de objetos.

Considere, por exemplo, as letras w , x , y , z . Quantas sequências diferentes podemos formar com apenas essas quatro letras, sem repetição? Podemos colocar a primeira letra, w , em qualquer uma das posições, mas posicionada uma letra, somente temos três espaços para preencher. Escolhido o lugar da terceira letra, temos duas posições restantes e assim preenchemos as posições até que não

tenhamos nenhuma letra disponível. Aplicando a Eq. (2) a esse problema, temos $W = 4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$.

Exemplo: considere uma caixa contendo cada uma das 26 letras do alfabeto. Qual é a probabilidade de extrair as letras uma a uma na ordem alfabética? Há duas formas de se pensar nesse problema:

(a) A probabilidade de extrair a letra “A” é de $1/26$. Extraída uma letra, temos somente 25 letras restantes em que a probabilidade de extrair a letra “B” é de $1/25$. Aplicando esse raciocínio para cada letra do alfabeto, temos que a probabilidade de obter uma sequência em ordem alfabética é:

$$P = \frac{1}{26} \cdot \frac{1}{25} \cdot \frac{1}{24} \cdots \frac{1}{1} = \frac{1}{26!}$$

(b) Para descobrir a probabilidade, precisamos saber o número total de sequências possíveis. Uma delas será a sequência em ordem alfabética. A multiplicidade, isto é, o número total de arranjos possíveis, é $W = 26!$. A probabilidade de escolher a sequência em ordem alfabética em um universo de $26!$ configurações é:

$$P = \frac{1}{W} = \frac{1}{26!}$$

Quando lidamos com N objetos distinguíveis, a multiplicidade de estados é dada por $N!$. Essa fórmula, entretanto, não se aplica a sistemas com objetos indistinguíveis em seu meio. Considere a palavra **banana**. Três letras **a** e duas letras **n** são indistinguíveis. Se fôssemos calcular a multiplicidade desconsiderando que algumas letras são repetidas, teríamos $W = 6! = 720$ arranjos distintos. Como três **a** e dois **n** são indistinguíveis, então temos que descontar as permutações dessas letras com respeito a todas as outras já que elas são idênticas. Assim:

$$W = \frac{6!}{3!2!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \times (2 \times 1)} = 60$$

Podemos generalizar essa fórmula para:

$$W = \frac{N!}{n_1!n_2! \dots n_k!}, \quad (6)$$

onde N é o total de objetos e n_i é o número de objetos do tipo i repetidos e indistinguíveis. Quando há somente duas categorias (e.g, sucesso/falha, cara/coroa, *spin up/spin down* etc):

$$W = \binom{N}{n} = \frac{N!}{n!(N-n)!}, \quad (7)$$

A interpretação dessa fórmula é simples: num jogo de “cara ou coroa” se todas as N sequências são distinguíveis entre si, então $W = N!$. Se todas as n jogadas resultantes em caras forem indistinguíveis, somente, então teríamos $W = N!/n!$. Se todas as jogadas resultantes em coroas também forem indistinguíveis, como $n_{coroa} = N - n$, então $W = N!/n!(N - n)!$

3 Coleções de probabilidade são descritas por funções de distribuição.

Para $i = 1, 2, \dots, t$ resultados mutualmente exclusivos, a função de distribuição é $P(i)$, o conjunto das probabilidades de todos os resultados. Se tal conjunto for coletivamente exaustivo, então:

$$\sum_{i=1}^t P(i) = 1. \quad (8)$$

Em diversas situações em física e química, lidamos com variáveis contínuas como posição (x), momentum (p), energia (E). Um caso famoso de Química Geral é a dependência espacial da função de onda eletrônica, $\psi(x)$. Sabemos que enquanto $\psi(x)$ é interpretado como “amplitude de probabilidade”, o quadrado do valor absoluto de ψ , $|\psi(x)|^2$ é identificado como a **densidade de probabilidade** em que a probabilidade de encontrar o elétron em um intervalo dx é igual a:

$$P(x) = |\psi(x)|^2 dx. \quad (9)$$

De forma mais geral, podemos interpretar a a probabilidade de encontrar um valor entre x e $x + dx$ é dada por

$$P(x) = \rho(x) dx, \quad (10)$$

onde ρ é chamada de densidade de probabilidade. A sua integral definida sobre todo o espaço deve ser:

$$\int \rho(x) dx = 1, \quad (11)$$

isto é, a densidade de probabilidade deve ser normalizada, isto é, os pesos estatísticos somados devem dar 1. Se a função de distribuição não for normalizada, podemos fazê-lo por:

$$\rho(x) = \frac{g(x)}{\int g(x) dx}. \quad (12)$$

A grande utilidade de funções de distribuição de probabilidade está em conter toda a informação que podemos saber de um sistema probabilístico. Em geral experimentos somente podem medir certos **momentos** de uma distribuição. Definimos o n -ésimo momento de uma distribuição $\rho(x)$ como:

$$\langle x^n \rangle = \int x^n \rho(x) dx = \frac{\int x^n g(x) dx}{\int g(x) dx}. \quad (13)$$

$\langle \rangle$ indicam os momentos ou valores esperados de uma função de distribuição. O 0-ésimo momento é sempre 1, pois a soma de todas as probabilidades (Eq. (11)) é sempre 1. O primeiro momento é chamado de **média** (*mean* ou *average* em inglês):

$$\langle x \rangle = \int x \rho(x) dx. \quad (14)$$

Se quisermos calcular a média de uma função $f(x)$:

$$\langle f(x) \rangle = \int f(x) \rho(x) dx. \quad (15)$$

Relações semelhantes valem para distribuições sobre t valores discretos:

$$\langle i \rangle = \sum i \cdot p(i), \quad (16)$$

$$\langle f(i) \rangle = \sum f(i) \cdot p(i). \quad (17)$$

Outra quantidade importante que podemos obter da distribuição é a **variância** (σ^2), uma medida da largura da distribuição: se ρ for largo, sua variância é grande; se ρ for estreito, sua variância é pequena. Ela é definida como:

$$\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle, \quad (18)$$

ou seja, o desvio quadrático médio com relação à média. Expandindo o polinômio dentro da Eq. (18):

$$\sigma^2 = \langle x^2 - 2\langle x \rangle x + \langle x \rangle^2 \rangle. \quad (19)$$

Como a integração é uma operação linear:

$$\sigma^2 = \langle x^2 \rangle - \langle 2\langle x \rangle x \rangle + \langle \langle x \rangle^2 \rangle. \quad (20)$$

A média da distribuição, $\langle x \rangle$, é um valor constante, logo:

$$\begin{aligned} \sigma^2 &= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2. \end{aligned} \quad (21)$$

A variância, σ^2 , é uma quantidade importante para entender propriedades termodinâmicas como capacidades caloríficas e sua raiz quadrada é chamada de **desvio padrão**, σ :

$$\sigma = \sqrt{\sigma^2}. \quad (22)$$