

1. Probabilidade

1.1 Frequência e probabilidade

Muitos conhecem o fato de que, quando realizamos o mesmo experimento várias vezes, o resultado não é sempre o mesmo, ainda que tentemos manter as condições do experimento o mais semelhantes possível. A razão disto é que alguns dos fatores que contribuem para o resultado do experimento não são (ou não podem ser) completamente **controladas**. Exemplos simples são “experimentos” de rolar dados, escolher uma carta do baralho, sortear o lado de uma moeda, ou qualquer outro dos chamados jogos de azar. Se atiramos quatro vezes uma moeda, como “cara” é tão provável quanto “coroa”, se a moeda não estiver “viciada”, deveríamos obter duas vezes “cara” e duas vezes “coroa”, mas isto dificilmente ocorre. Vamos chamar os resultados distintos de um experimento de **eventos simples**. Desta forma, o resultado de cada experimento é somente um e *apenas um* evento simples. Para dar uma representação matemática destes eventos, vamos atribuir um índice i a cada um dos eventos simples, chamados a partir de agora somente evento. Assim, os dois eventos possíveis quando atiramos uma moeda são “cara” ou “coroa” ($i = ca, co$), ao passo que na jogada de um dado há seis eventos possíveis ($i = 1, 2, 3, \dots, 6$).

Agora, quando fazemos o mesmo experimento N vezes, um dos eventos i pode ocorrer mais de uma vez, vamos imaginar que o evento i se repita n_i vezes. Se repetimos o mesmo experimento em uma outra ocasião, esperamos obter aproximadamente o mesmo valor para n_i . Por exemplo, se jogarmos uma moeda 100 vezes, esperamos obter em torno de 50 caras e 50 coroas (talvez obtenhamos 51 caras e 49 coroas ou 47 caras e 53 coroas). Para descrever esse fato, vamos definir a razão

$$F_i = \frac{n_i}{N} \quad (1)$$

Esta razão é a fração dos N experimentos que resultou no evento i e recebe o nome de frequência do evento i . Embora conhecer o valor de F_i obtido em algum conjunto de experimentos seja útil, é importante perceber que se fizermos novamente N experimentos, o número de eventos i pode ser diferente do anterior. O novo resultado pode ser m_i . Isto quer dizer que F_i , em geral, é diferente para cada conjunto de experimentos. Assim, no exemplo acima,

teríamos três valores diferentes para $F_{ca} = 0,50; 0,51; 0,47$. Ou seja, a frequência de um evento depende do grupo de experimentos que estamos considerando.

Mas gostaríamos que houvesse uma grandeza que não dependesse do conjunto específico de experimentos, mas que indicasse o valor esperado para frequência em qualquer conjunto de experimentos. Por exemplo, se jogarmos a moeda 20 vezes, poderíamos obter $F_{ca} = 0,4$ (8 caras); se jogarmos a moeda 100 vezes, poderíamos obter $F_{ca} = 0,54$ (54 caras); se jogarmos a moeda 1000 vezes, poderíamos obter $F_{ca} = 0,51$ (510 caras). Na medida em que N vai se tornando maior, esperamos, se a moeda não for viciada, que F_{ca} se aproxime de 0,5. Chamamos este valor limite de probabilidade do evento (neste caso, “cara”). Assim, temos um método para obter esta quantidade que denominamos probabilidade P_i , que está relacionada com a frequência que esperamos obter em experimentos futuros.

Podemos dar uma definição matemática para a probabilidade de um evento i da seguinte maneira:

$$P_i = \lim_{N \rightarrow \infty} F_i = \lim_{N \rightarrow \infty} \frac{n_i(N)}{N} \quad (2)$$

Estamos estabelecendo que P_i é o valor limite de F_i quando N torna-se arbitrariamente grande. Que, por sua vez, é o limite de n_i / N , onde n_i depende do valor de N .

Há duas maneiras de interpretar a equação (2), e ambas são utilizadas na mecânica estatística. Dissemos que N representa o número de experimentos e que n_i é o número de resultados em que ocorreu o evento i . Essa afirmação pode ser interpretada de duas maneiras equivalentes:

- Há um sistema físico, sobre o qual foram realizados N experimentos, um depois do outro (1 moeda, N jogadas, uma depois da outra);
- Há N sistemas físicos idênticos e um experimento para cada um, todos ao mesmo tempo (N moedas idênticas, que são jogadas ao mesmo tempo).

Por exemplo, podemos jogar uma moeda 1000 vezes, e obter 510 caras. Ou de posse de 1000 moedas idênticas, jogarmos todas ao mesmo tempo, nas mesmas condições em que jogamos a moeda da primeira situação, e obtermos 510 caras. Em qualquer das duas situações temos $n_{ca} = 510$ e $F_{ca} = 0,51$. Vamos, de agora em diante, supor que as duas situações são completamente

equivalentes, ou seja, que os dois métodos levam ao mesmo resultado. Por isto o tempo não é considerado nesta probabilidade.

Na mecânica estatística, o conjunto de sistemas idênticos é denominado **ensemble**. Mas como definir *sistemas idênticos*? Definimos dois sistemas como idênticos se não tivermos nenhum método *macroscópico* para distingui-los.

Embora a definição (2) seja interessante, em princípio, na prática sempre realizamos um experimento um número finito de vezes. Ou construímos um ensemble com um número finito de sistemas. Isso quer dizer que o limite N tendendo ao infinito não pode ser realizado em nenhuma situação física. O que podemos fazer é obter valores aproximados para a probabilidade de um evento. Apesar disso, dizemos que a probabilidade de obter “cara” é $\frac{1}{2}$, e que a probabilidade de obter um ás de copas no baralho é $\frac{1}{52}$. Estas afirmações baseiam-se na **hipótese de equiprobabilidade**, isto é, que as probabilidades de determinados eventos são iguais. Vamos ver como essa hipótese justifica as afirmações sobre a probabilidade de “cara” ou de um “ás de copas”. Para isto, precisamos considerar duas propriedades da probabilidade:

1. A frequência de qualquer evento é um número positivo, portanto, devido à definição (2), todas as probabilidades P são números positivos;
2. Se jogamos uma moeda N vezes (ou N moedas uma vez), e obtemos n_{ca} “caras”, o número de “coroas” será $n_{co} = N - n_{ca}$. Como $F_{ca} = n_{ca}/N$ e $F_{co} = n_{co}/N$, isso significa que $n_{co} + n_{ca} = N$. E $F_{ca} + F_{co} = 1$ Generalizando, a soma de todos os n_i é igual ao número total N de experimentos, e a soma de frequências F_i de um conjunto de eventos deve ser igual à unidade, e a soma de todos os P_i também.

Podemos descrever matematicamente estas duas propriedades importantes da probabilidade da seguinte maneira:

$$P_i \geq 0, \text{ para todos os eventos simples.} \quad (3)$$

$$\sum_i P_i = 1, \text{ soma sobre todos os possíveis eventos.}$$

Agora veremos porque a hipótese de equiprobabilidade aplicada sobre a segunda propriedade leva ao resultado $\frac{1}{2}$ para a probabilidade de “cara”. Temos $P_{ca} + P_{co} = 1$ e $P_{ca} = P_{co} = P_0$, portanto $P_{ca} + P_{co} = 2P_0 = 1$ e $P_{ca} = P_{co} = P_0 = \frac{1}{2}$. Da mesma maneira, se o baralho tem todas as cartas, 52, supomos que a probabilidade de tirar qualquer carta independe do valor da carta, isto é

$P_i=1/52$, qualquer que seja a carta i . Embora estes casos que consideramos pareçam muito simples, o mesmo tipo de lógica é utilizado em casos mais complexos.

O ponto importante aqui, é que, mesmo sem realizar os experimentos, podemos *fazer hipóteses sobre as probabilidades* e, utilizando suas propriedades, deduzir valores para as mesmas. O procedimento que utilizamos acima é muito comum.

Se há não há razão aparente para que um evento ocorra com maior frequência que outro, então supomos que suas probabilidades são iguais.

Falamos então de dois tipos de probabilidade:

1. Probabilidades baseadas enunciada acima, são chamadas de probabilidades *a priori* (antes do experimento).
2. Probabilidades resultantes de conjuntos de experimentos, com a extrapolação para um número infinito de experimentos, como na definição (2), são chamadas de *a posteriori* (depois do experimento).

As teorias físicas são desenvolvidas em termos de probabilidades *a priori*. Se a teoria for capaz de predizer o resultado de experimentos, concluímos que as hipóteses adotadas são boas. Caso contrário, será necessário rever nossas hipóteses e propor outras, que levarão a outros valores de probabilidade. No caso da moeda, se imaginamos probabilidades iguais para “cara” e “coroa” e testarmos essa hipótese em uma moeda, e obtemos muito mais caras do que coroas, então concluímos que nossa hipótese não é boa para aquela moeda. Mas quando nosso interesse é o comportamento molecular, acreditamos que todas as moléculas são iguais, então uma hipótese boa (testada em suas previsões em experimentos) vale para qualquer conjunto de moléculas.

1.2 Probabilidade de eventos compostos: eventos independentes

Um conceito importante em teoria de probabilidade é o espaço amostral. Este “espaço” é constituído de todos os eventos possíveis e é representado em termos de um conjunto de pontos, como na figura abaixo. Cada ponto representa um evento simples, ou seja um resultado possível do experimento. Neste “espaço”, não estamos interessados em distâncias, nem no arranjo espacial dos pontos, mas apenas nos pontos em si.

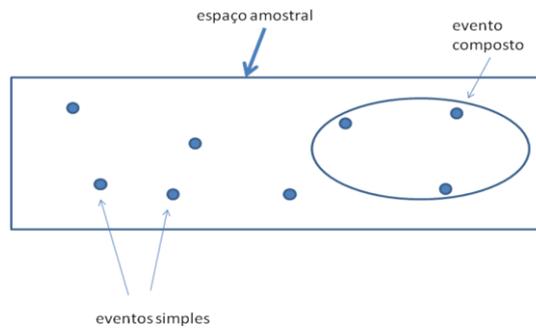


Figura 1

A figura 1 poderia ilustrar os eventos possíveis em um lançamento de três moedas. O evento composto poderia ser aquele em que o resultado do experimento são duas caras e uma coroa.

Um evento composto é uma coleção de pontos no espaço amostral. A coleção de pontos do evento composto é definida por uma característica comum dos eventos pertencentes à coleção. Vamos chamar o evento composto da figura acima de A . Todos os pontos no interior da fronteira pertencem a A . A probabilidade deste evento composto $P(A)$ é definida por

$$P(A) = \sum_{i \in A} P_i \quad (4)$$

onde a soma é sobre todos os pontos do evento composto (indicados simbolicamente por $i \in A$, que significa todos os i s contidos no evento composto A).

Exercício: Represente o espaço amostral correspondente aos eventos possíveis ao retirar-se uma carta do baralho de 52 cartas. Represente neste espaço os seguintes eventos compostos: (1) i: as cartas de espadas; (2) ii: as cartas de número 2; (3) iii: ases pretos. Calcule a probabilidade de cada um dos eventos compostos, utilizando a equação (4).

Dois eventos compostos podem ter eventos em comum. No caso acima, os eventos i (cartas de espadas) e ii (cartas de número 2) possuem um evento simples em comum, a carta de espadas de número 2. Já os eventos compostos ii (cartas de número 2) e iii (ases pretos) não possuem pontos comuns.

Eventos compostos que não tem pontos em comum são denominados disjuntos. Na figura abaixo, os eventos A e B são disjuntos.

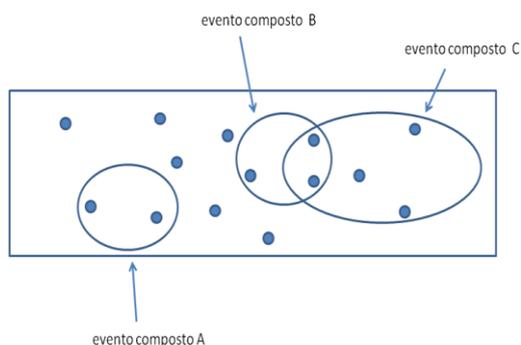


Figura 2

É útil também definir o conceito de união: o conjunto composto dos pontos de A e B representa um novo evento composto. Mas qual o significado deste novo evento? Este evento composto (pontos de A + pontos de B) representa os possíveis resultados de um experimento em que obtemos OU eventos A OU eventos B. Por exemplo, ases pretos e cartas de copas são eventos disjuntos.

Se estivermos interessados em obter, em uma jogada qualquer, um ás preto OU uma carta de espadas, o evento que nos interessa é a união de A com B. Representamos matematicamente a união através da expressão $A \cup B$. De acordo com a definição (4), a probabilidade deste evento, $P(A \cup B)$, é a soma das probabilidades de todos os pontos do evento. Como A e B não tem pontos em comum (são eventos disjuntos), essa probabilidade é a soma de $P(A)$ e $P(B)$, isto é,

$$P(A \cup B) = P(A) + P(B). \quad (5)$$

Nos casos em que dois eventos **não** são disjuntos (por exemplo, B e C da figura acima), $P(B \cup C)$ ainda é a soma das probabilidades de todos os pontos em B e C, e no entanto a expressão (5) para $P(B \cup C)$ **não** é válida. O problema é que os pontos de B e C **em comum** seriam contados **duas** vezes. Para consertar a contagem dupla, temos que subtrair a probabilidade de todos os pontos que B e C possuem em comum. Então definimos a intersecção de B e C como o evento composto que contém todos os pontos comuns a B e C, que descrevemos como $B \cap C$.

Na figura 3 podemos ver a diferença entre união e intersecção.

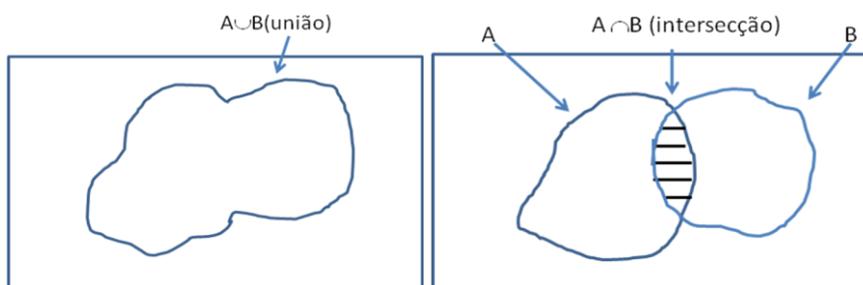


Figura 3

Na figura 2, a probabilidade de cada um dos eventos simples é $1/14$, e a probabilidade $P(B \cap C) = 2/14$. Por outro lado, como A e B não possuem pontos em comum, a probabilidade $P(A \cap B) = 0$. Podemos escrever uma relação geral para a intersecção,

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y). \quad (6)$$

No exemplo da figura 2, $P(A \cup B) = 2/14 + 3/14 = 5/14$ e $P(B \cup C) = 3/14 + 5/14 - 2/14 = 6/14$.

Uma característica importante no caso de eventos compostos é que, diferentemente de eventos simples, dois eventos compostos podem ter pontos em comum (sua intersecção). Eventos simples, por outro lado, não possuem pontos em comum, pois, por definição, o resultado de um experimento é *um e apenas um* evento simples. Assim, eventos simples são sempre disjuntos.

Para determinar a probabilidade da união de alguns eventos simples (de um evento composto), simplesmente somamos as probabilidade relevantes, usando a definição da equação (4). A equação (5) mostra que não há diferença essencial entre eventos compostos *disjuntos* e eventos simples, pois adicionamos as probabilidades da mesma maneira nos dois casos. É apenas quando eventos compostos possuem alguma intersecção que devemos tomar alguns cuidados. Vimos que (6) descreve a forma correta de somar probabilidades. De agora em diante iremos nos referir simplesmente a eventos, tanto no caso de eventos simples como no caso de eventos compostos, e utilizaremos sempre a equação (8). Se os eventos forem disjuntos, a equação (8) se reduz à equação (7).

O objetivo desta seção é levar à noção de eventos independentes. Para dar precisão a este conceito vamos primeiro considerar a ideia de probabilidade condicional. Vamos supor que um evento A pode ocorrer depois de um evento B. Queremos saber qual a probabilidade do evento A se o evento B tiver algum

efeito sobre a probabilidade do evento A. Medimos este efeito em termos da probabilidade condicional. Por exemplo, podemos perguntar “Qual é a probabilidade de que tiremos do baralho completo uma carta de “copas” (evento A) se sabemos que ela é um 3 (evento B)?”.

O fato de que a carta é um “3” não tem efeito nenhum sobre o naipe, portanto a probabilidade de que esta carta seja de “copas” é $\frac{1}{4}$, e não importa em nada o fato de que seja um “3”. Por outro lado, se soubermos que a carta é vermelha, ela tem que ser ou “copas” ou “ouro”, e a probabilidade de que seja “copas” aumenta para $\frac{1}{2}$. Neste segundo caso, sabermos que a carta vermelha (evento C) tem efeito sobre a probabilidade de obtermos o evento B (carta de “copas”). Definimos a probabilidade condicional por

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} \quad (7)$$

A expressão $P(X|Y)$ significa a probabilidade de que o evento X ocorra, dado que o evento Y ocorreu no experimento. Essa definição está de acordo com nossos dois exemplos:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{52}}{\frac{4}{52}} = \frac{1}{4},$$

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{1/52}{2/52} = \frac{1}{2}.$$

Assim, embora a definição (7) seja boa para qualquer caso, é interessante que esteja de acordo com nossa intuição nos casos em que a intuição é suficiente para nos dar a resposta.

É claro que se X e Y não tem pontos em comum (são eventos disjuntos), então $P(X \cap Y) = 0$, e o evento X não pode ocorrer, se Y ocorreu. O que está de acordo com a definição (7). Por exemplo, se sabemos que a carta que tiramos do baralho completo é preta, então a probabilidade de que ela seja “copas” é zero, pois a intersecção cartas pretas com cartas vermelhas é zero. Observe que, em geral $P(X|Y)$ não é igual a $P(Y|X)$. No exemplo acima, $P(B|A)=1/13$, ao passo que $P(A|B)=1/4$.

Note que no nosso exemplo com as cartas do baralho, temos

$$P(A \cap B) = P(A) \quad \text{e} \quad P(B \cap A) = P(B). \quad (8)$$

De acordo com a equação (7), essas duas relações implicam que

$$P(A \cap B) = P(A)P(B). \quad (9)$$

Essa relação só é verdadeira para A e B se os dois eventos forem independentes. No caso de eventos independentes, o fato de acontecer A não tem efeito nenhum sobre a probabilidade de obter B.

Mas, cuidado! Como vimos, nem todos os eventos são independentes. No exemplo das cartas,

$$P(A \cap C) = \frac{1}{52} \neq P(A)P(C) = \left(\frac{1}{4}\right)\left(\frac{1}{26}\right) = \frac{1}{104}.$$

O ponto essencial que é preciso entender é o conceito da independência, porque ele vai ser crucial no desenvolvimento da mecânica estatística. Por definição, dois eventos A e B são independentes apenas no caso em que a equação (9) é satisfeita, o que por sua vez exige que a equação (8) seja válida.

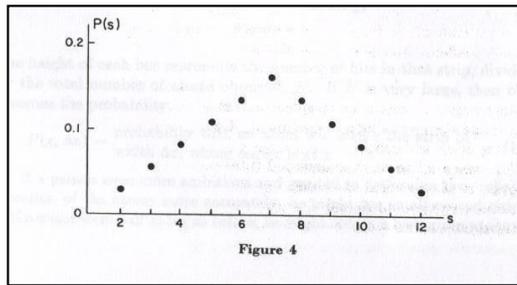
Podemos definir a independência de eventos da seguinte maneira. **Se a ocorrência de um evento não afeta as chances de ocorrência de outro evento, então os dois eventos são independentes.**

1.3 Funções de distribuição

A representação do espaço amostral por meio de pontos é muito útil para organizarmos as ideias. No entanto, como todo instrumento, tem suas limitações, e há muitas situações em que ela não é muito útil. Se os pontos do espaço amostral tem probabilidade diferentes, precisamos de uma representação melhor do que simplesmente escrever a probabilidade ao lado de cada ponto. Vamos considerar um exemplo com dois dados: em muitos jogos, o que importa é a soma dos números que ficam com a face para cima nos dois dados. Se todas as faces dos dois dados tem igual probabilidade de ficarem “para cima”, pode-se mostrar (você saberia fazê-lo?) que a probabilidade da soma ser s $P(s)$

s	2	3	4	5	6	7	8	9	10	11	12
$P(s)$	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

Nesse caso, o espaço amostral tem 11 pontos, com 6 valores diferentes de probabilidade. O espaço amostral não serve para mostrar as probabilidades $P(s)$. Podemos utilizar a tabela, mas há outra forma gráfica, que tem a vantagem da visualização. O gráfico que representa as probabilidades desse exemplo está mostrado na figura 4.



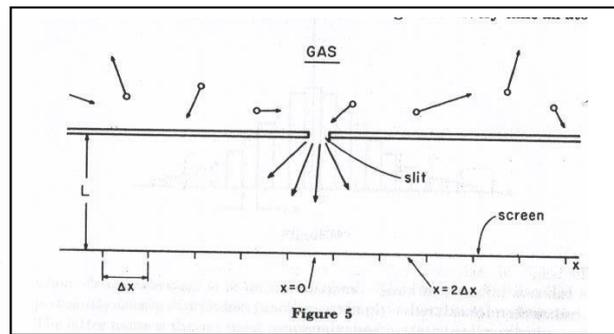
Se representamos $P(s)$ dessa maneira, percebemos que a probabilidade é na verdade uma *função* de s . E isto é verdade em geral. Isto é, a probabilidade $P(i)$ de um evento descrito pelo número i é na verdade uma *função* de i e recebe o nome de **função de distribuição de probabilidades**. O nome deriva do fato de que $P(i)$ determina como a probabilidade se distribui sobre os eventos i . Essa função tem as propriedades da probabilidade

$$P(i) \geq 0, \quad \sum_i P(i) = 1 \quad (10)$$

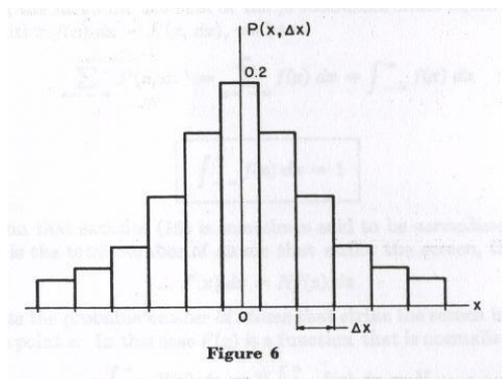
Vamos agora considerar uma generalização importante da função de distribuição de probabilidades que acabamos de discutir. Até aqui, para simplicidade, discutimos “experimentos” muito simples. Um experimento mais realista poderia envolver um tanque de gás em que há uma pequena fenda que liga o tanque a um compartimento onde é mantido vácuo através de uma bomba.

Gostaríamos de estabelecer a probabilidade de que um átomo do gás que escapa do tanque para o vácuo esteja se movendo em dada direção. Poderíamos utilizar uma tela fluorescente, como ilustrado na figura 5. Cada vez que um átomo atinge a tela, esta cintila o que nos permite determinar a direção de que veio o átomo, pois sabemos que veio do furo. A primeira diferença importante em relação aos experimentos discutidos anteriormente é que há um espectro **contínuo** de eventos – isto é, **qualquer ponto** da tela pode ser atingido por um átomo. As chances de que um átomo atinja um ponto (matemático) qualquer são **nulas**, por mais tempo que esperemos que ocorra (já que o número total de eventos simples é infinito!). O que podemos definir **não** é a probabilidade de que o átomo viaje em uma determinada **direção precisa**, mas sim a **probabilidade de que se desloque “aproximadamente” em uma**

determinada direção. Mas o que quer dizer isto, do ponto de vista matemático?



Vamos dividir a tela em faixas paralelas à direção da fenda e de largura Δx (por exemplo, 1cm). Cada vez que um átomo atinge **qualquer ponto dentro** da faixa, contabilizamos o evento. Agora classificamos os resultados do experimento em termos das faixas atingidas pelos átomos. Os eventos do nosso espaço amostral são as faixas atingidas. Para determinar experimentalmente a probabilidade (ou a frequência) com que um átomo atinge determinada faixa com centro em x , precisamos contar o número de átomos que atingem esta faixa, $N(x)$, e dividir pelo número total de átomos que atingem a tela (retome a definição de frequência da seção 1). Os resultados do experimento podem ser colocados em um histograma, como o da figura 6.

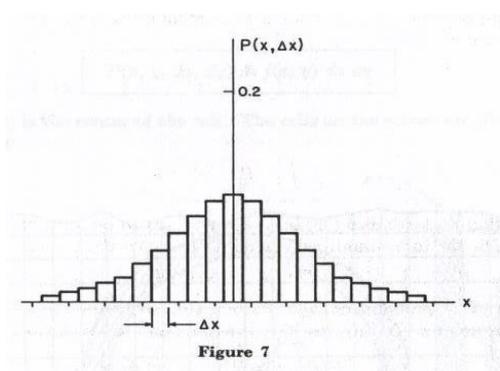


A altura de cada barra representa o número de átomos que atingiu a faixa, dividido pelo número total de átomos, N . Se N for muito grande, esta frequência pode ser identificada com a probabilidade.

$$P(x, \Delta x) = \text{probabilidade de um átomo atingir a faixa com centro em } x \text{ e de largura } \Delta x. \quad (11)$$

Se necessitarmos de maior precisão para o experimento, podemos definir faixas de largura menor, isto é, podemos diminuir Δx . Se fizermos Δx duas vezes

menor, poderíamos obter o histograma da figura 7. A informação é agora mais detalhada do que no caso anterior. Note que o número de átomos que cintila em determinada faixa é menor, e portanto a probabilidade de que um átomo atinja determinada faixa é menor. O que está de acordo com nossa afirmação anterior, de que se quisermos definir exatamente a direção de movimento do átomo obteríamos uma probabilidade nula (situação que corresponderia a uma faixa de largura zero, ou seja, uma linha).

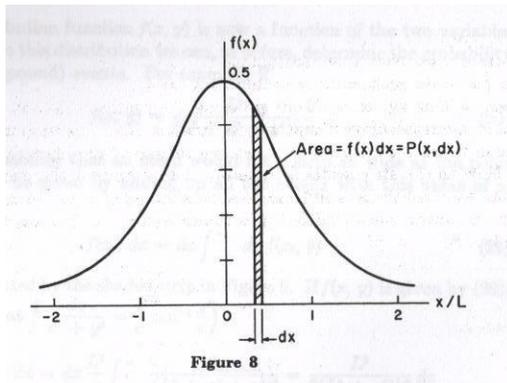


A medida que diminuimos Δx , verifica-se que o histograma vai tornando-se mais homogêneo e podemos aproximá-lo por uma função de x . Como, no entanto, $P(x)$ tende a zero, à medida que Δx tende a zero, consideramos a seguinte função

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x, \Delta x)}{\Delta x}, \quad \text{ou}$$

$$f(x)dx = P(x, dx). \quad (12)$$

em que dx é infinitesimal. A função $f(x)$ é chamada de função distribuição de densidade de probabilidade, ou simplesmente, **função distribuição**. Mas notem que $f(x)$ **não** é a **probabilidade de que um átomo atinja o ponto x** , pois esta seria nula! O que definimos é a probabilidade $P(x, dx)$ de que um átomo atinja a faixa centrada em x , com largura infinitesimal dx , e que é dada por $f(x)dx$. Isto está ilustrado na figura 8 abaixo.



Como vimos antes, a definição de probabilidade requer que soma de todas as probabilidades seja igual à unidade. Matematicamente, temos

$$\sum_{x=-\infty}^{\infty} P(x, dx) = \sum_{x=-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} f(x)dx,$$

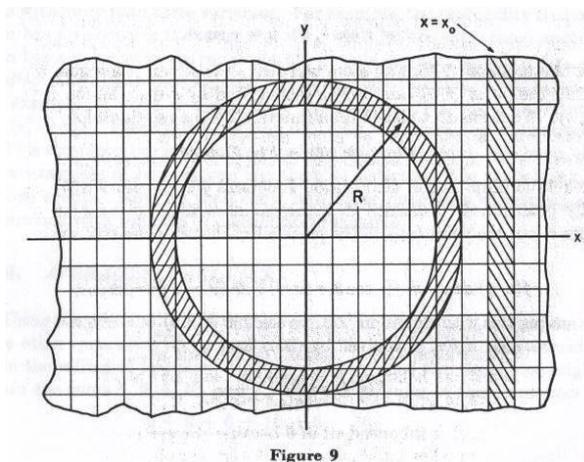
assim temos como condição

$$\int_{-\infty}^{\infty} f(x)dx = 1. \quad (13)$$

Costumamos denominar uma função que satisfaz a equação (13) de **normalizada para 1**.

A função distribuição $f(x)$ pode ser utilizada para calcular a probabilidade de eventos compostos da mesma forma que a probabilidade $P(x, dx)$.

Vamos imaginar uma situação em que o furo seja circular, ao invés de retangular. Neste caso, temos que dividir a tela em pequenos quadradinhos de lados Δx e Δy (veja figura 9 abaixo) e registrar o número de átomos que chega em cada quadradinho centrado em (x,y) .



Para um número muito grande de eventos, isto é, de átomos, podemos determinar a probabilidade de que um determinado quadradinho infinitesimal seja atingido, $P(x, y, dx, dy)$, que será dada por $P(x, y, dx, dy) = f(x, y)dxdy$, em

que $f(x, y)$ é uma função bem definida, a função de distribuição de densidade de probabilidade.

Neste último caso, podemos perguntar qual a probabilidade de encontrarmos um átomo a uma distância R do centro da tela. Este é um evento composto, que corresponde a soma de todos os eventos em que um átomo atinge qualquer ponto (x, y) , desde que $x^2 + y^2 = R^2$.

Para um ponto qualquer (x, y) , podemos escrever, em coordenadas polares, que $f(x, y)dx dy = f(r \cos(\theta), r \sin(\theta))r dr d\theta$. Mas se supomos que não há direções privilegiadas, isto é, que os átomos desviam em todas as direções θ **com igual probabilidade**, então podemos escrever que $f(x, y)$ tal que $x^2 + y^2 = R^2$, seja $f(r \cos(\theta), r \sin(\theta))$ para qualquer θ , Então podemos simplesmente obter $f(R)$. Então,

$$P_R = \sum_{x, y, x^2 + y^2 = R^2} f(R) r dr d\theta = \int f(R) r dr d\theta. \quad (14)$$

Há muitas situações em que é necessário utilizar mais de duas variáveis. No caso de uma molécula de gás, por exemplo, podemos perguntar qual a probabilidade de que possua a componente x de sua velocidade no intervalo dv_x , a componente y de sua velocidade no intervalo dv_y e a componente z de sua velocidade no intervalo dv_z . Em termos numéricos poderíamos perguntar qual a probabilidade de que sua velocidade seja $(200 \pm dv_x, -511 \pm dv_y, 103 \pm dv_z)$?

A probabilidade pode ser escrita em termos de uma função distribuição das 3 variáveis (v_x, v_y, v_z) , da seguinte forma

$$P(v_x, v_y, v_z, dv_x, dv_y, dv_z) = f(v_x, v_y, v_z) dv_x dv_y dv_z. \quad (15)$$

Um exemplo desta função distribuição é a função distribuição de Maxwell para as velocidades das moléculas de um gás, que veremos adiante,

$$f(\mathbf{v}) = \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} e^{-\alpha v^2}.$$

Mas há pode haver função distribuição de qualquer número de variáveis. Podemos perguntar qual a probabilidade de encontrarmos uma molécula na posição $(x \pm dx, y \pm dy, z \pm dz)$ com velocidade $(v_x \pm dv_x, v_y \pm dv_y, v_z \pm dv_z)$, então a função de distribuição será uma função de 6 variáveis, $f(x, y, z, v_x, v_y, v_z)$.

1.4 Médias e variâncias

Em um experimento repetido várias vezes (vários eventos) podemos obter resultados diferentes em cada experimento, ou pode ser que um resultado ocorra um certo número de vezes. Qual é o “resultado médio”?

Vamos retomar o exemplo do jogo de dois dados, da seção anterior. Vamos supor que, em 10 jogadas, obtemos os seguintes resultados: 2,8,6,7,7,11,3,6,8 e 2. O valor médio destes números, que representamos por \bar{x} , é $(2+8+6+7+7+11+3+7+8+2)/10 = 6,1$. Mas note que podemos reorganizar esta soma na forma $(2 \times 2 + 1 \times 3 + 1 \times 6 + 3 \times 7 + 2 \times 8 + 1 \times 11)/10 = 6,1$. Os resultados 3, 6 e 11 ocorreram com frequência 1/10, os resultados 2 e 8 ocorreram com frequência 2/10 e o resultado 7 ocorreu com frequência 3/10. As duas formas de organizar os resultados do experimento estão apresentadas na tabela:

Resultado x_i do experimento i	Número total de experimentos	Resultado x_k possível do experimento	Frequência do resultado x_k	Número total de resultados possíveis
2	n=10	2	2	N=12
8		3	1	
6		4	0	
7		5	0	
7		6	1	
11		7	3	
3		8	2	
7		9	0	
8		10	0	
2		11	1	
		12	0	

As diferentes organizações dos resultados dos experimentos correspondem diferentes formas de escrevermos o **valor médio** de x

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}, \quad (17)$$

$$\bar{x} = \sum_{k=1}^n x_k F_k, \quad (18)$$

onde $x_{i,k}$ corresponde ao resultado de um dos experimentos, n é o número total de experimentos e F_k é a frequência do evento.

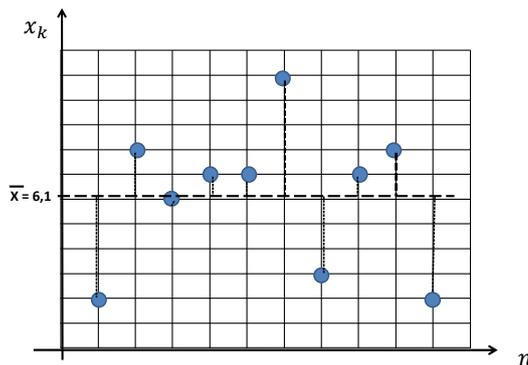
Como vimos na seção 1, a probabilidade está diretamente relacionada com a frequência que esperamos obter em um experimento. Utilizando os conceitos da seção 1, podemos escrever a expressão para o valor médio de x^m em termos de probabilidades

$$\overline{x^m} = \sum_{k=1}^n x_k^m P_k. \quad (19)$$

Esta é a definição de média que utilizaremos daqui por diante.

Muitas vezes queremos saber se os resultados dos experimentos se afastam muito ou pouco da média. No “experimento” com os dois dados, que discutimos acima, a “distância” de cada medida da média dos resultados pode ser representada na forma apresentada no gráfico abaixo. Observe que esta “distância”, $x_i - \bar{x}$, pode ser positiva ou negativa. Então, se somarmos todas essas distâncias, para obter a média dos afastamentos, podemos obter um afastamento pequeno, por causa do cancelamento. Em nosso exemplo, teríamos um afastamento dos resultados do valor médio nulo,

$$\overline{x_i - \bar{x}} = \sum_{k=1}^n \frac{x_i - \bar{x}}{n} = 0, \quad (20)$$



Por isso, para termos uma ideia melhor deste afastamento é interessante definir o quadrado desta distância, que é chamado de variância. A variância ou dispersão é definida por

$$\sigma = \overline{(x - \bar{x})^2} = \overline{x^2} - \bar{x}^2. \quad (21)$$

Calcule o valor de σ para nosso exemplo do jogo de dois dados.

Como definir o valor médio e a variância **quando o resultado do experimento é um número real**? Nesse caso, temos que trabalhar com a **função distribuição**. A média é definida pela expressão

$$\bar{x}^m = \int x^m f(x) dx. \quad (22)$$

A variância é definida por da mesma forma que (21).

Um exemplo importante de função distribuição é a **função gaussiana**, dada por

$$f(x) = C e^{-\alpha x^2} \quad (23)$$

Para que $f(x)$ represente uma função de distribuição de probabilidades é preciso que a soma das probabilidades correspondentes seja igual a 1, isto é, que

$$\int f(x) dx = \int C e^{-\alpha x^2} dx = 1, \quad (24)$$

o que impõe uma condição sobre a constante C , que deve ser dada por

$$C = \sqrt{\frac{\pi}{\alpha}}. \quad (25)$$

Quanto a média, temos

$$\bar{x} = \int x f(x) dx = \int C x e^{-\alpha x^2} dx = 0 \quad (26)$$

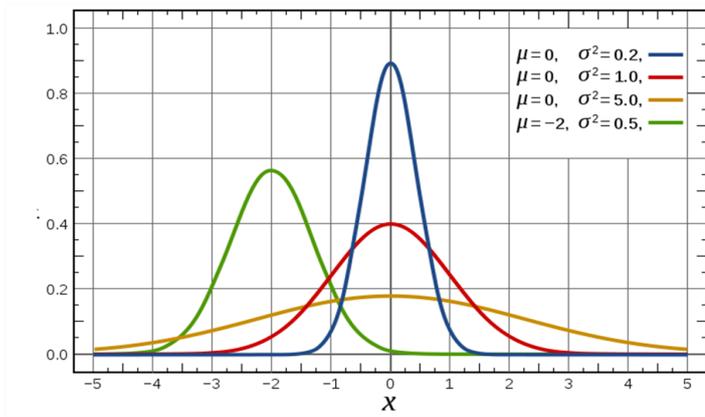
e para a variância, obtemos

$$\sigma = \overline{x^2} = \int x^2 f(x) dx = \int C x^2 e^{-\alpha x^2} dx = \frac{1}{2\alpha} \quad (27)$$

Uma outra forma de escrever a função gaussiana, muito utilizada na descrição de probabilidades é:

$$f(x) = C e^{-\frac{(x-x_0)^2}{2\sigma_0^2}} \quad (28)$$

Verifique que para esta função distribuição o valor médio de x é x_0 , e o valor da variância é σ_0 . Alguns exemplos desta função estão representados na figura abaixo. Algumas curvas estão associadas a valor médio nulo, com diferentes variâncias (ou “larguras”), e uma delas apresenta valor médio negativo (analise a legenda da figura).



O conhecimento desta função e de suas propriedades é muito importante para compreender a matemática da física estatística.