

1. EVENTOS ALEATORIOS Y PROBABILIDAD

1.1 El concepto de probabilidad

Todos los conceptos básicos de la teoría de la probabilidad pueden ser ilustrados mediante un sencillo ejemplo. El ejemplo que veremos a continuación tiene un alto interés para la llamada física estadística. Consideremos la molécula A que se mueve caóticamente dentro de un receptáculo que tiene la forma de una caja (Fig. 1.1)

Un evento aleatorio se define como un fenómeno que en la realización experimental su observación ocurre o no ocurre. Por ejemplo, la posición de la molécula A en un determinado instante se encuentre dentro del volumen $\Delta\tau$ ubicado al interior de la caja (Fig. 1.1). Si la molécula A pudiese ser fotografiada, la fotografía obtenida determinará dos posibles resultados, que se encuentre al interior de $\Delta\tau$, en cuyo caso consideramos que el evento ha ocurrido, o que la molécula se encuentre fuera de $\Delta\tau$, entonces el evento no ha ocurrido. *Un experimento que determina la observación de un evento aleatorio es llamado una prueba.*

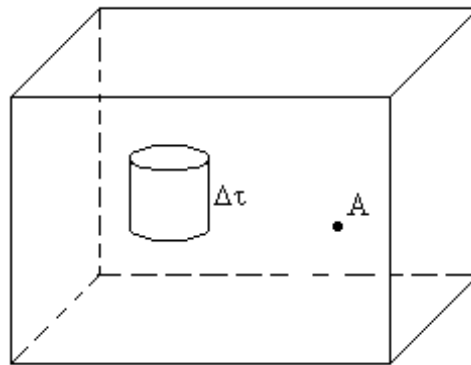


Fig. 1.1

Habitualmente, se entiende por probabilidad de un evento aleatorio como la razón entre el número de pruebas m en que el evento ha ocurrido y el número total M de pruebas, de tal manera que M sea suficientemente grande. Si designamos la probabilidad de que el evento A ocurra por $W(A)$, tenemos que

$$W(A) = \frac{m}{M} \quad \text{o} \quad W(A) = \lim_{M \rightarrow \infty} \frac{m}{M}$$

¿Por qué se impone el requerimiento de que el número de pruebas M sea suficientemente grande? Y ¿cuán grande debe ser? El hecho de que M debe ser grande es obvio conforme a la precisión para determinar el valor de la probabilidad. Supongamos que fotografiamos la molécula en la caja y paramos después de obtener la primera fotografía mostrando a la molécula dentro de la región $\Delta\tau$. Si el número total de fotografías fue de 127, podría ser prematuro concluir que $W(A) = 1/127$. Ahora, si tomamos 60 fotografías

más, podremos ver, por ejemplo, que el evento en que estamos interesados no ocurrió aún cuando se realizaron pruebas adicionales y, consecuentemente, de acuerdo a las nuevas mediciones la probabilidad es de $1/187$. Para obtener un resultado suficientemente preciso, debemos realizar pruebas hasta que la razón m/M difiera una de otra en un valor pequeño determinado por la precisión en que nosotros queremos conocer la probabilidad del evento A . Rigurosamente hablando, puesto que el evento A ocurre aleatoriamente, no podemos excluir de nuestras consideraciones el caso en que la razón m/M obtenida experimentalmente no nos entrega el correcto valor de la probabilidad. Aunque tales casos son posibles en principio, son eventos cuya chance de ocurrencia es extremadamente pequeña, y su probabilidad disminuye cuando aumenta el número de pruebas.

Daremos un interesante ejemplo histórico que ilustra lo que hemos dicho anteriormente. El científico francés Pierre Simón de Laplace, uno de los fundadores de la teoría de la probabilidad, estuvo interesado en los nacimientos de niños y niñas. El material por el estudiado, que incluía un número de ciudades europeas, y así para Francia completa, mostraba que la razón que la razón entre el número m de niños nacidos y el número total de infantes nacidos M estaba cercano a 0.5116 . Tomando en cuenta que el número de niños nacidos m que el había analizado era muy grande, Laplace llegó a la conclusión que la discrepancia que había descubierto no podía deberse a la suerte. Descubrió que el número de infantes nacidos en París incluía a los infantes abandonados, y la población de los suburbios abandonaba principalmente a niñas. Después de excluir los datos relacionados a infantes abandonados, la estadística de los nacidos en París se ajustaron a los conclusiones de otras ciudades europeas.

La importancia de cuántas pruebas son necesarias no es de modo alguno trivial. Como veremos, en el ejemplo con la molécula, las pruebas pueden ejecutarse de dos formas, y que en principio son diferentes.

De la primera forma, fotografiamos la molécula en la caja consecutivamente en diferentes instantes. No resulta difícil ver que que estos instantes deben estar separados en intervalos de tiempo suficientemente grandes. Si una serie de fotografías es tomada a una velocidad muy alta, durante ese tiempo la molécula no logrará moverse a una distancia considerable; al evaluar la razón m/M sobre la base de tal serie, inevitablemente llegaremos a un resultado impropio. Los intervalos de tiempo entre fotografías deben ser, por ejemplo, de tal forma que la molécula tenga tiempo para viajar a cualquier punto de la caja. Desde el punto de vista experimental, el criterio de elección del intervalo de tiempo entre fotografías consiste en que en series repetidas de pruebas a intervalos más grandes que el original deben converger al mismo valor límite m/M .

La segunda forma de ejecutar una prueba es considerar M cajas idénticas, de modo que cada una de ellas contenga una molécula de la misma especie A . En un instante definido, todas las moléculas son fotografiadas simultáneamente, y el valor de la razón m/M se establece analizando cada una de las diferentes fotos de cada caja. *Una colección de sistemas idénticos usados para el estudio de las características de la probabilidad es llamado un **ensemble***. Entonces, podemos usar un ensemble de cajas con moléculas para determinar la probabilidad.

Ambas formas conducen a idénticos resultados si solamente en la primera forma la condición de intervalos de tiempo suficientemente grande es mantenida, y en la segunda

forma todos los sistemas del ensemble deben ser necesariamente idénticos.

Se sigue de la definición de probabilidad que estos valores están confinados entre 0 y 1. Necesariamente, m y M son números positivos y, además, el menor valor posible para m es 0, y el máximo valor posible es el número total de pruebas M . *Un evento que ocurre en cada ensayo y cuya probabilidad es además igual a la unidad se llamado un evento seguro.* Un ejemplo es el evento consistente en la ubicación de la molécula A en cualquier lugar del interior el recipiente. Es natural que en cada fotografía que tomemos siempre la molécula se encontrará en el interior del recipiente. Esto es, para cada prueba siempre la molécula estará en el interior de la caja. En el caso opuesto, cuando un evento no ocurre en cualquier prueba, y además su probabilidad es cero, se denominará **imposible**. Podemos dar un ejemplo trivial: el evento de encontrar la molécula fuera del receptáculo. Puesto que la molécula, físicamente siempre estará en la caja, es imposible encontrarla fuera de ella.

1.2 Conjuntos de eventos mutuamente excluyentes

De gran significado en la construcción y aplicación de la teoría de la probabilidad es el concepto de conjuntos de eventos **mutuamente excluyentes**.

Dos eventos se dicen mutuamente excluyentes si la ocurrencia de uno excluye la posibilidad de ocurrencia del otro.

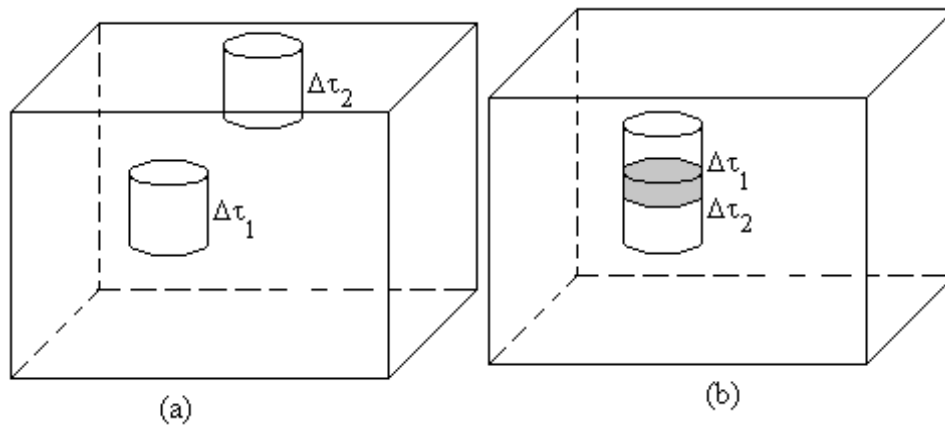


Fig. 1.2

Consideremos un ejemplo. Dos volúmenes, $\Delta\tau_1$ y $\Delta\tau_2$, se seleccionan de la caja para observar si se encuentra en uno de ellos la molécula. Si los volúmenes no se intersectan, como se muestra en la Fig. 1.2a, el evento 1 consistirá en que en el instante t la molécula A se encuentre en $\Delta\tau_1$ y el evento 2 consistirá en que la molécula se encuentre, en el mismo instante t , en $\Delta\tau_2$. En este caso ambos eventos son mutuamente excluyentes. Es evidente que cuando los dos volúmenes se intersectan, como lo indica la Fig. 1.2b, los dos eventos son compatibles puesto que en instante t la molécula puede encontrarse en la región sombreada de la intersección, y esto significa que la molécula puede encontrarse en ambos

volúmenes.

El valor de aplicación del concepto de eventos mutuamente excluyentes se debe al siguiente teorema: *la probabilidad de ocurrencia de uno de dos eventos mutuamente excluyentes es igual a la suma de las probabilidades de ocurrencia de cada uno de ellos*. Este teorema es fácil de probar considerando el ejemplo dado anteriormente. Supongamos que la situación de que la molécula se encuentre en el primero de los dos volúmenes que no se intersectan, está caracterizada por la probabilidad

$$W(1) = \frac{m_1}{M}$$

donde m_1 es el número de veces que la molécula cae dentro del primer volumen en M pruebas. Similarmente, la probabilidad de que la molécula se encuentre en el segundo volumen es

$$W(2) = \frac{m_2}{M}$$

El evento consistente en encontrar la molécula en al menos uno de los dos volúmenes sucede $m_1 + m_2$ veces. Ahora, de acuerdo con la definición general, podemos ver que la probabilidad de que suceda al menos uno de los dos eventos es

$$W = \frac{m_1 + m_2}{M} = \frac{m_1}{M} + \frac{m_2}{M} = W(1) + W(2)$$

y esto prueba el teorema. Con eventos compatibles (intersección no vacía), no podemos asegurar que el número de veces que la molécula se encuentra en el primer o segundo volumen es $m_1 + m_2$. Este número será más pequeño debido a los casos en que la molécula se encuentre en la intersección de los volúmenes, y tales eventos son incluidos simultáneamente en m_1 y m_2 .

Un colección completa de eventos mutuamente excluyentes se define como la colección de eventos mutuamente excluyentes en que la ocurrencia de uno de ellos es segura. Los eventos mutuamente excluyentes considerados en el ejemplo precedente (Fig. 1.2a) no forma un conjunto completo debido a que la molécula puede estar fuera de ambos volúmenes $\Delta\tau_1$ y $\Delta\tau_2$. es decir es una situación posible que no ocurra ninguno de ellos. Ahora, si complementamos los eventos 1 y 2 anteriores con el evento 3 consistente en que la molécula se encuentre en el resto del espacio luego de restar los volúmenes $\Delta\tau_1$ y $\Delta\tau_2$, el nuevo grupo de estos tres eventos forma un conjunto completo. Entonces, como los eventos no se intersectan, ellos son mutuamente excluyentes; y además la ocurrencia de cualquiera de ellos es una certeza, puesto que la molécula se encontrará ya sea en $\Delta\tau_1$, o en $\Delta\tau_2$, o fuera de ambos volúmenes.

Eventos igualmente posibles son aquellos en que la probabilidad de ocurrencia de cualquiera de ellos tiene el mismo valor. Para explicar este concepto, vamos a suponer que la caja está dividida en dos partes iguales, digamos 1 y 2 (Fig. 1.3) y consideremos una colección consistente en dos eventos: obtener la molécula A en 1 y obtener la molécula A en 2. Los volúmenes no se intersectan, y, consecuentemente los eventos son mutuamente excluyentes. La colección es completa porque los volúmenes 1 y 2 completan exhaustivamente todo el volumen de la caja. Finalmente, de la equivalencia completa de los dos volúmenes, es evidente de que la probabilidad de encontrar la molécula en 1 es igual a la probabilidad de encontrarla en 2. En este ejemplo, los eventos 1 y 2 forman una colección

completa mutuamente excluyente e igualmente posible.

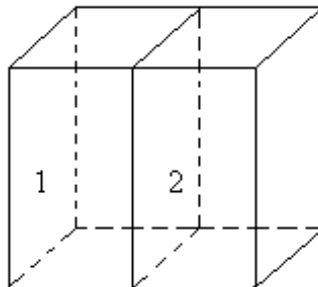


Fig. 1.3

Debemos cuidar que la conclusión de los eventos igualmente posibles esté fundado en nociones bien específicas de la naturaleza y condiciones del movimiento de la molécula. Por ejemplo, si la molécula tiene un momento magnético y un magneto que produce un campo magnético no homogéneo ubicado en la cara izquierda de la caja, la probabilidad de detectar la molécula en el volumen 1 será más alta, y en este caso los eventos no son igualmente posibles.

La Figura 1.4 muestra cinco volúmenes de tamaño idéntico en que una caja ha sido dividida virtualmente. Consideremos los eventos de encontrar la molécula en cada uno de estos volúmenes, ¿son estos volúmenes igualmente posibles?

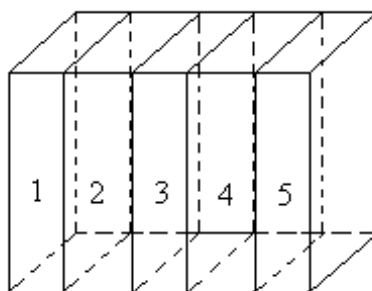


Fig. 1.4

La respuesta dependerá de lo que ocurrirá cuando la molécula colisiona con una pared. Si suponemos que la molécula después de la colisión se adhiere a la pared por algún tiempo, es natural que pasará más tiempo en el volumen 1 y 5. En consecuencia, las probabilidades $W(1)$ y $W(5)$, aunque siendo iguales, serán mayores que las otras probabilidades. Las condiciones del movimiento en cada uno de los volúmenes 2, 3 y 4, sin embargo, son absolutamente idénticos, y además las probabilidades $W(2)$, $W(3)$ y $W(4)$ son iguales unas a otras. En una situación real, lo más frecuente en lo que respecta a colisiones con una muralla es que los impactos se pueden considerar perfectamente elásticos. Las probabilidades de estos cinco eventos serán, entonces, iguales. En lo que sigue, vamos a considerar este

tipo de colisión, de modo que *los sucesos que consisten en detectar una molécula en un cierto instante en volúmenes iguales tendrá la misma posibilidad.*

La importancia del concepto de un conjunto completo de eventos igualmente posibles y mutuamente excluyentes consiste en que nos permite encontrar el valor de la probabilidad de un evento teóricamente. Consideremos el ejemplo de determinar la probabilidad de que la molécula se encuentre dentro del volumen $\Delta\tau$ que está en la caja.

Dividamos virtualmente la caja en n paralelepípedos idénticos (Fig. 5) de manera tal que el volumen $\Delta\tau$, con una alta precisión, esté contenido en un número m de estos paralelepípedos. Como hemos visto anteriormente, los eventos consistente en encontrar la molécula A en uno de estos paralelepípedos, en un instante t , forman un conjunto completo de eventos igualmente posibles y mutuamente excluyentes.

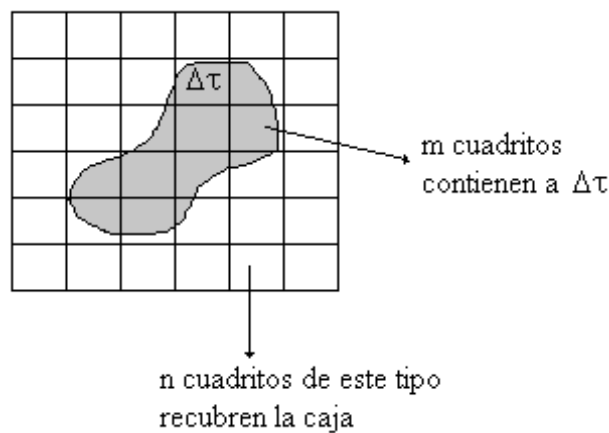


Fig. 1.5

La probabilidad de que la molécula A esté en un determinado paralelepípedo es la misma para todos los paralelepípedos (la condición de igualmente posible) y la podemos calcular de las condiciones de completitud y exclusión mutua. La probabilidad de que la molécula A se encuentre dentro de la caja es 1 (el evento seguro). Por otro lado, de acuerdo a la regla de la suma para eventos mutuamente excluyentes, la misma probabilidad puede ser representada como la suma de las probabilidades de todos los eventos (que son n), es decir $1 = nW$. Se concluye entonces que

$$W = \frac{1}{n}$$

La probabilidad $W(\Delta\tau)$ en la que estamos interesados será igual a la suma de las probabilidades de que la molécula se encuentre en alguno de los paralelepípedos que conforman el volumen $\Delta\tau$. Puesto que la suma total de estos paralelepípedos es m , usando nuevamente el teorema de la adición, tenemos que

$$W(\Delta\tau) = m \times \frac{1}{n} = \frac{m}{n} \tag{1.1}$$

Si multiplicamos el numerador y el denominador por el volumen de uno de los paralelepípe-

dos, obtenemos $\Delta\tau$ en el numerador y el volumen de la caja entera en el denominador. En consecuencia, *la probabilidad de que una molécula en un recipiente de volumen V se encuentre en el elemento $\Delta\tau$ de volumen en el instante t es*

$$W(\Delta\tau) \approx \frac{\Delta\tau}{V} \quad (1.2)$$

La fórmula (1.2) no contiene el número de paralelepípedos n , de modo que el resultado final no depende del tamaño de los volúmenes en el cual hicimos la división virtual del volumen del recipiente. Este tamaño se puede considerar infinitamente pequeño y, consecuentemente, será más exacto formar el volumen $\Delta\tau$ de modo que la fórmula (1.2) es absolutamente precisa, más que la fórmula aproximada de (1.1)

1.3 Eventos independientes

En los cálculos que tenemos que desarrollar con probabilidades, la propiedad de **independencia** de eventos aleatorios es frecuentemente utilizada. *Dos eventos se dicen que son independientes si la ocurrencia de uno de ellos no afecta la probabilidad de ocurrencia del segundo.* Consideremos un ejemplo. Supongamos que el primer evento consiste en que la molécula A está dentro del volumen $\Delta\tau_1$ en el instante t ; y el segundo, que otra molécula B se ubique dentro del volumen $\Delta\tau_2$ en el mismo instante. Sin importar si la molécula A está o no dentro del volumen $\Delta\tau_1$, la probabilidad de que la molécula B se encuentre en el volumen $\Delta\tau_2$ es igual a la cantidad $\Delta\tau_2/V$, luego estos eventos son independientes. Esto ocurrirá si es que las moléculas no interactúan una con la otra. Y en este caso, aún si los volúmenes $\Delta\tau_1$ y $\Delta\tau_2$ coinciden, los eventos son independientes. La presencia de interacción, sin embargo, puede cambiar la situación. Por ejemplo, bajo la mutua repulsión de las moléculas, la probabilidad de que la molécula B aparezca en el volumen $\Delta\tau_1$ junto con la molécula A es más baja que si fuese en la ausencia de A . En condiciones generales, todas las moléculas interactúan una con la otra, pero la fuerza de interacción disminuye rápidamente cuando la distancia entre ellas aumenta entre ellas, de modo que a distancias de orden de varios diámetros de las moléculas la interacción pueden ser despreciada completamente. Esto es posible en bajas concentraciones (gases), cuando la distancia media entre moléculas es mucho más grande que su diámetro, luego si consideramos esta situación, los eventos tratados en el ejemplo anterior son efectivamente independientes. Para grandes concentraciones (gases fuertemente comprimidos, líquidos) la situación es diferente.

La probabilidad de la ocurrencia conjunta de eventos independientes es igual al producto de las probabilidades de cada uno de ellos.

Para explicar esta propiedad, consideraremos un gas enrarecido. Supongamos que hasta n pruebas la molécula A fue encontrada m_1 veces en el volumen $\Delta\tau_1$, y la molécula B fue encontrada m_2 veces en el volumen $\Delta\tau_2$, entonces

$$W(A) = \frac{m_1}{n} \quad y \quad W(B) = \frac{m_2}{n}$$

De todas las pruebas cuyo número total es m_1 en que A cae dentro de $\Delta\tau_1$, dejemos de lado aquellas pruebas en que B cae dentro de $\Delta\tau_2$. Puesto que la probabilidad de que el evento B ocurra es m_2/n , el número de eventos dejados fuera debe ser igual a $m_1(m_2/n)$. Si ahora relacionamos el número encontrado de eventos con el número total de pruebas, la

probabilidad de la ocurrencia conjunta de los eventos A y B es

$$W(AB) = \frac{m_1(m_2/n)}{n} = \frac{m_1}{n} \frac{m_2}{n} = W(A) \cdot W(B)$$

Esta fórmula es la expresión matemática de lo que se asevero anteriormente.

1.4 Probabilidad condicional

La pregunta que asoma ahora es: ¿cómo podemos calcular la probabilidad conjunta de dos sucesos que son dependientes? Para responder a esto volvamos a un ejemplo. Tomemos el evento 1 como la ubicación de la molécula A dentro del volumen $\Delta\tau_1$ en el instante t , y el evento 2 como la ubicación de la misma molécula dentro del volumen $\Delta\tau_2$ en el mismo instante t . Estos eventos son independientes. Supongamos, por ejemplo, que los volúmenes $\Delta\tau_1$ y $\Delta\tau_2$ no se intersectan, es decir que los eventos 1 y 2 son mutuamente excluyentes. Ahora, si la molécula se encuentra en $\Delta\tau_1$, ella no podría estar, naturalmente, en el volumen $\Delta\tau_2$. Luego si la probabilidad del evento 2 es $\Delta\tau_2/V$, la misma probabilidad, cuando el evento 1 ocurre, es ahora 0. Es exactamente el cambio en el valor de la probabilidad del evento 2 debido a la ocurrencia del evento 1 que indica la **dependencia** de estos eventos en el sentido probabilístico.

En el caso más general de intersección arbitraria entre los volúmenes $\Delta\tau_1$ y $\Delta\tau_2$, los eventos 1 y 2 siguen siendo dependientes. para convencernos que esto es cierto, calculemos cuanto cambia la probabilidad del evento 2 dado que ocurre el evento 1. La ocurrencia del evento 1 significa que la molécula está en el interior de $\Delta\tau_1$, y además se puede considerar como un nuevo recipiente que contiene a la molécula. La probabilidad de detectar a la molécula, en el mismo instante, en el volumen $\Delta\tau_2$ es la probabilidad de que la molécula se encuentre en el volumen $\Delta\tau$, que es la parte común de $\Delta\tau_1$ y $\Delta\tau_2$ (vea Fig. 1.2b). Esta probabilidad es igual a la razón entre $\Delta\tau$ y el volumen del "nuevo recipiente" $\Delta\tau_1$. Luego la probabilidad $W(2)$ del evento 2 sin la condición 1, o, de otra forma, la probabilidad incondicional, tiene el valor de

$$W(2) = \frac{\Delta\tau_2}{V}$$

mientras que la *probabilidad condicional* $W_1(2)$ es la probabilidad de que el evento 2 ocurra dado que el evento 1 ha ocurrido: esto es

$$W_1(2) = \frac{\Delta\tau}{\Delta\tau_1}$$

Es obvio que en el caso general las probabilidades condicionales e incondicionales no coinciden en su valor, y esto es justamente lo que significa la dependencia entre dos eventos. Para eventos independientes, las probabilidades condicional e incondicional son iguales.

Podemos ahora formular una regla casi evidente. *La probabilidad $W(1, 2)$ de la ocurrencia conjunta de los dos eventos 1 y 2 es igual a la probabilidad $W(1)$ del evento 1 multiplicado por la probabilidad condicional $W_1(2)$, o la probabilidad $W(2)$ del evento 2 multiplicado por la probabilidad condicional $W_2(1)$.* Esto es

$$W(1, 2) = W(1)W_1(2) = W(2)W_2(1)$$

Tomemos el ejemplo con intersección de ambos volúmenes para ilustrar este regla. La

probabilidad $W(1,2)$ de ocurrencia de ambos eventos 1 y 2 es, en esencia, la probabilidad de que ocurra la intersección de estas dos regiones, es decir que la molécula se encuentre en $\Delta\tau$, de modo que

$$W(1, 2) = \frac{\Delta\tau}{V} \quad (1.3)$$

la fórmula (1.3) pues ser escrita ya sea como

$$W(1, 2) = \frac{\Delta\tau}{V} = \frac{\Delta\tau}{\Delta\tau_1} \frac{\Delta\tau_1}{V} = W(1) \cdot W_1(2)$$

o en la forma

$$W(1, 2) = \frac{\Delta\tau}{V} = \frac{\Delta\tau}{\Delta\tau_2} \frac{\Delta\tau_2}{V} = W(2) \cdot W_2(1)$$

lo cual describe matemáticamente la regla anunciada más arriba.

1.5 Distribución binomial de probabilidades

La información dada anteriormente nos permite resolver muchos problemas de la física estadística. En la presente sección, consideraremos uno de ellos que tiene gran relevancia tanto desde el punto de vista de la aplicación como del desarrollo de la teoría de la probabilidad.

Supongamos que un tubo contiene un gas con un determinado número de moléculas N . Mentalmente fijemos una parte del volumen del vaso $\Delta\tau$ y determinemos la probabilidad de encontrar n moléculas en el volumen $\Delta\tau$. La solución de este problema hará posible, particularmente, conocer la probabilidad de un cierto valor de la densidad del gas, que es de interés en algunas aplicaciones físicas.

Comencemos con dos moléculas ($N = 2$) y asignemos los números 1 y 2 a las moléculas. Para determinar la probabilidad en la que estamos interesados, debemos considerar el siguiente conjunto completo de eventos mutuamente excluyentes:

- 1 las moléculas 1 y 2 están en $\Delta\tau$;
- 2 la molécula 1 está en $\Delta\tau$, y la 2 está fuera de este volumen;
- 3 la molécula 2 está en $\Delta\tau$, y la 1 está fuera de este volumen;
- 4 ambas moléculas están fuera de $\Delta\tau$.

Las probabilidades de estos eventos se evalúan como sigue.

La probabilidad de que la molécula 1 esté en $\Delta\tau$ es $W = \Delta\tau/V$, mientras que la probabilidad de que se encuentre fuera de ella es

$$1 - W = 1 - \frac{\Delta\tau}{V}$$

en virtud de que la suma de las probabilidades de ambos eventos debe ser 1. Las probabilidades relevantes para la molécula 2 tiene los mismos valores. Volvamos ahora al primer evento en que ambas moléculas están en $\Delta\tau$. Puesto que hemos supuesto que la intersección entre ambas moléculas puede ser despreciada, la ubicación de las moléculas 1 y 2 en $\Delta\tau$ son eventos independientes, y además la probabilidad W_1 del primer evento es el

producto de las probabilidades de que las moléculas 1 y 2 estén en $\Delta\tau$, es decir

$$W_1 = \frac{\Delta\tau}{V} \frac{\Delta\tau}{V} = \left(\frac{\Delta\tau}{V}\right)^2$$

Similarmente, podemos encontrar

$$W_2 = \frac{\Delta\tau}{V} \left(1 - \frac{\Delta\tau}{V}\right) = W(1 - W)$$

$$W_3 = \frac{\Delta\tau}{V} \left(1 - \frac{\Delta\tau}{V}\right) = W(1 - W)$$

$$W_4 = \left(1 - \frac{\Delta\tau}{V}\right)^2 = (1 - W)^2$$

Estamos interesados en la probabilidad de que $\Delta\tau$ contenga n moléculas, es decir, para nuestro caso será 2, 1 ó 0. No es difícil verificar que

$$W(2) = W_1 = W^2$$

puesto que el evento 1 sucede solamente si ambas moléculas están en $\Delta\tau$. Además,

$$W(1) = W_2 + W_3 = 2W(1 - W)$$

puesto que en el segundo y tercer caso solo una molécula está en $\Delta\tau$. Finalmente,

$$W(0) = W_4 = (1 - W)^2$$

Si el número de moléculas es tres y ellas son designadas por 1, 2, y 3, los eventos listados en la Tabla 1.1 son posibles. Esa tabla entrega los valores de las probabilidades W_i correspondientes a las diferentes posibles distribuciones de las moléculas, y también entrega las probabilidades requeridas $W(3)$, $W(2)$, $W(1)$ y $W(0)$.

Yendo ahora al caso más general de N moléculas, podemos demostrar que

$$W(n) = \frac{N!}{n!(N-n)!} W^n (1 - W)^{N-n} \quad (1.4)$$

N° even.	mol. en $\Delta\tau$	mol. en $V - \Delta\tau$	N° mol. en $\Delta\tau$	W_i	$W(n)$
1	1, 2, 3	—	3	W^3	$W(3) = W^3$
2	1, 2	3			
3	1, 3	2	2	$W^2(1-W)$	$W(2) = 3W^2(1-W)$
4	2, 3	1			
5	1	2, 3			
6	2	1, 3	1	$W(1-W)^2$	$W(1) = 3W(1-W)^2$
7	3	1, 2			
8	—	1, 2, 3	0	$(1-W)^3$	$W(0) = (1-W)^3$

Tabla 1.1

Necesariamente, la probabilidad de que n moléculas estén en el volumen $\Delta\tau$, y el resto de

las $N - \Delta\tau$ moléculas se encuentren fuera de este volumen es

$$W^n(1 - W)^{N-n} \quad (1.5)$$

El número de eventos similares que difieren unos a otros en cuanto a determinar que n moléculas del total N se encuentren en $\Delta\tau$ es igual al número de combinaciones de N tomadas de n a la vez, esto es

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (1.6)$$

Luego, sumando las probabilidades de todos los eventos en que n moléculas se encuentran en $\Delta\tau$, esto es multiplicando (1.5) por (1.6), nos conduce a la fórmula general dada en (1.4).

Los eventos que indican el número de moléculas que pueden estar en $\Delta\tau$, que pueden ser $0, 1, 2, \dots, N$, conforman una colección mutuamente excluyente y completa. De manera que la suma de las probabilidades de estos eventos deber ser igual a la unidad. Es sencillo convencerse que es necesariamente correcto lo siguiente:

$$\sum_{n=0}^N W(n) = \sum_{n=0}^N \frac{N!}{n!(N-n)!} W^n(1 - W)^{N-n}$$

Usando el teorema del binomio, podemos escribir

$$\sum_{n=0}^N \frac{N!}{n!(N-n)!} W^n(1 - W)^{N-n} = [W + (1 - W)]^N = 1^N = 1$$

La relación entre las probabilidades $W(n)$ y el teorema del binomio explica el porque a este tipo de distribución se le conoce como **distribución binomial de probabilidades**.

Consideremos como ejemplo la distribución de seis moléculas en un recipiente, que lo hemos dividido mentalmente en dos partes iguales, esto es $\Delta\tau = V/2$. Una vez que hayamos calculado los coeficientes binomiales, podemos obtener los resultados que se grafican en la Fig. 1.6, donde los puntos indican los valores correspondientes de las probabilidades. Se puede observar que el evento que tiene la probabilidad mayor cuando n es igual a tres, es decir cuando el recipiente contiene la mitad del número total de moléculas. La ausencia de moléculas ($n = 0$) o cuando están todas en una de las mitades del recipiente ($n = 6$) son considerablemente menos probable y además se observan con menor frecuencia (unas 20 veces menos frecuente).

La distribución binomial es usada para la resolución de variados problemas, y no solamente en el tipo de ejemplo considerado aquí. Supongamos que conocemos la probabilidad de que una molécula tenga una característica definida (por ejemplo, la velocidad que oscila entre 100 y 150 m/s , o que ella viaja sin colisionar con otra molécula en una trayectoria que no excede 1 mm , etcétera). ¿Cuál es la probabilidad de que n moléculas de un total de N tenga la misma característica? Si los correspondientes eventos para las diferentes moléculas son independientes y su probabilidad es W , la respuesta está dada por la distribución binomial (1.4). Para convencernos de lo cierto de esta conclusión, es suficiente tomar en cuenta que el contenido específico del problema del problema dado anteriormente fue usado solamente para encontrar el valor de la probabilidad $W = \Delta\tau/V$ de tener una

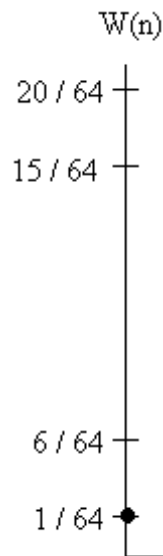


Figure 1: Fig. 1.6

característica definida (en este caso que la molécula esté dentro de $\Delta\tau$ en el instante t). Todas las otras razones no dependen sobre que otra característica podamos considerar.

1.6 La fórmula de Stirling

En la resolución de problemas físicos estadísticos, debemos trabajar con grandes números de moléculas. En consecuencia, el uso de algunas fórmulas, particularmente la distribución binomial, no es conveniente debido a los grandes factoriales de números muy grandes que vamos a encontrar. Por ejemplo, a temperatura y presión estándar, un centímetro cúbico de gas contiene 2.69×10^{19} moléculas. Si queremos determinar la probabilidad de que 10^{16} de ellas se encuentren en un volumen de $\Delta\tau = 1 \text{ mm}^3$, el uso de la fórmula (1.4) requiere que computemos $10^{16}!$ y $2.69 \times 10^{19}!$. Si tomamos en cuenta que ya es difícil el cómputo de $30!$ o de $40!$, es bastante obvio que es deseable tener una fórmula que rápidamente nos permita acceder a factoriales de grandes números. La correspondiente expresión, que encontraremos, fue obtenida por J. Stirling y lleva su nombre.

Es conveniente considerar no el valor de $N!$ sino que con el logaritmo natural de esta expresión

$$\ln N! = \ln 1 + \ln 2 + \cdots + \ln N = \sum_{n=1}^N \ln n$$

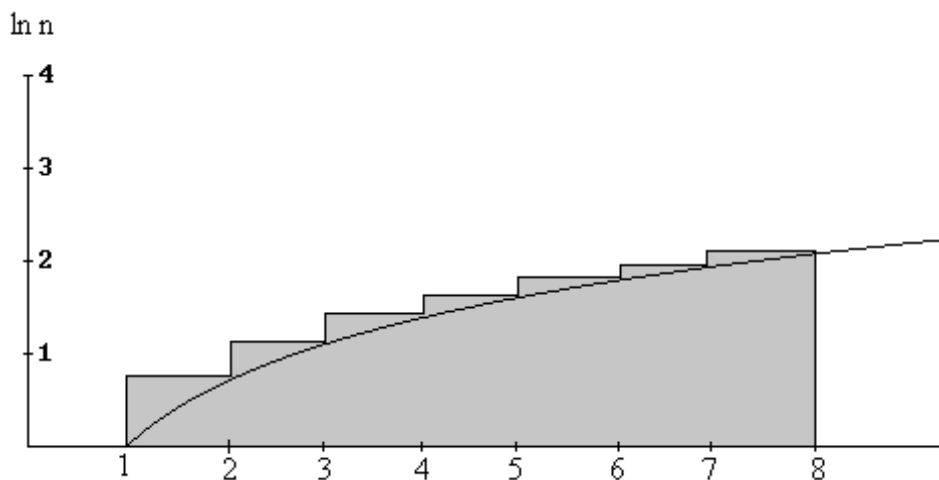


Fig. 1.7

Los valores de $\ln n$ han sido graficados en la Fig. 1.7. Y hemos agregado una suma de rectángulos de tal forma que cada uno de ellos tiene altura precisamente $\ln n$, y de base unitaria, por lo tanto el área de cada rectángulo es $\ln n$. De modo que el área es la figura escalonada está dada por $S = \sum_{n=1}^N \ln n$. Por otro lado, el área bajo la curva determinada por $y = \ln x$ desde $x = 1$ hasta $x = N$ está dada por

$$S' = \int_1^N \ln x \, dx = N \ln N - (N - 1) \quad (1.7)$$

La fórmula (1.7) da un valor aproximado para el área de S , esto es $S' \approx S$, pero S puede ser determinado con más precisión. Del gráfico se desprende claramente que $S' < S$, y una mejor aproximación puede obtenerse si agregamos al área S' el área de los "triángulos" que están sobre la curva. El área de cada triángulo, cuya base es 1 y altura $\ln(n+1) - \ln n$, está dada por

$$\frac{1}{2} [\ln(n+1) - \ln n]$$

Sumando las áreas de todos los triángulos nos queda

$$\frac{1}{2} \ln 2 + \frac{1}{2} [\ln 3 - \ln 2] + \frac{1}{2} [\ln 4 - \ln 3] + \dots + \frac{1}{2} [\ln N - \ln(N-1)] = \frac{1}{2} \ln N$$

Podemos entonces hacer la siguiente aproximación

$$S' \approx N \ln N - N + \frac{1}{2} \ln N + 1$$

Y puesto que $N! = e^S$,

$$N! = \exp \left(N \ln N - N + \frac{1}{2} \ln N + 1 \right) = e N^N e^{-N} \sqrt{N}$$

La estricta derivación, que se puede obtener en un curso de análisis, concluye que un resultado más preciso tiene la forma de

$$N! = \sqrt{2\pi} N^N e^{-N} \sqrt{N} \quad (1.8)$$

es decir, existe una leve diferencia en el valor del factor constante ($e = 2.718\dots$, $\sqrt{2\pi} = 2.506\dots$).

El error relativo de la fórmula 1.8 disminuye cuando N aumenta. Daremos dos ejemplos. Cuando $N = 4$, el valor exacto de $4!$ es 24, mientras que la fórmula de Stirling nos da

$$\sqrt{2\pi}\sqrt{4} \times 4^4 e^{-4} \approx 23.7$$

Cuando $N = 6$, el valor exacto de $6!$ es 720, mientras que si aplicamos la fórmula (1.8) obtenemos el valor de 716. En el primer ejemplo, el error relativo es del 1.25%, y en el segundo, 0.56%. El error relativo disminuye rápidamente cuando aumenta el valor de N .

2. VARIABLES ALEATORIAS Y SUS CARACTERÍSTICAS

2.1 Variable aleatoria

En las aplicaciones de la teoría de probabilidad a problemas de la física estadística, debemos trabajar con la noción de **variable aleatoria**. *Toda variable que toma valores numéricos con una probabilidad definida se dice que es aleatoria*. En el capítulo anterior, en esencia, estábamos trabajando con el concepto de variable aleatoria. En efecto, el número de moléculas en el elemento de volumen $\Delta\tau$ seleccionado de un recipiente de volumen V es el correspondiente ejemplo, puesto que este número toma valores en $0, 1, 2, \dots, N$, y cada uno de estos valores puede ocurrir con una probabilidad $W(n)$ determinada por la fórmula 1.4 de la distribución binomial. El número de moléculas en el volumen elemental $\Delta\tau$ es, entonces, una variable aleatoria. Tal variable aleatoria si dirá **discreta** debido a que toma valores sobre un conjunto finito (en nuestro ejemplo, los valores son $0, 1, 2, \dots, N$).

VARIABLES aleatorias **continuas** también son consideradas. Estas toman valores sobre una serie de valores continuos (un intervalo, por ejemplo). Posiblemente, el ejemplo más simple de variable aleatoria continua es la coordenada de la molécula A en el recipiente de volumen V . Ubiquemos el origen de las coordenadas en una de las esquinas del recipiente rectangular como lo muestra la Fig. 2.1. La coordenada z_A de la molécula a lo largo del eje Z puede tomar valores dentro del intervalo de $z = 0$ hasta $z = a$, donde a es la dimensión del recipiente en la dirección del eje Z .

No es correcto calcular la probabilidad de que una variable aleatoria continua tome un determinado valor prefijado. Lo adecuado es establecer la probabilidad de que una variable aleatoria tome un valor comprendido entre z y $z + dz$. Si el intervalo dz es infinitamente pequeño, la correspondiente probabilidad dW será también infinitamente pequeña, y, consecuentemente, será proporcional al intervalo dz . Podemos escribir la siguiente expresión

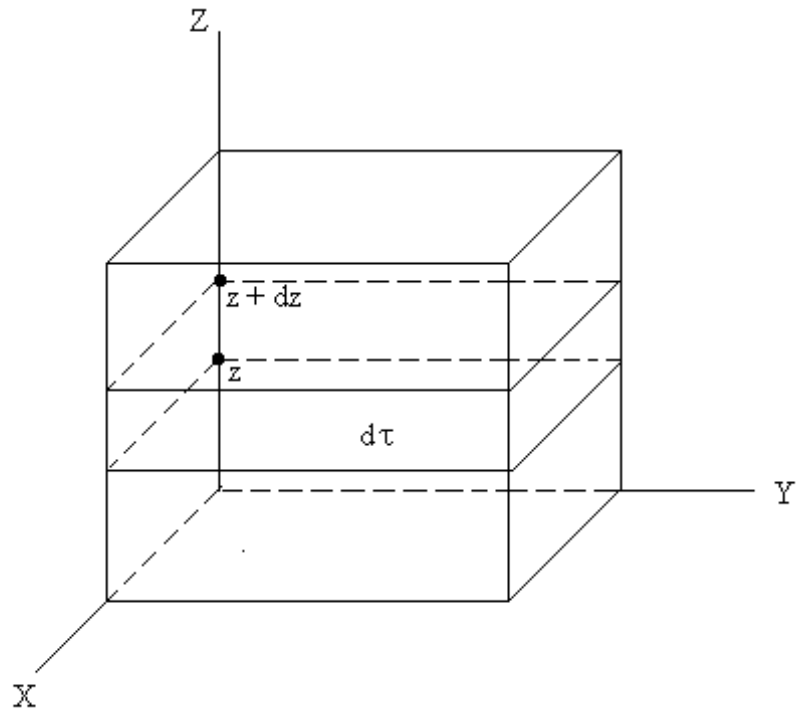


Figure 2: Fig. 2.1

para la probabilidad requerida:

$$dW = w(z) dz \quad (2.1)$$

donde $w(z)$ se conoce como la función de densidad de probabilidad. Su dimensión es inversamente proporcional a la variable aleatoria z puesto que la probabilidad dW es una cantidad adimensional. Entonces, *la función de densidad de probabilidad de una variable aleatoria z es una función $w(z)$ de tal forma que cuando se multiplica por dz , obtenemos la probabilidad de que el valor de la variable aleatoria se encuentre en el intervalo que va desde z hasta $z + dz$* . Si el intervalo dz es igual a cero, la correspondiente probabilidad también es cero, además no tiene sentido en interesarse por la probabilidad de un valor puntual de una variable aleatoria continua.

Para explicar el concepto que hemos introducido, volvamos al ejemplo en que la coordenada z de la molécula A es una variable aleatoria continua. Tenemos que encontrar la densidad de probabilidad para esta variable aleatoria. La coordenada de la molécula caerá dentro del intervalo dz si, como se puede ver en la Fig. 2.1, la molécula misma logra estar dentro del volumen $\Delta\tau$ que está confinado por dos planos paralelos que pasan por los puntos z y dz en un ángulo recto al eje Z . Supongamos que el área de la base del recipiente

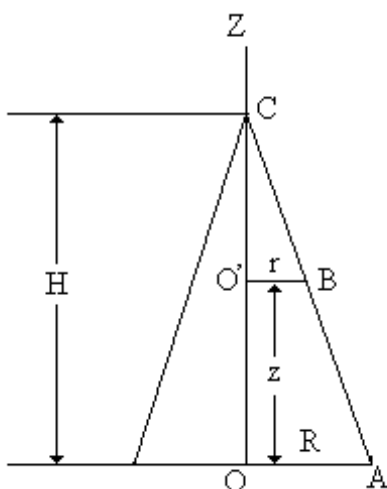


Figure 3: Fig. 2.2

es S , luego el volumen elemental es $d\tau = S dz$. Puesto que el volumen del recipiente completo es $V = a S$, obtenemos la siguiente expresión para la probabilidad:

$$dW = \frac{d\tau}{V} = \frac{S dz}{a S} = \frac{1}{a} dz$$

Comparando esta última igualdad con (2,1) podemos concluir que la densidad de probabilidad para la coordenada de la molécula es

$$w(z) = \frac{1}{a}$$

Como debía de esperarse, la dimensión de la función de densidad de probabilidad es el valor recíproco de la longitud debido a que la variable aleatoria z tiene unidades de longitud.

Que la función de densidad que encontramos sea una cantidad constante está asociado con las características específicas del problema planteado, en que el recipiente es rectangular y que todas las posiciones de la molécula A en el recipiente son igualmente posibles. Si el recipiente fuera ubicado en un campo de acción de fuerzas externas (por ejemplo, si tomamos en cuenta la acción de la fuerza de gravedad), o si la forma del recipiente no fuese rectangular, la función de densidad de probabilidad podría no ser una cantidad constante.

Por ejemplo, consideremos un recipiente de forma cónica (Fig. 2.2), despreciando la acción de la fuerza de gravedad, podemos encontrar la siguiente expresión para la probabilidad dW :

$$dW = \frac{d\tau}{V} = \frac{\pi r^2 dz}{(1/3)\pi R^2 H}$$

donde r es el radio de la sección de un cono circular a una altitud z de la base y, consecuentemente, $\pi r^2 dz$ es el volumen de la sección del cono confinado entre los planos que

pasan perpendiculares al eje Z por los puntos z y $z + dz$ respecto de la base. R es el radio de la base, H es la altura del cono. En virtud de la similaridad de los triángulos COA y COB se sigue que $\frac{r}{R} = \frac{H-z}{H}$, entonces tenemos

$$r^2 = \frac{(H-z)^2}{H^2} R^2$$

Si tomamos en cuenta que el volumen del cono es $V = (1/3) \pi R^2 H$, entonces

$$dW = 3 \frac{(H-z)^2}{H^3} dz$$

es decir, la densidad de probabilidad es

$$w(z) = 3 \frac{(H-z)^2}{H^3}$$

A menudo tendremos que trabajar con vectores de variables aleatorias, es decir con vectores que tengan varias opciones de longitud y dirección que puedan tomar con una probabilidad definida. Un ejemplo muy simple de un vector aleatorio es la posición de la molécula A que puede tomar en el recipiente de volumen V . Construyamos un sistema cartesiano de coordenadas cuyos ejes describirán las componentes x_A , y_A , y z_A aleatorias del vector aleatorio \mathbf{r}_A , y consideremos todos los puntos x, y, z que están contenidos en el volumen rectangular infinitamente pequeño $d\tau = dz dy dz$ (Fig. 2.3). ¿Cuál es la probabilidad de que la punta del vector \mathbf{r}_A esté en el interior del volumen $d\tau$? Puesto que el volumen es infinitamente pequeño, la probabilidad dW debe ser proporcional a tal volumen pequeño, esto es

$$dW = w(x, y, z) d\tau \quad (2.2)$$

La cantidad w , que depende de las coordenadas x, y, z , es llamada la densidad de probabilidad del vector de variables aleatorias \mathbf{r}_A . Entonces, *el producto de la densidad de probabilidad w de un vector de variables aleatorias y un volumen elemental $d\tau$ es la probabilidad de que la posición del vector \mathbf{r}_A esté en el interior de $d\tau$ o, en otras palabras, que las siguientes condiciones se observan simultáneamente: las componentes x, y, z de \mathbf{r}_A están contenidas dentro de los intervalos que van desde x hasta $x + dx$, desde y hasta $y + dy$, desde z hasta $z + dz$, respectivamente.*

En el ejemplo que hemos considerado, la densidad de probabilidad w es el recíproco del volumen del recipiente. En efecto, en la sección 1.2, establecimos que la probabilidad de que la molécula A esté en el volumen elemental $d\tau$, esto es la probabilidad de que la posición superior del vector esté en este elemento, es

$$dW = \frac{d\tau}{V}$$

Comparando esta ecuación con la ecuación (2.2), podemos concluir que

$$w = \frac{1}{V} \quad (2.3)$$

Si las componentes de un vector aleatorio son independientes, en este importante caso particular la densidad de probabilidad $w(x, y, z)$ está representada como el producto de tres

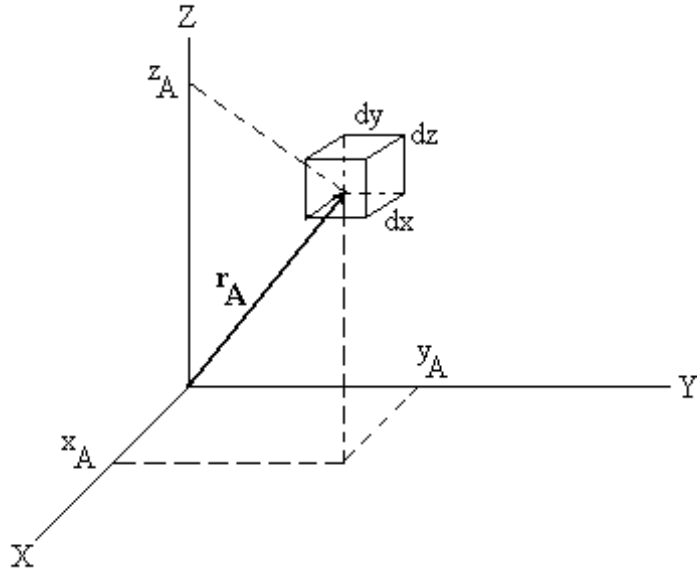


Figure 4: Fig. 2.3

densidades de probabilidad

$$w(x, y, z) d\tau = w_x(x) dx \cdot w_y(y) dy \cdot w_z(z) dz \quad (2.4)$$

de modo que la probabilidad de que la componente x se encuentre en el intervalo dx es $w_x(x) dx$ y no depende de los valores de las otras componentes.

La fórmula (2.4) tiene una sencilla interpretación, significa que la probabilidad de que el vector \mathbf{r}_A esté en $d\tau$ es igual al producto de tres eventos independientes, digamos a la probabilidad de que la componente x se encuentre en el intervalo dx , que la componente y se encuentre en el intervalo dy , y que la componente z se encuentre en el intervalo dz . La fórmula (2.3) obtenida más arriba ilustra este caso. Puesto que el volumen del recipiente tiene la forma de un paralelepípedo rectangular es igual al producto de sus lados, tenemos que

$$V = cba$$

donde c , b y a son las longitudes de los lados a lo largo de los ejes X , Y , y Z , respectivamente. Usando esta relación, podemos escribir la siguiente expresión para la densidad de probabilidad:

$$w = \frac{1}{V} = \frac{1}{c} \times \frac{1}{b} \times \frac{1}{a}$$

entonces

$$w_x = \frac{1}{c}; \quad w_y = \frac{1}{b}; \quad w_z = \frac{1}{a}$$

2.2 Valor medio

El conocimiento de las probabilidades de los valores que puede tomar una variable aleatoria discreta o la densidad de probabilidad de una variable aleatoria continua, nos permite encontrar su valor medio, o, dicho de otra forma, su **esperanza matemática**. Obtendremos la regla para su cálculo mediante un ejemplo específico. Supongamos que tenemos un volumen $\Delta\tau$ dentro del recipiente de volumen V , y la variable aleatoria en la que estamos interesados es el número de moléculas que pueden estar en el interior de $\Delta\tau$ en un cierto instante t . Supongamos además que realizamos un gran número de pruebas M y en cada una de ellas registramos el número de moléculas en $\Delta\tau$. Supongamos que m_1 veces fue registrado el valor n_1 , m_2 veces el valor n_2 , y así sucesivamente. Entonces, encontramos el valor medio del número de moléculas en $\Delta\tau$ mediante la fórmula

$$\begin{aligned}\langle n \rangle &= \frac{m_1 n_1 + m_2 n_2 + \dots}{m_1 + m_2 + \dots} = \frac{m_1 n_1 + m_2 n_2 + \dots}{M} \\ &= \frac{m_1}{M} n_1 + \frac{m_2}{M} n_2 + \dots\end{aligned}$$

Si el número de pruebas es suficientemente grande, la razón m_1/M , m_2/M , etcétera, llegará a ser igual a la probabilidad de los valores relevantes de n , es decir

$$\langle n \rangle = w(n_1)n_1 + w(n_2)n_2 + \dots = \sum_i w(n_i)n_i \quad (2.5)$$

donde la suma es evaluada sobre todos los valores posibles de la variable aleatoria n_i ($n_i = 0, 1, \dots, N$), donde N es el número de moléculas en el recipiente).

Cuando consideremos una variable aleatoria continua z , y puesto que la probabilidad de que su valor se encuentre en un intervalo dz es $w(z) dz$, debemos sumar la expresión $zw(z) dz$ sobre todos los valores de z , es decir sobre todos los intervalos dz debemos encontrar el valor medio. Esto significa que **la regla para determinar el valor medio** de una variable aleatoria continua puede escribirse como

$$\langle z \rangle = \int zw(z) dz \quad (2.6)$$

donde la integral (suma) es evaluada sobre todos los valores posibles de z .

Volvamos al número medio de partículas en el volumen $\Delta\tau$. Supongamos que el número total de partículas en el recipiente de volumen V es N . El valor medio está determinado por la fórmula (2.5) en que $w(n_i)$ debe entenderse como la probabilidad de que n_i partículas estén en $\Delta\tau$, esto es la expresión (1.4). Luego

$$\langle n \rangle = \sum_{n_i=0}^N n_i \frac{N!}{n_i!(N-n_i)!} \left(\frac{\Delta\tau}{V}\right)^{n_i} \left(1 - \frac{\Delta\tau}{V}\right)^{N-n_i}$$

En la evaluación de la suma anterior, debemos prestar atención al hecho de que la suma empieza desde $n_i = 1$ debido a que en $n_i = 0$ el término se anula (debemos recordar que $0! = 1$). Con esto en mente y cancelando n_i , podemos escribir

$$\langle n \rangle = \sum_{n_i=1}^N \frac{N!}{(n_i-1)!(N-n_i)!} \left(\frac{\Delta\tau}{V}\right)^{n_i} \left(1 - \frac{\Delta\tau}{V}\right)^{N-n_i}$$

Si designamos $n_i - 1$ por k , donde k toma los valores $0, 1, \dots, N - 1$, la expresión para $\langle n \rangle$ puede escribirse como

$$\langle n \rangle = \sum_{k=0}^{N-1} \frac{N!}{k!(N-1-k)!} \frac{\Delta\tau}{V} \left(\frac{\Delta\tau}{V} \right)^k \left(1 - \frac{\Delta\tau}{V} \right)^{N-1-k}$$

Removemos el factor común dentro de la suma, y obtenemos

$$\langle n \rangle = N \frac{\Delta\tau}{V} \sum_{k=0}^{N-1} \frac{(N-1)!}{k!(N-1-k)!} \left(\frac{\Delta\tau}{V} \right)^k \left(1 - \frac{\Delta\tau}{V} \right)^{N-1-k}$$

Mediante la fórmula del teorema del binomio y de manera análoga a los cálculos que realizamos en la sección 1.5, tenemos

$$\sum_{k=0}^{N-1} \frac{N!}{k!(N-1-k)!} \frac{\Delta\tau}{V} \left(\frac{\Delta\tau}{V} \right)^k \left(1 - \frac{\Delta\tau}{V} \right)^{N-1-k} = \left[\frac{\Delta\tau}{V} + \left(1 - \frac{\Delta\tau}{V} \right) \right]^{N-1} = 1$$

De modo que,

$$\langle n \rangle = N \frac{\Delta\tau}{V} \quad (2.7)$$

es decir, el número medio de partículas en el volumen $\Delta\tau$ es igual al número de partículas N del recipiente multiplicado por la probabilidad $(\Delta\tau/V)$ de que una partícula se encuentre en $\Delta\tau$.

Ahora calcularemos el valor medio de una variable aleatoria continua usando el ejemplo de la coordenada z de una molécula en un recipiente rectangular. Habíamos visto que la densidad de probabilidad para este caso es $w(z) = 1/a$. El valor medio de z es calculado mediante (2.6):

$$\langle z \rangle = \int_0^a z \frac{1}{a} dz = \left[\frac{1}{a} \frac{z^2}{2} \right]_0^a = \frac{a}{2} \quad (2.8)$$

y obtenemos un resultado que fue intuitivamente obvio desde el principio.

A menudo es importante en las aplicaciones conocer el valor medio de la función de una variable aleatoria, por ejemplo, el cuadrado (u otra potencia) del número de partículas en el volumen $\Delta\tau$ o el cuadrado de la coordenada de una molécula z^2 . Aplicando el razonamiento anterior no a la variable aleatoria n_i misma, sino a la función $\psi(n_i)$ de ella, podemos demostrar fácilmente que el valor medio puede ser determinado por la fórmula

$$\langle \psi(n_i) \rangle = \sum_i \psi(n_i) w(n_i) \quad (2.9)$$

El valor medio de la función de una variable aleatoria continua es evaluada con ayuda de la regla cuyo significado es similar

$$\langle \psi(z) \rangle = \int \psi(z) w(z) dz \quad (2.10)$$

la integral es evaluada sobre todos los valores posibles de la variable aleatoria z .

Entonces, *el valor medio (esperanza matemática) de la función de una variable aleatoria es evaluada como la suma del producto de esta función y la probabilidad del valor de su argumento*, es decir mediante la fórmula (2.9) para una variable aleatoria discreta y por

la fórmula (2.10) para una continua.

Usemos la regla anterior para calcular una importante característica conocida como **dispersión**. La dispersión es la esperanza matemática del cuadrado de la desviación de una variable aleatoria respecto de su valor medio, es decir

$$D(n_i) = \langle (n_i - \langle n_i \rangle)^2 \rangle, \quad D(z) = \langle (z - \langle z \rangle)^2 \rangle \quad (2.11)$$

La importancia de esta característica viene del hecho que ella determina el grado de dispersión de una variable aleatoria, es decir de una cierta manera la dispersión es una medida de la aleatoriedad o azar. Si una variable no aleatoria es considerada como aleatoria tomando un mismo valor con probabilidad 1, es claro que la desviación de su valor medio es cero, y, en consecuencia, la dispersión también es cero. Luego, la dispersión de una variable no aleatoria es cero, mientras que para una variable aleatoria ella será más grande, y su amplitud es la dispersión de sus valores.

las fórmulas (2.11) pueden escribirse de manera más conveniente para sus cálculos. Obtendremos más abajo su cálculo para un ejemplo de una variable aleatoria continua puesto que para el caso discreto solo tenemos que sustituir la integral por una suma.

De acuerdo con la definición de la dispersión,

$$\begin{aligned} D(z) &= \int (z - \langle z \rangle)^2 w(z) dz = \int (z^2 - 2z \langle z \rangle + \langle z \rangle^2) w(z) dz \\ &= \int z^2 w(z) dz - \int 2z \langle z \rangle w(z) dz + \int \langle z \rangle^2 w(z) dz \end{aligned}$$

Puesto que z^2 es simplemente un número, la última integral es igual a

$$\int \langle z \rangle^2 w(z) dz = \langle z \rangle^2 \int w(z) dz$$

La condición de normalización para la densidad de probabilidad de una variable aleatoria es

$$\int w(z) dz = 1$$

Esto proviene del hecho de que los valores que toma una variable aleatoria de diferentes intervalos dz forman una colección de eventos completo y mutuamente excluyente. La condición de normalidad significa que la variable aleatoria con probabilidad 1 tomará un valor cualquiera de todos los posibles valores.

La penúltima integral en la expresión para la dispersión adopta la siguiente forma:

$$\int 2z \langle z \rangle w(z) dz = 2 \langle z \rangle \int z w(z) dz = 2 \langle z \rangle \langle z \rangle = 2 \langle z \rangle^2$$

Puesto que la primera integral corresponde al valor medio del cuadrado de la variable puede ser denotada por $\langle z^2 \rangle$, y tomando en cuenta la normalización el resultado final es

$$D(z) = \langle z^2 \rangle - 2 \langle z \rangle^2 + \langle z \rangle^2 = \langle z^2 \rangle - \langle z \rangle^2 \quad (2.12)$$

Una fórmula similar se tiene para una variable aleatoria discreta:

$$D(n_i) = \langle n_i^2 \rangle - \langle n_i \rangle^2$$

Usando el ejemplo de una dispersión, se puede demostrar fácilmente que el valor medio

de una función de variable aleatoria no es igual a la función cuando el valor medio de la variable aleatoria es considerada en el argumento de la función. Necesariamente, en el caso dado, debemos evaluar el valor medio de la función $f(z) = z^2$. Su valor medio es $\langle z^2 \rangle$, y no es igual al cuadrado del valor medio $\langle z \rangle^2$. Si tal igualdad existiese, la dispersión debería ser nula. Veremos que esto no sucede, usando el ejemplo tratado en la presente sección.

Vamos a calcular la **dispersión de la coordenada** de una molécula en un recipiente rectangular. Para usar (2.12), calcularemos $\langle z^2 \rangle$:

$$\langle z^2 \rangle = \int_0^a z^2 \frac{1}{a} dz = \frac{a^2}{3}$$

Puesto que $\langle z \rangle = a/2$, por (2.12) tenemos

$$D(z) = \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12}$$

Las siguientes **propiedades de valor medio y dispersión** son muy importantes para su aplicación.

- (a) Si sumamos dos funciones $f_1(x, y, z)$ y $f_2(x, y, z)$, donde $r = \{x, y, z\}$ es un vector aleatorio, *el valor medio de la suma de sus funciones es igual a la suma de los valores medios.*

Además, cuando r es un vector aleatorio continuo, tenemos que

$$\begin{aligned} \langle [f_1(x, y, z) + f_2(x, y, z)] \rangle &= \int (f_1 + f_2) w(x, y, z) d\tau \\ &= \int f_1 w(x, y, z) d\tau + \int f_2 w(x, y, z) d\tau \\ &= \langle f_1 \rangle + \langle f_2 \rangle \end{aligned}$$

Esta propiedad es similar para el caso de variables aleatorias discretas.

- (b) Si una función de un vector aleatorio \mathbf{r} es multiplicado por una constante numérica a , el valor medio del producto es igual al valor medio de f_1 multiplicado por a . De otra manera, esta propiedad es formulada como sigue: *un factor constante puede ser removido del operador del valor medio.* La prueba es como sigue:

$$\langle a f_1 \rangle = \int a f_1 w d\tau = a \int f_1 w d\tau = a \langle f_1 \rangle$$

(la prueba es similar para el caso discreto)

- (c) Si existen varias funciones (por ejemplo, tres) f_1 , f_2 , y f_3 de las componentes x, y, z del vector aleatorio $r = \{x, y, z\}$, y *si las variables aleatorias x, y, z son independientes, de modo que*

$$w(x, y, z) d\tau = w_x dx \cdot w_y dy \cdot w_z dz$$

el valor medio del producto $f_1(x) \cdot f_2(y) \cdot f_3(z)$ es igual al producto de sus valores medios.

En efecto,

$$\begin{aligned}\langle f_1 \cdot f_2 \cdot f_3 \rangle &= \int \int \int f_1(x) f_2(y) f_3(z) w_x(x) w_y(y) w_z(z) dx dy dz \\ &= \int f_1(x) w_x(x) dx \cdot \int f_2(y) w_y(y) dy \cdot \int f_3(z) w_z(z) dz \\ &= \langle f_1 \rangle \cdot \langle f_2 \rangle \cdot \langle f_3 \rangle\end{aligned}$$

(d) *La dispersión de la suma de varias variables aleatorias independientes es igual a la suma de sus dispersiones.*

Sea x, y, z variables aleatorias independientes; luego, por definición, tenemos

$$D(x + y + z) = \langle (x + y + z)^2 \rangle - \langle x + y + z \rangle^2$$

Obteniendo el cuadrado del primer paréntesis y usando la propiedad (a) en el segundo paréntesis:

$$\begin{aligned}D(x + y + z) &= \langle (x^2 + y^2 + z^2 + 2xy + 2xz + 2yz) \rangle \\ &\quad - (\langle x \rangle + \langle y \rangle + \langle z \rangle)^2\end{aligned}$$

Obteniendo el cuadrado del segundo paréntesis y aplicando la propiedad (a) al primero:

$$\begin{aligned}D(x + y + z) &= (\langle x^2 \rangle + \langle y^2 \rangle + \langle z^2 \rangle + \langle 2xy \rangle + \langle 2xz \rangle + \langle 2yz \rangle) \\ &\quad - \langle x \rangle^2 - \langle y \rangle^2 - \langle z \rangle^2 \\ &\quad - 2 \langle x \rangle \langle y \rangle - 2 \langle x \rangle \langle z \rangle - 2 \langle y \rangle \langle z \rangle\end{aligned}$$

En virtud de la propiedad (b), removemos el factor 2 del operador valor medio de modo que, por ejemplo, $\langle 2xy \rangle = 2 \langle xy \rangle$. Y puesto que las cantidades x, y, z son independientes, de acuerdo a la propiedad (c) tenemos

$$\langle xy \rangle = \langle x \rangle \langle y \rangle; \quad \langle xz \rangle = \langle x \rangle \langle z \rangle; \quad \langle yz \rangle = \langle y \rangle \langle z \rangle$$

y ahora vemos que todos los términos contienen productos de los valores medios de diferentes variables aleatorias cancelándose unos con otros. En consecuencia,

$$\begin{aligned}D(x + y + z) &= \langle x^2 \rangle - \langle x \rangle^2 + \langle y^2 \rangle - \langle y \rangle^2 + \langle z^2 \rangle - \langle z \rangle^2 \\ &= D(x) + D(y) + D(z)\end{aligned}$$

■

Las propiedades anteriores nos permiten a menudo simplificar los cálculos de los valores medios y de dispersión. Nuevamente vamos a considerar el ejemplo del número de partículas en el volumen $\Delta\tau$. Si asociamos a cada molécula la variable aleatoria n_A que toma el valor de 1 cuando la molécula está en $\Delta\tau$ y el valor de 0 cuando está fuera de este volumen, el número de partículas en $\Delta\tau$ puede ser considerado como la suma de estas variables aleatorias para todas las moléculas del gas. En virtud de la propiedad (a), el valor medio del número de partículas es igual a la suma de la esperanza matemática de n_A . Esta última cantidad es calculada por la fórmula general

$$\langle n_A \rangle = 1 \times \frac{\Delta\tau}{V} + 0 \times \left(1 - \frac{\Delta\tau}{V}\right) = \frac{\Delta\tau}{V}$$

Puesto que la esperanza matemática para todas las moléculas es la misma, su suma es evaluada multiplicando la expresión obtenida por el número total de moléculas N . En consecuencia,

$$\langle n \rangle = N \frac{\Delta\tau}{V}$$

La dispersión del número de partículas en $\Delta\tau$ se puede evaluar rápidamente. Primero es necesario encontrar $\langle n_A^2 \rangle$:

$$\langle n_A^2 \rangle = 1^2 \times \frac{\Delta\tau}{V} + 0^2 \times \left(1 - \frac{\Delta\tau}{V}\right) = \frac{\Delta\tau}{V}$$

y entonces la dispersión de esta cantidad es

$$D n_A = \langle n_A^2 \rangle - \langle n_A \rangle^2 = \frac{\Delta\tau}{V} - \left(\frac{\Delta\tau}{V}\right)^2 = \frac{\Delta\tau}{V} \left(1 - \frac{\Delta\tau}{V}\right)$$

La dispersión del número total de partículas en $\Delta\tau$ está determinada con la ayuda de la propiedad (d). Esto es

$$D = N D n_A = N \frac{\Delta\tau}{V} \left(1 - \frac{\Delta\tau}{V}\right)$$

El cálculo directo de esta cantidad es considerablemente más incómodo.

2.3 Distribución de Poisson

En el Capítulo 1, mencionamos que el uso de la distribución binomial para calcular probabilidades, cuando el número de moléculas es muy grande, no era conveniente, y obteníamos la fórmula de Stirling para facilitar la evaluación de factoriales muy grandes. Usando la fórmula de Stirling, podemos simplificar la distribución binomial y reducirla a otra distribución que puede ser más conveniente para los cálculos.

En la presente sección, cuando el número de moléculas N es muy grande, y pequeños valores para el número n de moléculas en $\Delta\tau$ será el principal interés, es decir, asumimos que para grandes valores de n la probabilidad es muy baja de modo que puede ser ignorado. Supongamos que, por ejemplo, que un recipiente contiene N moléculas, donde N es del orden de 10^{19} . Separemos al volumen $\Delta\tau$ es muy pequeño en comparación con el volumen V del recipiente de modo que podemos encontrar $n = 0, 1, 2$ o 3 moléculas en él, y la probabilidad de que el número de moléculas exceda por ejemplo el valor de $n = 10$, es muy baja. Para este propósito, el volumen debe ser tal que el número medio de moléculas en él, dado por $\langle n \rangle = N (\Delta\tau/V)$ debe ser más pequeño que el número límite $n = 10$ que hemos elegido, por ejemplo, $\langle n \rangle = 3$. Naturalmente, el número 3 ha sido elegido arbitrariamente y perfectamente puede ser otro número que sea suficientemente pequeño en comparación con el número total de moléculas N ; sin embargo para valores relativamente grandes es más conveniente usar una aproximación diferente de la distribución binomial que daremos en la próxima sección y que a sabiendas es inadecuada para valores bajos de $\langle n \rangle$ (como $\langle n \rangle \leq 3$)

El valor exacto de la probabilidad según la fórmula de la distribución binomial es

$$W(n) = \frac{N!}{n!(N-n)!} \left(\frac{\Delta\tau}{V}\right)^n \left(1 - \frac{\Delta\tau}{V}\right)^{N-n}$$

Tomando en cuenta el hecho de que N y $N-n$ son números muy grandes, reemplazamos los factoriales usando la aproximación de la fórmula de Stirling

$$W(n) = \frac{N^N e^{-N} N^{1/2} \sqrt{2\pi}}{n!(N-n)^{N-n} e^{-N+n} (N-n)^{1/2} \sqrt{2\pi}} \times \left(\frac{\Delta\tau}{V}\right)^n \left(1 - \frac{\Delta\tau}{V}\right)^{N-n}$$

Dividiendo tanto el numerador como el denominador por $\sqrt{2\pi} \times N^{N-n} N^{1/2} e^N$, nos queda

$$W(n) = \frac{N^n (\Delta\tau/V)^n (1 - \Delta\tau/V)^{N-n}}{n! (1 - n/N)^{N-n} e^N (1 - n/M)^{1/2}} \quad (2.13)$$

El factor en el denominador, esto es

$$(1 - n/N)^{N-n} = \frac{(1 - n/N)^N}{(1 - n/N)^n}$$

es transformado como sigue. Se sabe del análisis que la expresión $(1 - n/N)^N$ cuando $N \rightarrow \infty$ (en términos prácticos, cuando N es suficientemente grande) es igual a e^{-n} . La expresión $(1 - n/N)^n$, cuando n/N es una cantidad muy pequeña, está muy cerca de la unidad. Si tenemos que $n \ll \sqrt{N}$, por ejemplo $n = 0.3\sqrt{N}$, entonces

$$\begin{aligned} \left(1 - \frac{n}{N}\right)^n &= \left(1 - \frac{0.3}{\sqrt{N}}\right)^{0.3\sqrt{N}} = \left[\left(1 - \frac{0.3}{\sqrt{N}}\right)^{\sqrt{N}}\right]^{0.3} \\ &= (e^{-0.3})^{0.3} = e^{-0.09} = 0.91 \end{aligned}$$

De ahí que

$$\left(1 - \frac{n}{N}\right)^{N-n} \approx e^{-n}$$

Naturalmente, con las suposiciones que hemos hecho, el otro factor del denominador de (2.13) también se aproxima a la unidad:

$$\left(1 - \frac{n}{N}\right)^{1/2} \approx 1$$

Para completar nuestros cálculos, nos queda por considerar la expresión encontrada en el numerador de (2.13):

$$\left(1 - \frac{\Delta\tau}{V}\right)^{N-n} = \frac{\left(1 - \frac{N(\Delta\tau/V)}{N}\right)^N}{(1 - \Delta\tau/V)^n} \quad (2.14)$$

Si $n\Delta\tau \ll V$, por analogía al razonamiento anterior, el denominador de (2.14) es igual a la unidad. La cantidad $N(\Delta\tau/V)$ es igual al valor medio del número de moléculas en $\Delta\tau$, y que denotamos por $\langle n \rangle$, es más pequeño que N , y además

$$\left(1 - \frac{\Delta\tau}{V}\right)^N = e^{-\langle n \rangle}$$

Para la probabilidad en la que estamos interesado, hemos encontrado la siguiente expresión:

$$W(n) = \frac{[N (\Delta\tau/V)]^n e^{-\langle n \rangle}}{n! e^{-n} e^n} = \frac{\langle n \rangle^n e^{-\langle n \rangle}}{n!} \quad (2.15)$$

la fórmula (2.15) se conoce como la **distribución de Poisson**. Ella determina la probabilidad de encontrar n moléculas en $\Delta\tau$ (cuando el valor medio del número de moléculas en $\Delta\tau$ es igual a $\langle n \rangle$) y esta distribución es correcta si el número de moléculas en el recipiente es más grande que $\langle n \rangle^2$.

Puesto que la distribución de Poisson es el caso límite de una binomial cuando $\langle n \rangle^2 \ll N$, o, en otras palabras, cuando $\langle n \rangle/N = \Delta\tau/V \rightarrow 0$, la dispersión de una variable aleatoria que se distribuye según una Poisson es obtenida de la fórmula (tanto tanto) por el mismo proceso de límite, es decir, la dispersión es igual al valor medio:

$$D = \langle n \rangle$$

2.4 Distribución Gaussiana

Otra distribución muy útil se puede utilizar no solo cuando el número de moléculas N es suficientemente grande si no que también cuando el número de molécula n en el volumen $\Delta\tau$ lo es, así como el número $N-n$ de moléculas que quedan fuera del volumen $\Delta\tau$. Este caso se realiza a menudo en la práctica puesto que para que ocurra es suficiente que el número medio de moléculas en el elemento

$$\langle n \rangle = \frac{N \Delta\tau}{V} \quad (2.16)$$

sea grande, pero no demasiado cercano al número total de moléculas, en otras palabras, el volumen $\Delta\tau$ no debe ser demasiado pequeño en comparación con V , ni tampoco muy cercano a este volumen. Veremos en lo que sigue que la distribución Gaussiana es suficientemente precisa si $\langle n \rangle$ y $N - \langle n \rangle$ son mayores que 3, y su precisión aumenta cuando $\langle n \rangle$ y $N - \langle n \rangle$ crecen.

La probabilidad de que el volumen $\Delta\tau$ contenga n partículas de modo que el número medio de partícula en $\Delta\tau$ sea $\langle n \rangle$ y el número medio de partículas $N - \langle n \rangle$ fuera de este volumen sean suficientemente grandes está dada por la fórmula conocida como **distribución Gaussiana**,

$$W(n) = \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(n - \langle n \rangle)^2}{2D} \right] \quad (2.17)$$

La distribución Gaussiana (2.17) es usada con más frecuencia cuando el número de moléculas en $\Delta\tau$ es muy grande. En la práctica, virtualmente no conocemos la probabilidad de encontrar n moléculas exactas en el volumen $\Delta\tau$; por lo general estamos interesados en la probabilidad δW de que el número de moléculas esté dentro del intervalo de valores desde n hasta $n + \delta n$. Si la cantidad δn es suficientemente pequeña (mucho más pequeña que \sqrt{D}), la probabilidad calculada por (2.17) para cualquier n del intervalo de longitud δn es virtualmente la misma. Desde esta condición, la probabilidad δW se obtiene multiplicando la probabilidad $W(n)$ por δn , considerando cualquier n arbitrario del intervalo de

longitud δn , esto es

$$\delta W = \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(n - \langle n \rangle)^2}{2D} \right] \delta n \quad (2.18)$$

La fórmula (2.20) nos muestra que cuando se trabaja con un gran número de moléculas la variable n puede considerarse como una variable aleatoria continua, y la probabilidad de que su valor se encuentre entre n y $n + \delta n$ está determinado por la fórmula (2.18), es decir, la densidad de la probabilidad es

$$w(n) = \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(n - \langle n \rangle)^2}{2D} \right]$$

Si es necesario determinar la probabilidad W del hecho de que n esté contenido en el intervalo que va desde n_1 hasta n_2 , y suponiendo que este intervalo no es pequeño en comparación con \sqrt{D} , será necesario tomar en cuenta los cambios en la densidad de probabilidad en este intervalo, y considerando el carácter de variable aleatoria continua que tiene n , tenemos que

$$W = \int_{n_1}^{n_2} w(n) dn = \int_{n_1}^{n_2} \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(n - \langle n \rangle)^2}{2D} \right] dn$$

Exactamente de la misma manera que cuando determinamos el valor medio de la función $f(n)$ de el número de moléculas n , debemos usar la fórmula para el cálculo de los valores medios de variables aleatorias continuas:

$$\langle f(n) \rangle = \int_{-\infty}^{\infty} f(n) \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(n - \langle n \rangle)^2}{2D} \right] dn \quad (2.19)$$

Hacemos notar que la distribución Gaussiana se utiliza muy a menudo y describe el comportamiento de un gran número de variables aleatorias continuas, y no solo en la situación descrita aquí. Existen razones para esto. El punto central es que para un gran número de pruebas la distribución Gaussiana es el límite de una serie completa de distribuciones. Existe un teorema, que debido a su importancia es conocido como el **teorema central del límite de la teoría de la probabilidad**, que establece condiciones generales suficientes para que la distribución Gaussiana sea el límite. A la distribución Gaussiana se le llama también **distribución normal**. La extendida dispersión de la ley normal proporcionó las bases para la ingeniosa observación que los físicos consideran la dispersión a gran escala de la ley normal para que sea un teorema matemático, mientras que los matemáticos consideran que es un hecho experimentalmente establecido. Debe ser tenido en cuenta naturalmente, que la distribución Gaussiana, aunque de uso frecuente, no es la única posible.

Si una variable aleatoria continua odedece la ley normal de distribución, la probabilidad de que los valores de esta variable z se encuentren en el intervalo que va desde z hasta $z + dz$, que llamaremos dW , está determinada por la fórmula

$$dW(z) = \frac{1}{\sqrt{2\pi D}} \exp \left[-\frac{(z - \langle z \rangle)^2}{2D} \right] dz \quad (2.20)$$

donde $\langle z \rangle$ es el valor medio, y D es la dispersión de la variable aleatoria z .

Podemos comparar la aproximación Gaussiana de la fórmula (1) con la fórmula de la distribución binomial, para los valores de $N = 6$ y $\Delta\tau/V = 0.5$, donde se considera que $\langle n \rangle = 3$, y $D = N(\Delta\tau/V)(1 - \Delta\tau/V) = 1.51$. Los valores comparativos se entregan en la siguiente tabla:

Normal	Binomial
0.016488	0.015625
0.086337	0.09375
0.23314	0.234375
0.32465	0.3125
0.23314	0.234375
0.086337	0.09375
0.016488	0.015625

La primera columna fue obtenida de la fórmula

$$W(n) = \frac{1}{\sqrt{2\pi} 1.51} \exp \left[-\frac{(n-3)^2}{21.51} \right]$$

y la segunda columna de

$$W(n) = \frac{6!}{n!(6-n)!} 0.5^n (1-0.5)^{6-n}$$

en los valores $n = 0, 1, 2, 3, 4, 5, 6$. Como podemos observar los valores están muy cercanos. Para valores más grande de $\langle n \rangle$ y $N - \langle n \rangle$ la diferencia entre los resultados es extremadamente insignificante.

2.5 Probabilidad como medida de incertidumbre

La probabilidad de un evento puede ser la medida de su **incertidumbre**. Esto es bastante evidente cualitativamente. Si la probabilidad es baja, un evento raramente sucede, consecuentemente, se puede decir que está en la categoría de los secuesos de incertidumbre. Inversamente, si la probabilidad es cercana a uno, es un evento que ocurre a menudo, y desde este punto de vista se relaciona con los evento esperados comunes.

En la teoría de la información *se conviene tomar como medida de la incertidumbre de un evento el logaritmo natural de su probabilidad W con el signo opuesto: $-\ln W$.*

La medida de incertidumbre elegida de esta manera transmite de manera correcta muchas de nuestras nociones intuitivas de las propiedades que debe poseer. Ciertamente, como la probabilidad es una cantidad menor o igual a 1, entonces su logaritmo tomado con el signo menos será positivo, en otras palabras, la incertidumbre es una cantidad positiva. Si un evento es seguro su probabilidad es 1, y su logaritmo es cero, de modo que su incertidumbre es cero, de modo que la incertidumbre de un evento seguro es cero. Finalmente, si un evento consiste de dos eventos independientes, y como la probabilidad es igual al producto de las probabilidades, la incertidumbre de un evento compuesto es igual a la suma de las incertidumbres de los eventos simples que lo componen.

Consideraremos un ejemplo para ilustrar este concepto. Supongamos que un recipiente-

contiene tres moléculas y que tenemos una porción de volumen del 0.1 del volumen total, esto es $\Delta\tau/V = 0.1$. ¿Cuál es la incertidumbre del evento que este volumen contenga las moléculas? La probabilidad del evento que estamos interesado, de acuerdo a la distribución binomial, es

$$W(3) = \left(\frac{\Delta\tau}{V}\right)^3 = 10^{-3}$$

De esto se concluye que su incertidumbre es

$$-\ln W(3) = 3 \cdot \ln(10) = 6.9078$$

El mismo resultado se podía haber obtenido de una forma diferente. El evento considerado consiste en la ocurrencia de tres eventos independientes, digamos, la obtención de cada una de las tres moléculas en $\Delta\tau$. La incertidumbre de obtener una molécula en el volumen, puesto que la probabilidad de este evento es $W = \Delta\tau/V = 0.1$, es

$$-\ln W = \ln 10 = 2.3$$

La incertidumbre de obtener las tres moléculas es tres veces este producto, y así obtenemos el mismo resultado.

Puesto que la razón $\Delta\tau/V$ es pequeña, la probabilidad de encontrar molécula allí es baja, y será menos sorprendente si las tres moléculas estén fuera de $\Delta\tau$. ¿Cuál es la probabilidad de este evento? Puesto que la probabilidad relevante es

$$W(0) = \left(1 - \frac{\Delta\tau}{V}\right)^3 = 0.9^3$$

y tomando en cuenta que $\Delta\tau \ll 1$, tenemos que la incertidumbre es aproximadamente

$$-3 \ln \left(1 - \frac{\Delta\tau}{V}\right) \approx 3 \frac{\Delta\tau}{V} = 0.3$$

El valor numérico en este caso es mucho menor que el precedente.

Disminución de la incertidumbre puede considerarse como un aumento de la información que se desprende de esta medición.

2.6 Entropía y aleatoriedad

En la sección precedente, introducimos una medida de incertidumbre para un evento aleatorio. Supongamos que estamos considerando un conjunto completo de eventos mutuamente excluyentes. Cada uno de estos eventos tiene su propia incertidumbre, y nos gustaría encontrar una característica que es común para el sistema entero. Tal característica, conocida como la **entropía** o la **entropía de la información**, es introducida como el valor medio de la incertidumbre para el sistema completo, es decir, la incertidumbre de cada evento es multiplicada por la probabilidad W_i de este evento, y todos estos productos se suman. Luego, designando la entropía por S , podemos llegar a una expresión matemática definiendo la cantidad:

$$S = - \sum_{i=1}^N W_i \ln W_i \quad (2.21)$$

donde la suma es desarrollada sobre todos los eventos que forman la colección completa. La

entropía caracteriza el grado de indeterminancia o aleatoriedad que ocurre en una situación dada. para explicar este significado de entropía, volvamos al ejemplo precedente considerado que la molécula A en un recipiente de volumen V puede estar o no estar en el volumen $\Delta\tau$.

Supongamos primeramente que el volumen $\Delta\tau$ es muy pequeño en comparación con V . Es intuitivamente obvio que en este caso la indeterminancia de la situación es muy pequeña porque la molécula con toda seguridad estará fuera de $\Delta\tau$. La entropía refleja esta noción. Necesariamente, la probabilidad de obtener la molécula en $\Delta\tau$ es $W_1 = \Delta\tau/V$, y la probabilidad de no encontrarla en $\Delta\tau$ es $W_2 = 1 - W_1 = (1 - \Delta\tau/V)$. Para la entropía, consecuentemente, tenemos

$$S = -\frac{\Delta\tau}{V} \ln \frac{\Delta\tau}{V} - \left(1 - \frac{\Delta\tau}{V}\right) \ln \left(1 - \frac{\Delta\tau}{V}\right)$$

Si $\Delta\tau/V$ es muy pequeño, no resulta complicado verificar que el valor de S será también muy pequeño. En efecto, basta verificar que la siguiente igualdad es correcta

$$\lim_{x \rightarrow 0^+} x \ln x + (1 - x) \ln(1 - x) = 0$$

No resulta complicado verificar que el resultado es el mismo cuando $\Delta\tau$ está muy cerca de V . En este caso la molécula A ciertamente deberá estar contenida en $\Delta\tau$. La indeterminancia de la situación es pequeña y por lo tanto la entropía será baja.

¿En que caso una gran indeterminancia puede conseguirse? Intentemos responder esta pregunta despues de establecer las condiciones en que la entropía alcanza su valor máximo considerándolo como una función de la probabilidad $W_1 = \Delta\tau/V$. Derivando la entropía S como función de W_1 , tenemos que

$$\begin{aligned} \frac{dS}{dW_1} &= \frac{d}{dW_1} [-W_1 \ln W_1 - (1 - W_1) \ln(1 - W_1)] \\ &= -\ln W_1 - 1 + \ln(1 - W_1) + 1 = 0 \end{aligned}$$

Se sigue de esta ecuación que $W_1 = (1 - W_1)$, esto es

$$W_1 = \frac{\Delta\tau}{V} = \frac{1}{2}$$

Entonces, el estado de mayor indeterminancia es alcanzado cuando $\Delta\tau$ es la mitad del volumen V .

Una medida de la indeterminancia o aleatoriedad de una situación caracterizada por una colección completa de eventos mutuamente excluyentes con probabilidades W_i como la que establece la fórmula (2.21) es llamada la entropía de la información.