Cap. 2

Conceptos básicos de probabilidad

Las consideraciones del capítulo anterior han indicado que los razonamientos sobre probabilidades son de importancia fundamental para la comprensión de sistemas macroscópicos compuestos de muchísimas partículas. Será útil, por consiguiente, revisar las nociones básicas de probabilidad y examinar cómo pueden aplicarse a algunos problemas sencillos pero importantes. Realmente, este estudio será probablemente de valor en temas mucho más amplios que los que nos interesan de modo inmediato. Por ejemplo, los conceptos de probabilidad son indispensables en todos los juegos de azar, en el negocio de seguros (puesto que necesita conocer la probabilidad de que se produzca la muerte o enfermedad entre sus clientes) y en los métodos de información, como el de consulta de la opinión pública. En biología son de profunda importancia en el estudio de genética. En física, son necesarios para el estudio de la desintegración radiactiva, la llegada de rayos cósmicos a la superficie de la tierra o la emisión aleatoria de electrones por el filamento caliente de un tubo de vacío; además, juegan un papel importante en la descripción cuántica de los átomos y moléculas. Y, lo que es más importante para nosotros, forman la base de nuestro estudio completo de los sistemas macroscópicos.

2.1 *Conjuntos estadísticos*

Consideremos un sistema A en el que podemos realizar observaciones o experimentos [1]. En muchos casos el resultado determinado que resulta de la realización de un simple experimento no puede predecirse con certidumbre, por ser intrínsecamente imposible [2] o porque la información disponible acerca del sistema es insuficiente para permitir esta predicción única. Aunque no se puede afirmar nada sobre un experimento aislado, es posible sin embargo, hacer algunas indicaciones significativas sobre los resultados de un gran número de experimentos análogos. Así llegamos a una descripción *estadística* del sistema, es decir, a una descripción expresada en *probabilidades.* Para su obtención seguiremos el siguiente camino:

En lugar de enfocar nuestra atención sobre el sistema aislado

[1] Un acto de observación puede considerarse como un experimento en el que el resultado de la observación constituye el resultado del mismo. Por lo tanto, no necesitamos hacer distinción entre experimentos y observaciones.

[2] Este es el caso, por ejemplo, en mecánica cuántica en donde el resultado de una medida en un sistema microscópico no es predecible con exactitud.

A de interés, contemplemos una reunión (o *conjunto* según la terminología más común) compuesta por un cierto número muy grande $\mathcal{N}$ de sistemas "semejantes". En principio, imaginemos que $\mathcal{N}$ es arbitrariamente grande (es decir, $\mathcal{N} \to \infty$). Los sistemas se supone que son semejantes en el sentido de que cada sistema satisface las mismas condiciones que sabemos satisface el sistema A. Esto significa que cada sistema se supone que ha sido preparado del mismo modo que A y ha sido sometido a los mismos experimentos que éste. Podemos preguntar entonces en qué fracción de casos se verificará un resultado determinado del experimento. Para ser precisos, los dispondremos de modo que podamos enumerar de algún modo conveniente todos los resultados posibles mutuamente excluyentes del experimento. (El número total de tales resultados posibles puede ser finito o infinito). Supongamos entonces que señalamos mediante una r un resultado determinado del experimento y que existen, entre los $\mathcal{N}$ sistemas del conjunto, $\mathcal{N}_r$ sistemas que presentan este resultado. Entonces la fracción

$$P_r \equiv \frac{\mathcal{N}_r}{\mathcal{N}} \text{ (en donde } \mathcal{N} \to \infty) \tag{1}$$

se denomina *probabilidad de ocurrencia del resultado r.* Cuanto mayor sea $\mathcal{N}$, lógicamente una repetición del mismo experimento en el conjunto conducirá con mayor reproducibilidad a la misma razón $\mathcal{N}_r/\mathcal{N}$. La definición (1) resulta así sin ambigüedad en el límite, cuando se hace $\mathcal{N}$ infinitamente grande.

Lo expuesto anteriormente, muestra cómo puede medirse la probabilidad de presentación de cualquier resultado posible de un experimento en un sistema repitiéndolo en un gran número $\mathcal{N}$ de sistemas semejantes [3]. Aunque el resultado del experimento en un sistema solo no puede predecirse, la misión de una teoría *estadística* es entonces la de predecir la *probabilidad* de que se presente cada uno de los resultados posibles del experimento. Las probabilidades predichas pueden compararse entonces con las realmente medidas, experimentando con un conjunto de sistemas semejantes.

Diversos ejemplos nos servirán para aclarar esta descripción estadística en cierto número de casos concretos.

[3] Si el caso que estamos viendo es independiente del tiempo, podríamos igualmente bien repetir el mismo experimento $\mathcal{N}$ veces en sucesión con el sistema particular en consideración. (Teniendo cuidado de iniciar cada vez el experimento con el sistema en la misma condición inicial.)

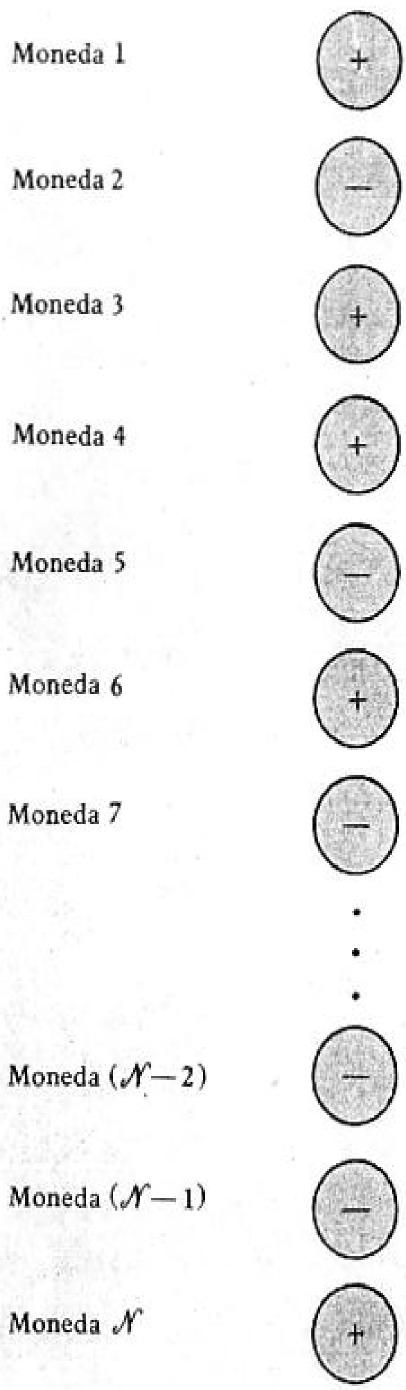


Fig. 2.1 Con objeto de obtener conclusiosiones sobre las probabilidades al lanzar una moneda *aislada*, consideraremos un conjunto compuesto por un número grande $\mathcal{N}$ de monedas semejantes. El diagrama indica esquemáticamente el aspecto de este conjunto después de haber lanzado cada moneda. El símbolo + significa "cara" y el símbolo — "cruz".

Lanzamiento de monedas o dados

Consideremos el experimento de lanzar al aire una moneda. Existen únicamente dos resultados posibles de este experimento o "cara" o "cruz", según que la moneda quede o no sobre una mesa con la efigie grabada hacia arriba. En principio podría predecirse totalmente el resultado del experimento si supiésemos exactamente cómo se arroja la moneda y las fuerzas con las que interacciona en la mesa. Sería necesario entonces únicamente hacer los complicados cálculos necesarios basados en las leyes de la mecánica clásica. Pero en la práctica no se dispone de esta información tan detallada sobre la tirada. No es posible, por lo tanto, hacer una predicción única sobre el resultado de una jugada concreta. (Ciertamente, aunque pudiese obtenerse toda la información precisa y pudiesen realizarse todos los cálculos necesarios, no estaríamos normalmente interesados en obtener estas predicciones tan precisos a costa de complejidad tan grande.) La formulación estadística del experimento es, sin embargo, muy sencilla y familiar. Necesitamos únicamente considerar un conjunto compuesto por un número muy grande $\mathcal{N}$ de monedas semejantes. Cuando se lanzan estas monedas de una manera semejante, podemos contar la fracción de casos en los que el resultado es *cara* y los que el resultado es *cruz*[5]. Estas fracciones dan entonces, respectivamente, la probabilidad medida p de obtener cara y la probabilidad q de obtener cruz. Una teoría estadística debería intentar la predicción de estas probabilidades. Por ejemplo, si el centro de masas de la moneda coincide con su centro geométrico, la teoría podría basarse en el razonamiento de simetría según el cual no hay nada en las leyes de la mecánica que distinga entre caras y cruces. En este caso, la mitad de los resultados experimentales deberían ser caras y la otra mitad cruces, de modo que $p = q = 1/2$. La comparación con el experimento puede comprobar o no esta teoría. Por ejemplo, si se observa que salen caras con mayor frecuencia que cruces, podemos llegar a la conclusión de que la teoría no está justificada al suponer que el centro de masas de la moneda coincide con su centro geométrico.

Consideremos ahora el experimento un poco más complicado de lanzar un grupo de N monedas. Como la tirada correspondiente a cualquier moneda puede tener 2 resultados posibles el lanzamiento de las N monedas puede tener entonces $2 \times 2 \times \ldots \times 2 = 2^N$ resultados posibles[6]. La formulación estadística del experimento requiere nuevamente que, en vez de considerar un simple grupo de N monedas, examinemos un conjunto formado por $\mathcal{N}$ de dichos grupos de N monedas, en los que todos se lanzan de la misma manera. Una cuestión de posible interés es entonces la probabilidad de que se produzca en el conjunto uno cualquiera de los 2^N resultados posibles. Otra cuestión de interés, menos detallada, puede consistir en la probabilidad de hallar en el conjunto un resultado en el que n monedas resulten caras y las restantes $(N - n)$, cruces.

El problema de lanzar un grupo de N dados es, naturalmente, semejante. La única diferencia consiste en que el lanzamiento de un dado cualquiera puede dar 6 resultados posibles, dependientes de cuál de las seis caras del dado cúbico queda hacia arriba.

[4] Despreciamos la remota posibilidad de que la moneda quede en reposo sobre su canto.

[5] De otro modo, podemos tomar la misma moneda y lanzarla $\mathcal{N}$ veces sucesivamente, contando la fracción de casos que presentan cara o cruz.

[6] En el caso especial en que $N = 4$, los $2^4 = 16$ resultados posibles están relacionados explícitamente en la tabla 1.1, pág. 11, si se interpreta la letra I como cara y la letra D como cruz.

Resulta con frecuencia conveniente utilizar la palabra *suceso* para designar el resultado de un experimento o de una observación. Obsérvese que la probabilidad de que se produzca un suceso depende inicialmente de la información disponible sobre el sistema en consideración. Ciertamente, esta información determina la clase de conjunto estadístico que ha de ser estudiado, puesto que este conjunto debe consistir únicamente de sistemas que cumplan todas las condiciones satisfechas por el sistema determinado en consideración.

Ejemplo

Supongamos que estamos interesados en la siguiente cuestión: ¿Cuál es la probabilidad de que una persona viviendo en nuestro país deba hospitalizarse en algún momento entre las edades de 23 y 24 años? Entonces debemos considerar un conjunto compuesto por un gran número de personas y debemos averiguar la fracción de las mismas hospitalizadas en algún instante entre dichas edades. Si imponemos ahora la condición de que la persona debe ser del sexo femenino, la respuesta a nuestra pregunta es distinta, porque ahora debemos examinar un conjunto de *mujeres* y debemos averiguar cuántas de ellas se hospitalizan entre las edades mencionadas. (Realmente, hay una mayor tendencia a que mujeres de estas edades se hospitalicen debido a los alumbramientos.)

Aplicación a sistemas de muchas partículas

Consideremos un sistema macroscópico compuesto por muchas partículas. Por ejemplo, el sistema podría ser un gas ideal de N moléculas, un sistema de N spines, un líquido o un bloque de cobre. En ninguno de estos casos es posible hacer una predicción única sobre el comportamiento de cada una de las partículas en el sistema[7], ni tiene interés. Acudiremos, por tanto, a una descripción estadística del sistema A en consideración. En lugar de considerar el sistema aislado A, examinaremos un conjunto compuesto de un gran número $\mathcal{N}$ de sistemas semejantes al A. Para hacer cualquier afirmación estadística sobre el sistema en el instante t, observemos los $\mathcal{N}$ sistemas en dicho momento. Así podemos determinar la probabilidad $P_r(t)$ de que la observación ofrezca un resultado concreto r en el instante t. El procedimiento puede comprenderse con mayor facilidad imaginando que tomamos una película de cada sistema del conjunto. Nos encontraremos al final con $\mathcal{N}$ películas que contienen los resultados de todas las obser-

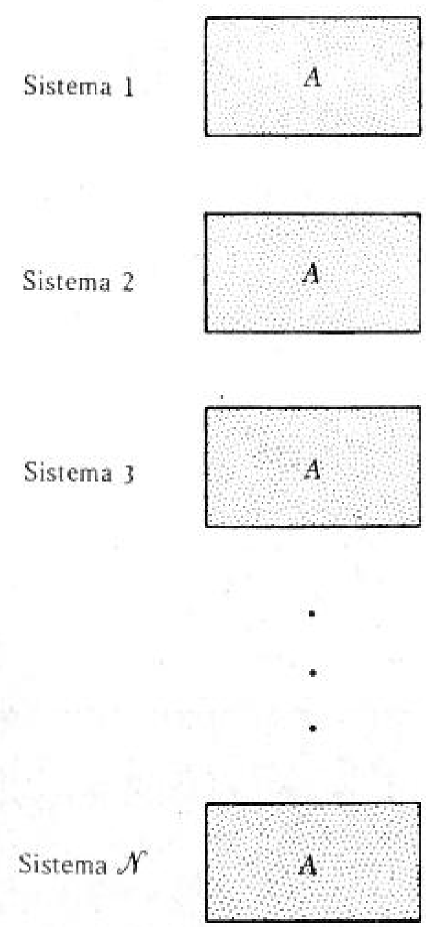


Fig. 2.2 Descripción estadística de un sistema A formado por un gas en una caja. El diagrama representa esquemáticamente un conjunto estadístico de $\mathcal{N}$ sistemas semejantes al sistema A considerado.

[7] En la descripción cuántica correcta del sistema no son posibles, incluso en principio, predicciones no estadísticas. En una descripción clásica, una predicción única sobre un sistema requerirá saber la posición y velocidad de todas las partículas a la vez, información que no disponemos.

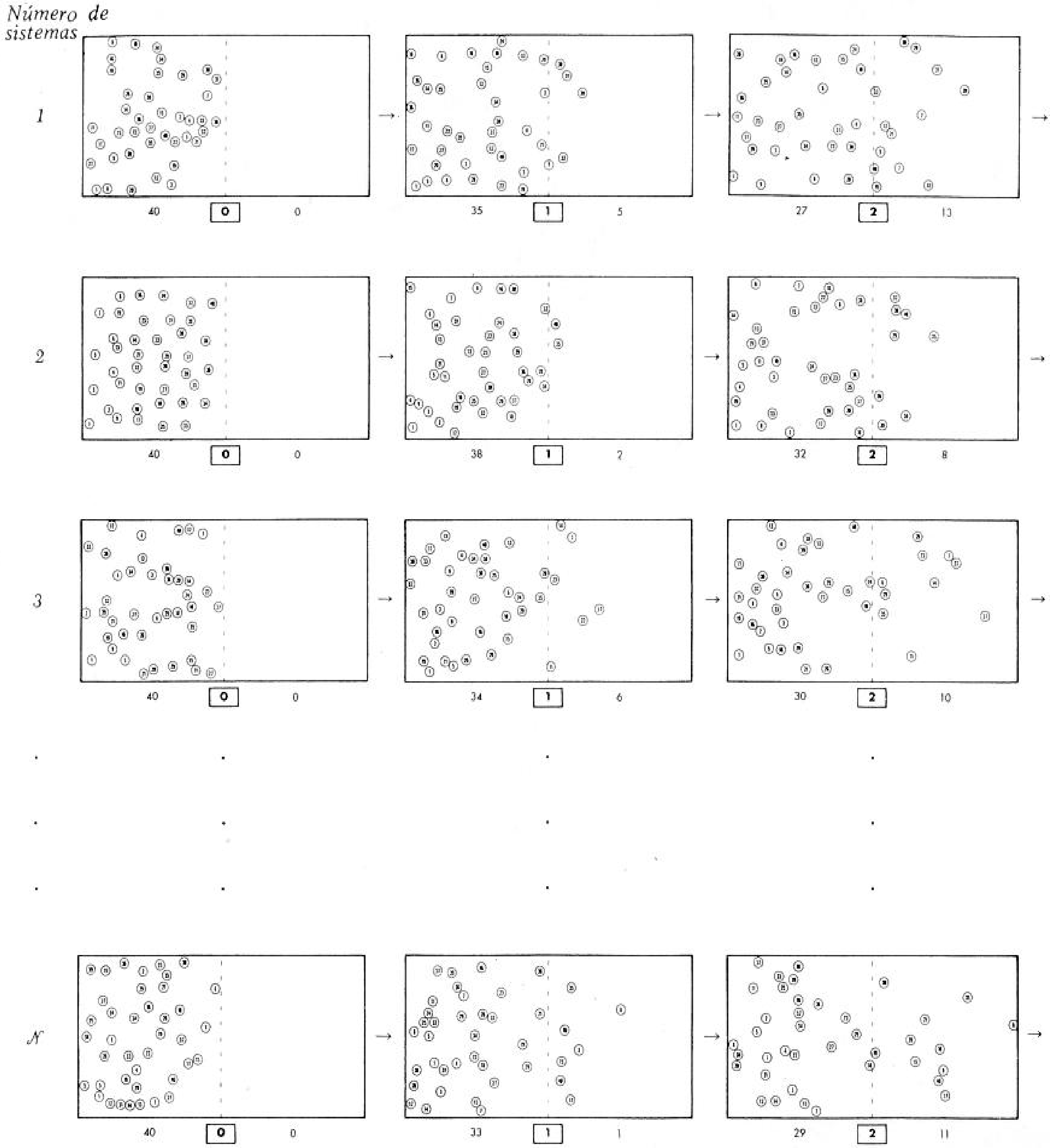


Fig. 2.3 Figuras construidas con ayuda de un ordenador electrónico mostrando un conjunto estadístico de sistemas. Este conjunto se preparó para representar un sistema compuesto por 40 partículas en una caja con la información siguiente: se sabe que todas las partículas están en la mitad izquierda de la caja en un determinado instante inicial correspondiente a la imagen $j = 0$, pero no se sabe nada sobre sus posiciones ni sobre sus velocidades.

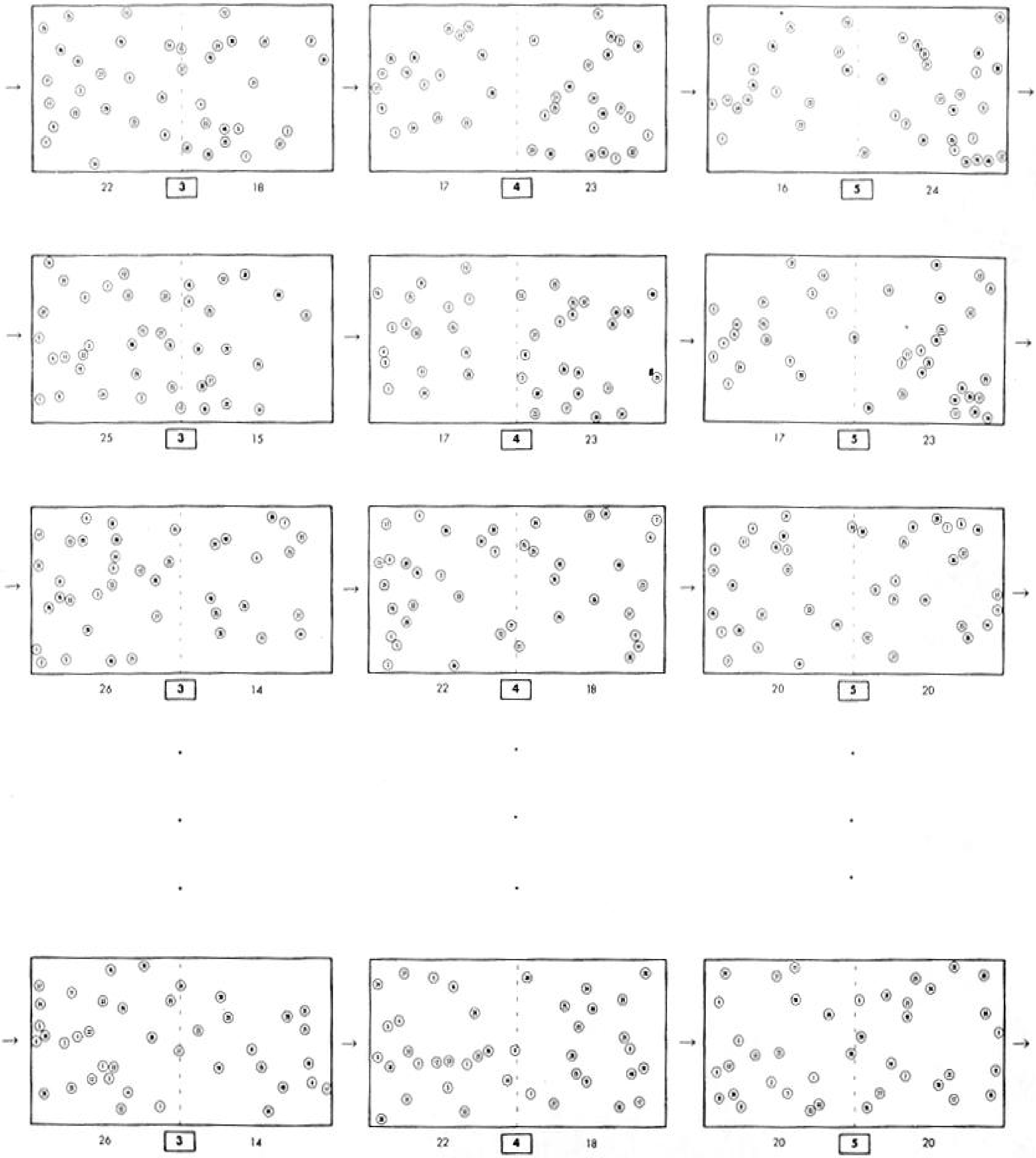


Fig. 2.3 **(cont.)**

La evolución en el tiempo del sistema k del conjunto puede seguirse examinándolo horizontalmente en las figuras sucesivas $j = 0, 1, 2 \ldots$ Pueden hacerse consideraciones estadísticas sobre el sistema en un instante cualquiera correspondiente a la figura j examinando verticalmente todos los sistemas en dicha imagen, y haciendo los recuentos precisos para determinar las probabilidades.

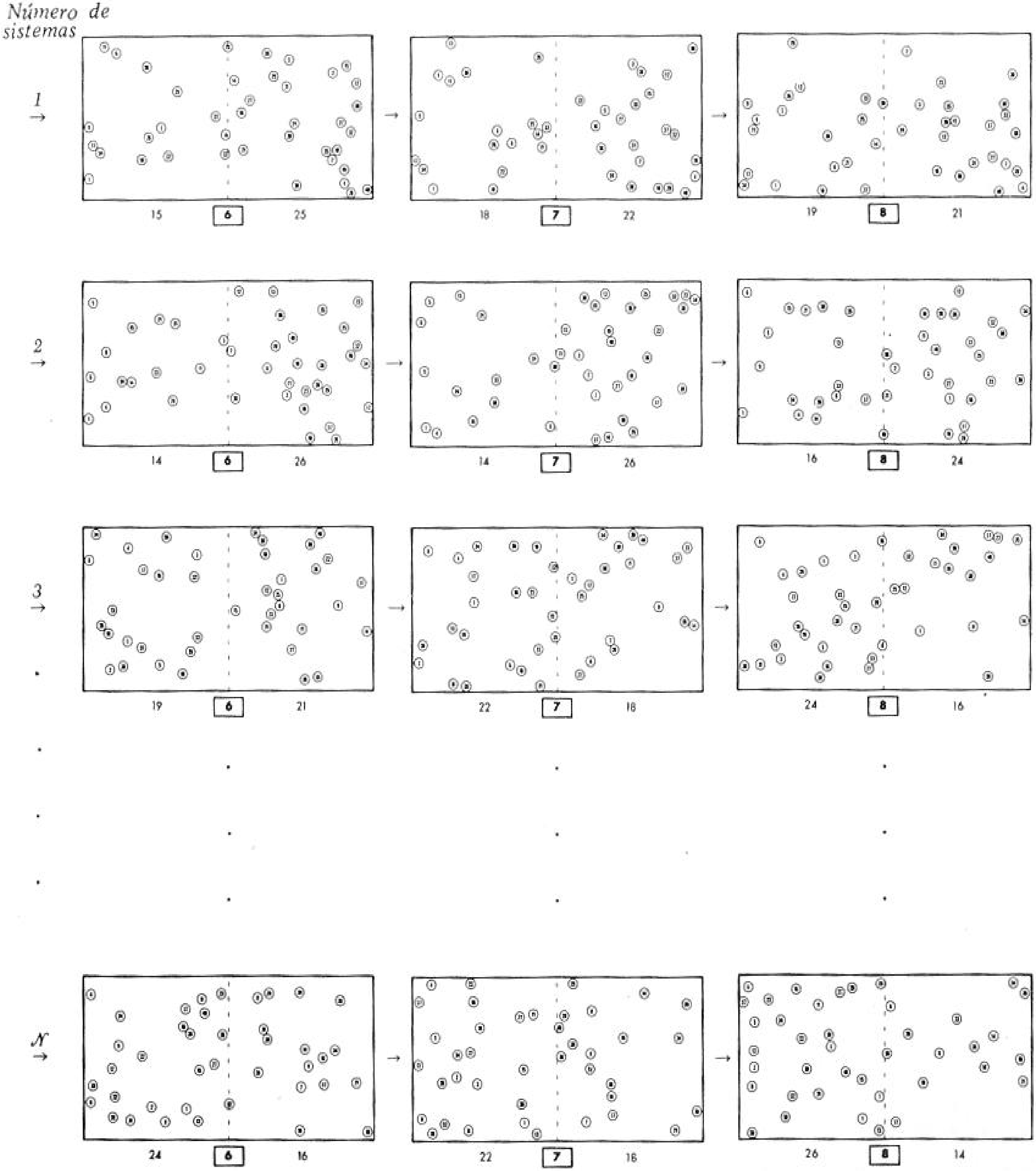


Fig. 2.4 Continuación de la Fig. 2.3. El conjunto ha resultado ser ahora independiente del tiempo, es decir, el sistema ha alcanzado el equilibrio.

vaciones sobre los sistemas del conjunto. El comportamiento de cualquiera de ellos, por ejemplo el sistema k, en función del tiempo puede obtenerse entonces examinando la cinta k (es decir, examinando una línea horizontal de la Fig. 2.3). Por otra parte se obtienen conclusiones de probabilidad sobre el sistema en cualquier instante t examinando todas las imágenes de película tomadas en dicho momento particular t (es decir, mirando a lo largo de una línea vertical en la Fig. 2.3 y contando la fracción de sistemas que presentan un resultado determinado en dicho momento).

Se dice que un conjunto estadístico de sistemas es independiente del tiempo si el número de sistemas que presentan un suceso cualquiera es el mismo en todo momento (o, de modo equivalente, si la probabilidad de que se verifique un suceso particular en este conjunto es independiente del tiempo). La descripción estadística proporciona, pues, una definición de equilibrio muy clara: *Se dice que un sistema macroscópico está en equilibrio si un conjunto estadístico de este sistema es independiente del tiempo.*

Ejemplo

Consideremos un gas ideal de N moléculas. En un cierto instante inicial t_0, inmediatamente después de haber retirado un tabique, se sabe que todas las moléculas de este gas están en la mitad izquierda de una caja. ¿Cómo haremos para dar una descripción estadística de lo que sucede en todos los momentos siguientes? Necesitamos considerar únicamente un conjunto compuesto por un gran número $\mathcal{N}$ de cajas semejantes de gas, teniendo concentradas todas las moléculas en su mitad izquierda en el instante t_0. En la Fig. 2.3 se muestra esquemáticamente un conjunto de éstos. Podemos considerar entonces este conjunto en cualquier momento $t > t_0$ y suscitar varias cuestiones de interés. Por ejemplo, centrando nuestra atención sobre una molécula cualquiera, ¿cuál es la probabilidad $p(t)$ de que esta molécula esté en la mitad izquierda de la caja o la probabilidad $q(t)$ de que esta molécula esté en la mitad derecha? O ¿cuál es la probabilidad $P(n,t)$ de que, en cualquier instante t, estén situadas en la mitad izquierda de la caja n de las N moléculas? En el instante inicial t_0, sabemos que $p(t_0) = 1$ mientras que $q(t_0) = 0$. [Análogamente, $P(N,t_0) = 1$, mientras que $P(n,t_0) = 0$ para $n \neq N$.] Al pasar el tiempo, cambian todas estas probabilidades hasta que las moléculas se distribuyen uniformemente por toda la caja, de modo que $p = q = \frac{1}{2}$. A partir de aquí las probabilidades permanecen invariables a lo largo del tiempo, es decir, el conjunto resulta independiente del tiempo y el sistema ha alcanzado el equilibrio (véase Fig. 2.4)[8]. Esta última situación es, como es lógico, especialmente simple. Ciertamente, el problema del gas

[8] Obsérvese que, a pesar de las fluctuaciones irregulares que ocurren en cualquier sistema *aislado* cuando transcurre el tiempo, la probabilidad P en el *conjunto* tiene en cualquier instante siempre un valor único bien definido, puesto que el número $\mathcal{N}$ de sistemas en el conjunto es arbitrariamente grande. Esta nota aclara la gran simplicidad conceptual que se obtiene al pensar en los conjuntos y no en un sistema solo.

de N moléculas resulta ser entonces análogo al anteriormente estudiado del conjunto de N monedas. En particular, la probabilidad p de hallar una molécula en la mitad izquierda de la caja es análoga a la probabilidad p de que una moneda presente una cara al ser lanzada al aire; de modo semejante, la probabilidad q de hallar una molécula en la mitad izquierda es análoga a la probabilidad q de que la moneda presente cruz. Además como en el caso de la moneda, estas probabilidades son independientes del tiempo y valen $p = q = \frac{1}{2}$.

Nota

En la ecuación (1.4) del Cap. 1 calculábamos probabilidades considerando únicamente un sistema aislado. Este procedimiento es válido ordinariamente en el caso especial de un sistema en equilibrio. Como un conjunto de estos sistemas es independiente del tiempo, un número grande de observaciones sucesivas sobre un sistema único es equivalente, pues, a un gran número de observaciones simultáneas sobre todos los sistemas del conjunto. En otras palabras supongamos que una cinta de película de un solo sistema tomada durante un tiempo τ se corta en $\mathcal{N}$ trozos grandes cada uno de ellos con una duración $\tau_1 \equiv \tau/\mathcal{N}$ (en donde τ_1 es tan grande que el comportamiento del sistema en un trozo es independiente del comportamiento en un trozo adyacente). Entonces la colección de estas $\mathcal{N}$ cintas de un solo sistema no se distingue de una colección de películas $\mathcal{N}$ tomadas de todos los sistemas del conjunto durante cualquier intervalo de tiempo de duración τ_1.

2.2. *Relaciones elementales entre probabilidades*

Las probabilidades satisfacen algunas relaciones sencillas que son casi evidentes por sí mismas, pero muy importantes. Será interesante deducirlas partiendo directamente de la definición (1) de la probabilidad. Debe entenderse a lo largo de todo el estudio siguiente que el número $\mathcal{N}$ de sistemas del conjunto es siempre infinitamente grande.

Supongamos que los experimentos realizados sobre un determinado sistema A pueden conducir a uno cualquiera de α resultados posibles mutuamente excluyentes. Señalemos cada resultado o *suceso* mediante un subíndice r que puede indicar cualquiera de los α números $r = 1, 2, 3, \ldots$ o α. En un conjunto de sistemas semejantes, $\mathcal{N}_1$ de ellos presentarán el suceso 1, $\mathcal{N}_2$ de ellos el suceso 2, ... y $\mathcal{N}_\alpha$ el suceso α. Como estos α sucesos son mutuamente excluyentes y se agotan todas las posibilidades se deduce que

$$\mathcal{N}_1 + \mathcal{N}_2 + \cdots + \mathcal{N}_\alpha = \mathcal{N}$$

Dividiendo por $\mathcal{N}$ esta expresión se reduce a

$$P_1 + P_2 + \cdots + P_\alpha = 1 \tag{2}$$

en donde $P_r \equiv \mathcal{N}_r/\mathcal{N}$ designa la probabilidad de que se presente el suceso r de acuerdo con la definición (1). La relación (2), que establece simplemente que la suma de todas las probabilidades es igual a la unidad, se denomina *condición de normalización* para las probabilidades. Empleando el símbolo sumatorio Σ definido en (M. 1), esta relación puede escribirse también abreviadamente como

$$\sum_{r=1}^{\alpha} P_r = 1 \tag{3}$$

¿Cuál es la probabilidad de que se presente el *suceso* r o el suceso s? Existen $\mathcal{N}_r$ sistemas en el conjunto que presentan el suceso r y $\mathcal{N}_s$ que presentan el suceso s. Por tanto, existen $(\mathcal{N}_r + \mathcal{N}_s)$ que presentan o bien el suceso r o bien el suceso s. En correspondencia, la probabilidad $P(r \text{ o } s)$ de presencia de cualquiera de ambos, el suceso r o el s viene dada simplemente por

$$P(r \text{ o } s) = \frac{\mathcal{N}_r + \mathcal{N}_s}{\mathcal{N}}$$

de modo que

$$\boxed{P(r \text{ o } s) = P_r + P_s} \tag{4}$$

Ejemplo

Supóngase que consideramos el lanzamiento de un dado que, en virtud de su simetría, tiene probabilidades iguales a 1/6 de quedar con cualquiera de sus caras hacia arriba. La probabilidad de que el dado presente el número 1 es entonces 1/6, igual a la probabilidad de que salga el número 2. La probabilidad de que salga el número 1 ó el 2 es pues, según (4), simplemente $1/6 + 1/6 = 1/3$. Este resultado es evidente, naturalmente, puesto que los sucesos en los que el 1 ó el 2 saldrán hacia arriba representa un tercio de todos los casos posibles 1, 2, 3, 4, 5 ó 6.

La relación (4) es generalizable inmediatamente a más de dos sucesos alternativos. Así pues, la probabilidad de que se presenten varios sucesos entre un conjunto de ellos es simplemente igual a la suma de sus probabilidades respectivas. En particular, vemos que la condición de normalización (2) establece sencillamente el resultado evidente de que la suma de probabilidades de la izquierda (es decir, la probabilidad de que se *presente* el suceso 1 *o* el 2, ... *o* el α) es simplemente igual a la unidad (es decir, equivalente a la certeza) puesto que hemos tenido en cuenta todos los posibles sucesos en la enumeración de las α alternativas posibles.

Probabilidades compuestas

Supongamos que el sistema en consideración puede presentar dos tipos diferentes de sucesos, por ejemplo α sucesos posibles del tipo señalado con una r (en donde el índice $r = 1, 2, 3, \ldots, \alpha$) y β sucesos posibles del tipo señalado con una s (en donde el índice $s = 1, 2, 3, \ldots \beta$). Llamaremos P_{rs} la probabilidad de que se presenten conjuntamente *ambos* sucesos, el r y el s. Es decir, en un conjunto compuesto por un gran número $\mathcal{N}$ de sistemas análogos, $\mathcal{N}_{rs}$ de ellos están caracterizados por la presencia conjunta de un suceso r del primer tipo y a la vez un suceso s del segundo. Entonces $P_{rs} \equiv \mathcal{N}_{rs}/\mathcal{N}$. Como es corriente llamaremos P_r a la probabilidad de que se presente un suceso r (independientemente de que se presenten o no sucesos del tipo s). Esto es, si en el conjunto anterior no se presta atención a los sucesos del tipo s y se cuentan $\mathcal{N}_r$ sistemas del conjunto que presentan el suceso r, entonces $P_r \equiv \mathcal{N}_r/\mathcal{N}$. De modo semejante, llamaremos P_s a la probabilidad de que se presente un suceso s (independientemente de la presencia de sucesos de tipo r).

Un caso especial importante es aquel en que la probabilidad de que se verifique un suceso de tipo s no se ve afectada por la presencia o ausencia de un suceso de tipo r. Se dice entonces que los sucesos de tipo r y s son *estadísticamente independientes* o no *correlacionados*. Consideremos ahora en el conjunto los $\mathcal{N}_r$ sistemas que presentan un suceso particular cualquiera r. Independientemente del valor de r, una fracción s de ellos presentará también el suceso s. Así pues, el número $\mathcal{N}_{rs}$ de sistemas que presentan conjuntamente r y s es simplemente

$$\mathcal{N}_{rs} = \mathcal{N}_r P_s$$

En correspondencia, la probabilidad compuesta de que se presenten ambos, r y s, viene dada por

$$P_{rs} \equiv \frac{\mathcal{N}_{rs}}{\mathcal{N}} = \frac{\mathcal{N}_r P_s}{\mathcal{N}} = P_r P_s$$

De aquí podemos sacar la conclusión de que

$$\boxed{\begin{array}{c}\text{si los sucesos } r \text{ y } s \text{ son estadísticamente independientes,}\\ P_{rs} = P_r P_s\end{array}} \tag{5}$$

Obsérvese que el resultado (5) no es cierto si los sucesos r y s no son estadísticamente independientes. La relación (5) puede generalizarse inmediatamente; así pues, la probabilidad compuesta de más de dos sucesos estadísticamente independientes es el producto de sus probabilidades respectivas.

Ejemplo

Supongamos que el sistema A en consideración se compone de dos dados A_1 y A_2. Un suceso de tipo r puede ser la aparición hacia arriba de cualquiera de las 6 caras del dado A_1; análogamente, un suceso de tipo s puede ser lo mismo respecto al dado A_2. Por consiguiente, un suceso del sistema A se especifica indicando las caras del dado A_1 y del dado A_2 que están hacia arriba. Un experimento en el que intervenga el lanzamiento de los dos dados tendría, por tanto, $6 \times 6 = 36$ resultados posibles. Podemos estudiar las probabilidades en un conjunto compuesto de un gran número $\mathcal{N}$ de pares de dados semejantes. Supongamos que todos los dados son simétricos de modo que existe la misma probabilidad de que el resultado de la tirada sea cualquiera de las 6 caras.

La probabilidad P_r de que cada dado caiga con una cara determinada r hacia arriba es entonces simplemente 1/6. Si el dado no interacciona con el otro (es decir, si no están imantados de modo que se ejerzan fuerzas que tiendan a alinearlos) y si se lanzan del mismo modo, pueden considerarse estadísticamente independientes. En este caso la probabilidad compuesta P_{rs} de que el dado A_1 salga con una cara determinada r hacia arriba y el dado A_2 quede con la cara s hacia arriba es sencillamente

$$P_{rs} = P_r P_s = \tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{36}$$

Este resultado es, naturalmente, evidente puesto que el suceso que se observa representa uno de los $6 \times 6 = 36$ resultados posibles.

2.3 *Distribución binómica*

Nos hemos familiarizado ahora suficientemente con los métodos estadísticos para hacer un estudio cuantitativo de algunos problemas físicamente importantes. Consideremos, por ejemplo, un sistema ideal de N spines $\frac{1}{2}$, con un momento magnético asociado μ_0. Este sistema es de particular interés puesto que es un sistema muy simple y se describe con facilidad en función de la mecánica cuántica; por lo tanto, se utilizará con frecuencia como prototipo de sistemas más complicados. Para mayor generalidad, supongamos que el sistema de spines está situado en un campo magnético externo $\mathbf{B}$. Cada momento magnético puede señalar entonces "hacia arriba" (es decir, paralelo al campo $\mathbf{B}$) o "hacia abajo" (es decir, antiparalelo al campo $\mathbf{B}$). Se supone que el sistema de spines está en equilibrio. Un conjunto estadístico compuesto por $\mathcal{N}$ de dichos sistemas es, pues, independiente del tiempo. Enfocando nuestra atención sobre un spin cualquiera, llamemos p a la probabilidad de que su momento magnético señale hacia arriba y q a la probabilidad de que señale hacia abajo. Como estas

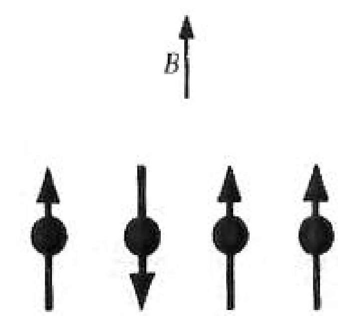


Fig. 2.5 Sistema compuesto por N spines $\frac{1}{2}$ en el caso especial en que $N = 4$. Las flechas indican el sentido del momento magnético del spin. El campo magnético externo está indicado por B.

dos orientaciones agotan todas las posibilidades, el requisito de normalización (3) tiene por consecuencia evidente que

$$p + q = 1 \tag{6}$$

o $q = 1 - p$. En ausencia de campo, cuando $\mathbf{B} = 0$, no existe dirección preferente en el espacio, de modo que $q = p = \frac{1}{2}$. Pero en presencia de un campo, los momentos magnéticos tendrían una mayor probabilidad de señalar en el mismo sentido que el campo, es decir, $p > q$ [9]. Como el sistema de spines es ideal, la interacción entre los spines es casi despreciable, de modo que pueden considerarse sus orientaciones como estadísticamente independientes. La probabilidad de que cualquier momento determinado señale hacia arriba no se ve, por tanto, influida por la orientación de los demás momentos.

Entre los N momentos magnéticos del sistema de spines, llamemos n al número de ellos que señalan hacia arriba y n' al número de los que señalan hacia abajo. Naturalmente

$$n + n' = N \tag{7}$$

de modo que $n' = N - n$. Consideremos entonces los sistemas de spines del conjunto estadístico. El número n de momentos que señalan hacia arriba no es el mismo en cada sistema, sino que puede adquirir cualquiera de los valores posibles $n = 0, 1, 2, \ldots, N$. La cuestión de interés es entonces la siguiente: para cada valor posible de n, ¿cuál es la probabilidad $P(n)$ de que n de los N momentos magnéticos señalen hacia arriba?

El problema de encontrar la probabilidad $P(n)$ se resuelve fácilmente mediante el razonamiento siguiente. La probabilidad de que un momento magnético cualquiera señale hacia arriba es p y de que señale hacia abajo es $q = 1 - p$. Como todos los momentos magnéticos son estadísticamente independientes, la relación general (5) nos permite decir inmediatamente que

$$\left[\begin{array}{l}\text{la probabilidad de que se pre-}\\ \text{sente } \textit{una} \text{ configuración parti-}\\ \text{cular en la que } n \text{ momentos}\\ \text{señalen hacia arriba y los res-}\\ \text{tantes } n' \text{ momentos señalen ha-}\\ \text{cia abajo.}\end{array}\right] = \underbrace{pp\cdots p}_{n \text{ factores}} \quad \underbrace{qq\cdots q}_{n' \text{ factores}} = p^n q^{n'} \tag{8}$$

[9] Consideremos a p y q como cantidades determinadas mediante experimentos. En el Cap. 4 aprenderemos cómo calcular p y q para cualquier valor de B si se sabe que el sistema de spines está a una temperatura dada.

Pero una situación en la que n momentos señalen hacia arriba puede obtenerse normalmente de muchos modos diferentes, como se indica en la Tabla 2.1. Introduzcamos, por tanto, la notación

$C_N(n) \equiv$ número de configuraciones diferentes de N momentos en donde n cualquiera de ellos señalan hacia arriba (y los restantes n' hacia abajo)[10] (9)

La probabilidad deseada $P(n)$ de que n de los N momentos señalen hacia arriba, es igual a la probabilidad de que se verifique la primera, o la segunda, ...o la última de las $C_N(n)$ posibilidades diferentes. De acuerdo con la relación general (4), la probabilidad $P(n)$ se obtiene, pues, sumando la probabilidad (8) para las $C_N(n)$ configuraciones posibles que tienen n momentos señalando hacia arriba, es decir multiplicando la probabilidad (8) por $C_N(n)$. Así se obtiene

$$P(n) = C_N(n)p^n q^{N-n} \tag{10}$$

en donde hemos puesto $n' = N - n$.

Falta sólo calcular el número de configuraciones $C_N(n)$ en el caso general de valores arbitrarios de N y n. Supongamos entonces que se considera una Tabla T, semejante a la Tabla 2.1 en la que se relacionan todas las configuraciones posibles de los N momentos y designemos por U cada momento que señala hacia arriba y por D los que señalen hacia abajo. El número $C_N(n)$ es entonces el número de entradas en las que la letra U aparece n veces. Para examinar cuántas entradas existen, consideremos n momentos diferentes que señalan hacia arriba y marquémoslos con las letras $U_1, U_2, \ldots, U_n$. ¿De cuántas maneras pueden escribirse en una Tabla T' (como, por ejemplo, la Tabla 2.2, en el caso especial en que $N = 4$ y $n = 2$)?

La letra U_1 puede escribirse en una línea de la Tabla en cualquiera de los N lugares diferentes;

para cada situación posible de U_1, puede escribirse entonces la letra U_2 en cualquiera de los $(N - 1)$ lugares restantes;

para cada situación posible de U_1 y U_2, la letra U_3 puede escribirse en cualquiera de los $(N - 2)$ lugares restantes;

... ...

1	2	3	4	n	n'	$C_N(n)$
U	U	U	U	4	0	1
U	U	U		3	1	4
U	U		U	3	1	
U		U	U	3	1	
	U	U	U	3	1	
U	U			2	2	6
U		U		2	2	
U			U	2	2	
	U	U		2	2	
	U		U	2	2	
		U	U	2	2	
U				1	3	4
	U			1	3	
		U		1	3	
			U	1	3	
				0	4	1

Tabla 2.1 Tabla T que relaciona todas las orientaciones posibles de N momentos magnéticos en el caso especial en que $N = 4$. La letra U indica un momento que señala hacia arriba y la D cuando señala hacia abajo. El número de momentos señalando hacia arriba se designa por n y el de los que señalan hacia abajo por n' Las $C_N(n)$ configuraciones posibles en las que n de los N momentos señalan hacia arriba están indicadas en la última columna. (Obsérvese que esta Tabla es análoga a la 1.1.)

[10] El número $C_N(n)$ se suele llamar también número de combinaciones de N objetos de n en n.

1	2	3	4	
U_1	U_2			a
U_1		U_2		b
U_1			U_2	c
U_2	U_1			a
	U_1	U_2		d
	U_1		U_2	e
U_2		U_1		b
	U_2	U_1		d
		U_1	U_2	f
U_2			U_1	c
	U_2		U_1	e
		U_2	U_1	f

Tabla 2.2 Tabla T' en la que se relacionan todas las ordenaciones posibles de $N = 4$ momentos idénticos, de los que $n = 2$ señalan hacia arriba. Para facilitar su enumeración se especifican estos últimos mediante U_1 y U_2, respectivamente, aunque son físicamente indiscernibles. De aquí que las filas que difieren únicamente en los subíndices son equivalentes y están indicadas por letras iguales en la última columna. La Tabla contiene así $n! = 2$ veces demasiadas filas si se está interesado únicamente en filas que sean físicamente distintas.

para cada situación posible de $U_1, U_2, \ldots U_{n-1}$ la letra última U_n puede escribirse entonces en cualquiera de los $(N - n + 1)$ lugares restantes.

El número $J_N(n)$ de posibles entradas o filas diferentes en la Tabla T' se obtiene entonces multiplicando todos los números de situaciones diferentes de las letras $U_1, U_2, \ldots, U_n$; es decir,

$$J_N(n) = N(N-1)(N-2)\cdots(N-n+1) \tag{11}$$

que puede escribirse más concisamente mediante el empleo de factoriales. Así pues,[11]

$$\begin{aligned} J_N(n) &= \frac{N(N-1)(N-2)\cdots(N-n+1)(N-n)\cdots(1)}{(N-n)\cdots(1)} \\ &= \frac{N!}{(N-n)!} \end{aligned} \tag{12}$$

Los símbolos $U_1, U_2, \ldots, U_n$ se consideran diferentes en la enumeración anterior, mientras que realmente los subíndices carecen de importancia puesto que todos los momentos que señalan hacia arriba son equivalentes, es decir cualquier U_i designa un momento que señala hacia arriba, independientemente de i. Las filas en la Tabla T' que difieren únicamente por una permutación de los subíndices, corresponden, por lo tanto, a situaciones físicamente indiscernibles (véase, por ejemplo, la Tabla 2.2). Como el número de posibles permutaciones de los n subíndices viene dado por $n!$, la Tabla T' contiene $n!$ veces demasiadas filas si han de considerarse únicamente filas diferentes no equivalentes[12]. El número deseado $C_N(n)$ de configuraciones distintas de momentos hacia arriba y hacia abajo viene dado entonces dividiendo $J_N(n)$ por $n!$ de modo que

$$C_N(n) = \frac{J_N(n)}{n!} = \frac{N!}{n!(N-n)!} \tag{13}$$

La probabilidad (10) resulta ser entonces

$$\boxed{P(n) = \frac{N!}{n!(N-n)!}p^n q^{N-n}} \tag{14}$$

o en forma más simétrica

[11] Por definición, $N! \equiv N(N-1)(N-2)\ldots(1)$; además, $0! \equiv 1$.

[12] El primer subíndice puede adquirir cualquiera de los n valores posibles, el segundo de los restantes $(n-1)$ valores posibles, ..., y el subíndice n-ésimo el valor restante. De aquí que los subíndices puedan ordenarse en $n(n-1)\ldots(1) \equiv n!$ maneras posibles.

$$P(n) = \frac{N!}{n!n'!} p^n q^{n'} \qquad \text{en que } n' \equiv N - n \tag{15}$$

En el caso especial

en que: $p = q = \frac{1}{2}$
$$P(n) = \frac{N!}{n!n'!} \left(\frac{1}{2}\right)^N \tag{16}$$

Para un número dado N, la probabilidad $P(n)$ es una función de n y se llama *distribución binómica.*

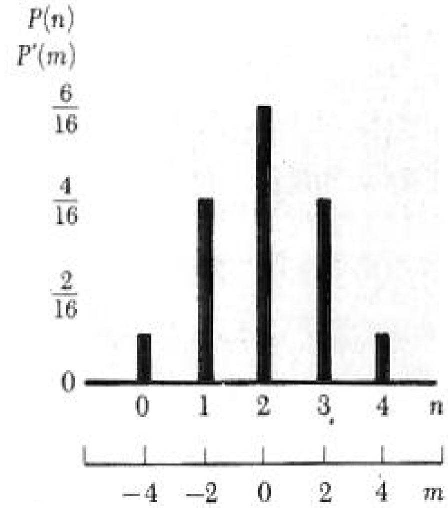


Fig. 2.6 Distribución binómica para $N=4$ momentos magnéticos cuando $p = q = \frac{1}{2}$. El gráfico muestra la probabilidad $P(n)$ de que n momentos señalen hacia arriba o, lo que es equivalente, la probabilidad $P'(m)$ de que el momento magnético en sentido hacia arriba sea igual a m (cuando se mide en unidades de μ_0).

Nota

Al desarrollar un binomio de la forma $(p+q)^N$, el coeficiente del término $p^n q^{N-n}$ es igual simplemente al número $C_N(n)$ de términos posibles que comprenden el factor p exactamente n veces y al factor q $(N - n)$ veces. De aquí se obtiene el resultado puramente matemático conocido como *Teorema del binomio*.

$$(p+q)^N = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n q^{N-n} \tag{17}$$

Comparando con (14) se ve que cada término del segundo miembro es precisamente la probabilidad $P(n)$. Esta es la razón para el nombre de "distribución binómica". Incidentalmente, como $p + q = 1$ cuando p y q son las probabilidades que nos interesan, la ecuación (17) es equivalente a

$$1 = \sum_{n=0}^{N} P(n)$$

Esto comprueba que la suma de las probabilidades para todos los valores posibles de n es igual a la unidad, como exigía la condición de normalización (3).

Discusión

Para examinar la dependencia de $P(n)$ con n, investiguemos primero el comportamiento del coeficiente $C_N(n)$ dado por (13). Observemos primero que $C_N(n)$ es simétrico cuando se cambia n por $N - n = n'$. Así pues,

$$C_N(n') = C_N(n) \tag{18}$$

Además

$$C_N(0) = C_N(N) = 1 \tag{19}$$

Observemos también que

$$\frac{C_N(n+1)}{C_N(n)} = \frac{n!(N-n)!}{(n+1)!(N-n-1)!} = \frac{N-n}{n+1} \tag{20}$$

Partiendo de $n = 0$, la razón de los coeficientes sucesivos es, por consiguiente, grande inicialmente, del orden de N; decrece a continuación monótonamente con n, siendo mayor que (o como máximo igual que) la unidad siempre que $n < \frac{1}{2}N$ y resultando menor que la unidad para $n \geq \frac{1}{2} N$. Este comportamiento, combinado con (19), muestra que $C_N(n)$ tiene un máximo próximo a $n = \frac{1}{2} N$ y que su valor es muy grande comparado con la unidad, en tanto N sea grande.

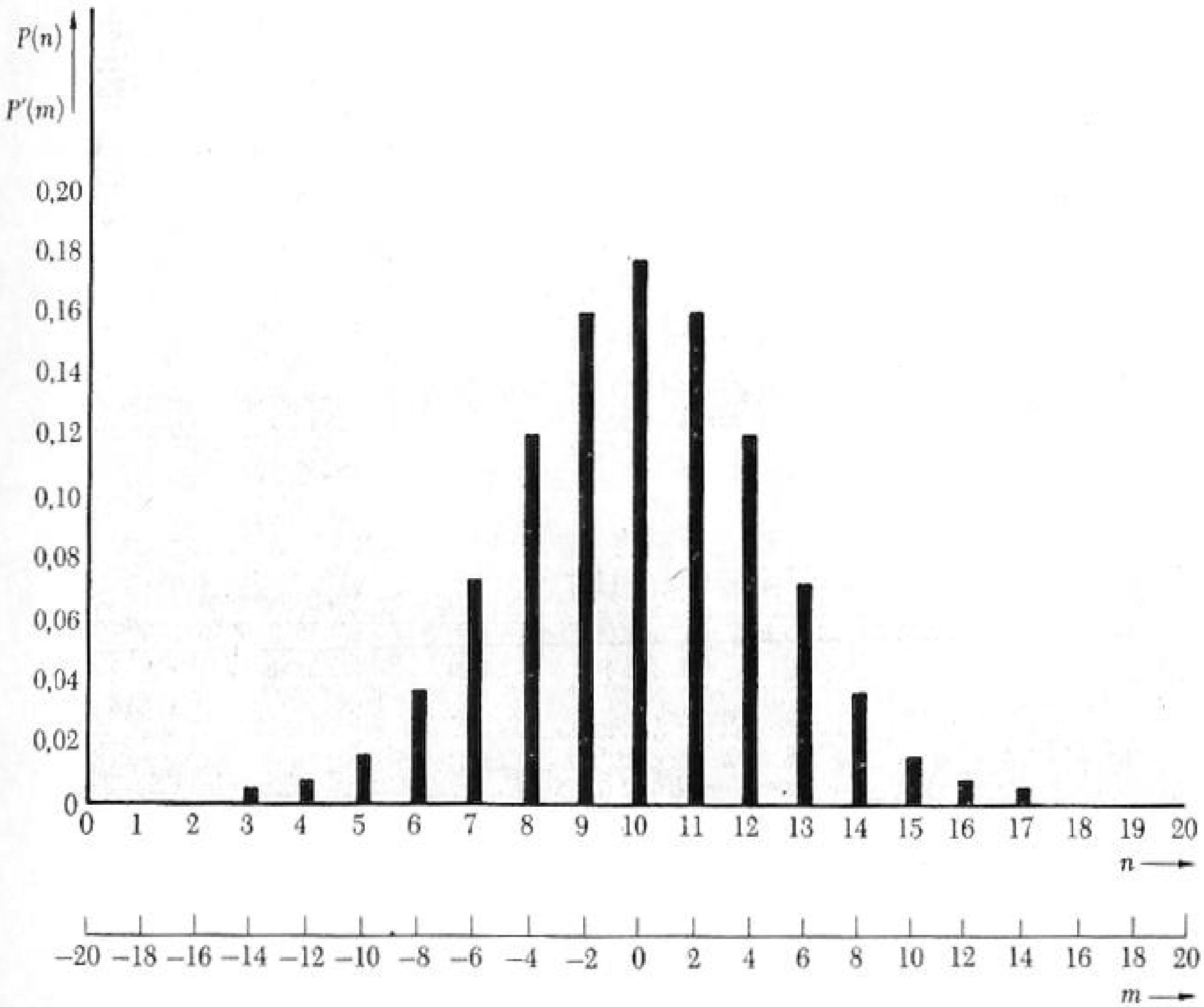
El comportamiento de la probabilidad $P(n)$ resulta entonces claro. Según (16) resulta que

$$\text{si} \qquad p = q = \tfrac{1}{2} \qquad\qquad P(n') = P(n) \tag{21}$$

Este resultado debe, naturalmente, ser cierto por simetría, puesto que no existe ninguna orientación espacial preferida si $p = q$ (es decir, en ausencia de un campo aplicado **B**). En este caso la probabilidad $P(n)$ tiene un máximo[13] cerca de $n = \frac{1}{2} N$. Por otra parte, si $p > q$, el coeficiente $C_N(n)$ tiende todavía a producir un máximo de $P(n)$, pero este máximo se desplaza ahora hacia un valor $n > \frac{1}{2} N$. Las figs. 2.6 y 2.7 muestran el comportamiento de la probabilidad $P(n)$ en algunos casos sencillos.

[13] El máximo está en $\frac{1}{2}N$ cuando N es par y en caso contrario solapa dicho valor.

Fig. 2.7 Distribución binómica para $N = 20$ momentos magnéticos cuando $p = q = \frac{1}{2}$. El gráfico muestra la probabilidad $P(n)$ de que n momentos señalen hacia arriba o lo que es equivalente, la probabilidad $P'(m)$ de que el momento magnético en sentido hacia arriba sea igual a m (cuando se mide en unidades de μ_0).



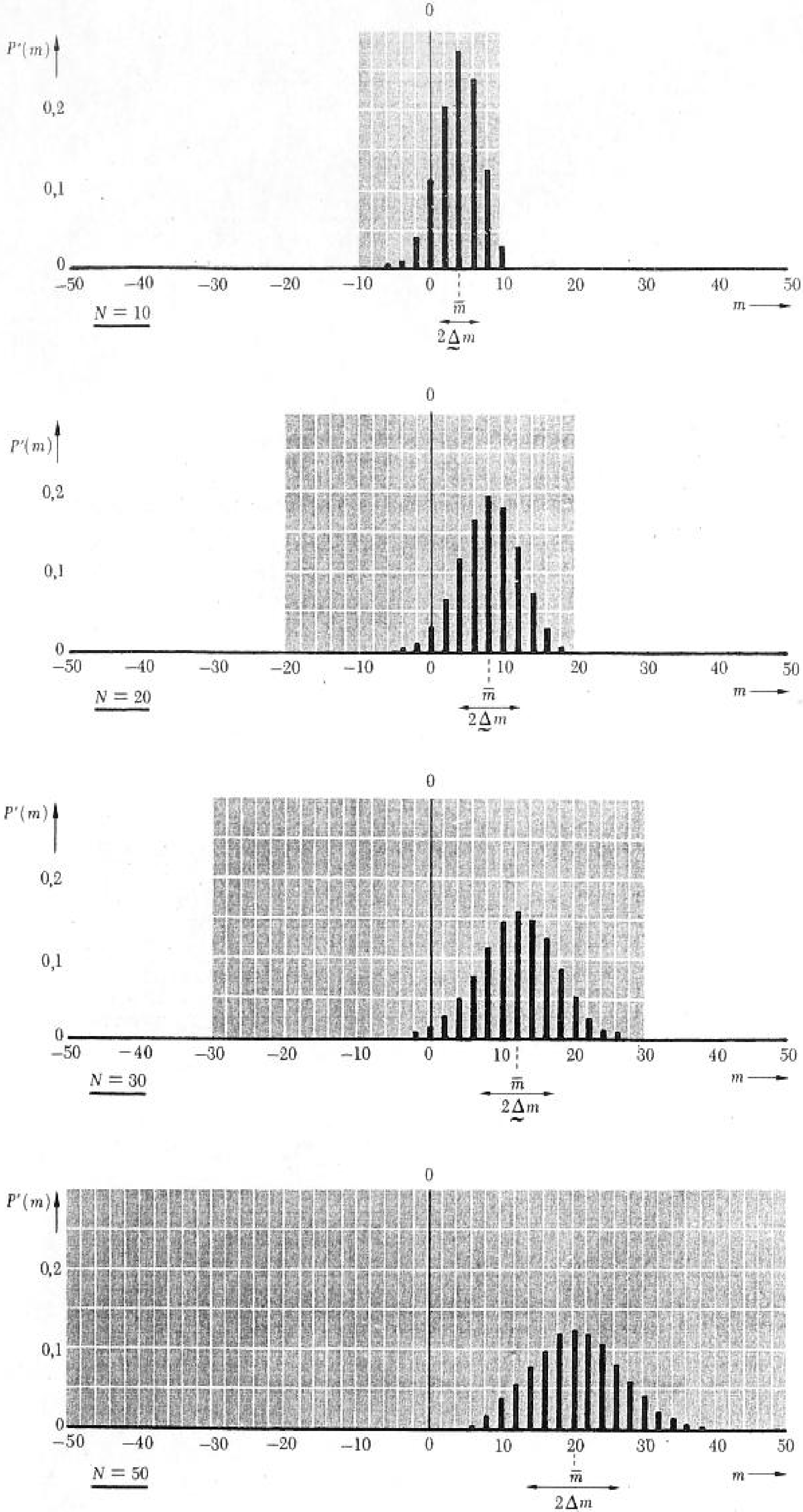


Fig. 2.8 Probabilidad $P(m)$ de que el momento magnético total de un sistema de N spines $\frac{1}{2}$ sea igual a m (medido en unidades de μ_0). Debido a la presencia de un campo magnético aplicado, $p = 0,7$ y $q = 0,3$. Los gráficos muestran a $P'(m)$ en cuatro casos diferentes correspondientes a $N = 10$, $N = 20$, $N = 30$ y $N = 50$.

El momento magnético total de un sistema de spines es una magnitud que se mide experimentalmente con facilidad. Llamemos M al momento magnético total en el sentido "hacia arriba". Como M es simplemente igual a la suma algebraica de los componentes en este sentido de los momentos magnéticos de todos los N spines, resulta que:

$$M = n\mu_0 - n'\mu_0 = (n - n')\mu_0$$

o bien

$$M = m\mu_0 \tag{22}$$

en donde

$$m \equiv n - n' \tag{23}$$

y μ_0 es el módulo del momento magnético de un spin. Según (22), $m = M/\mu_0$ es sencillamente el momento magnético total medido en unidades de μ_0. La relación (23) puede escribirse también como

$$m = n - n' = n - (N - n) = 2n - N \tag{24}$$

Esto muestra incidentalmente que los valores posibles de m deben ser impares si N es impar y pares si N es par. De acuerdo con (24), un valor definido de n corresponde a un valor único de m, e inversamente

$$n = \tfrac{1}{2}(N + m) \tag{25}$$

La probabililad $P'(m)$ de que m adquiera un cierto valor debe ser entonces la misma que la probabilidad $P(n)$ de que n adquiera el valor correspondiente dado por (25). Así pues,

$$P'(m) = P\left(\frac{N + m}{2}\right) \tag{26}$$

Esta expresión da la probabilidad de que se presente cualquier valor posible del momento magnético total del sistema de spin. En el caso especial en que $p = q = \frac{1}{2}$, las expresiones (16) y (26) nos dan explícitamente

$$P'(m) = \frac{N!}{\left(\frac{N + m}{2}\right)!\left(\frac{N - m}{2}\right)!}\left(\frac{1}{2}\right)^{N}$$

la situación más probable tiene lugar cuando $m = 0$ (o próximo a cero), en donde $M = 0$.

Generalidades de la distribución binómica

Aunque nuestro estudio se ha referido al problema específico de un sistema de spines, puede expresarse de un modo más abstracto. Realmente hemos resuelto el siguiente problema general:

dados N sucesos que son estadísticamente independientes, supongamos que cada uno de ellos se verifica con una probabilidad p; la probabilidad de que no se presente viene dada, pues, por $q = 1 - p$. ¿Cuál es entonces la probabilidad $P(n)$ de que se verifiquen n cualquiera de estos N sucesos (mientras que no se presentan los restantes $n' = N - n$ sucesos)? A esta cuestión se responde inmediatamente utilizando la distribución binómica (14). Realmente, en nuestro ejemplo específico de N spines independientes, la presencia de un suceso estaba representada simplemente por un spin señalando hacia arriba, mientras que la falta de un suceso se representaba por un spin no señalando hacia arriba, es decir, señalando hacia abajo.

Otros ejemplos adicionales servirán para resolver algunos problemas comunes que se solucionan inmediatamente en la distribución binómica.

Gas ideal de N moléculas

Consideremos un gas ideal de N moléculas encerrado en una caja de volumen V_0. Como las moléculas de un gas de este tipo interaccionan de un modo prácticamente despreciable, su movimiento es estadísticamente independiente. Supóngase que la caja se subdivide imaginariamente en dos partes de volumen respectivos V y V', siendo

$$V + V' = V_0 \qquad (27)$$

Consideremos un conjunto de muchas cajas de gas. Llamemos p a la probabilidad de que una molécula dada se encuentre en el volumen V y q a la probabilidad de que se halle en el volumen restante V'. Si el gas está en equilibrio, cada molécula tiende a distribuirse uniformemente por toda la caja, de modo que

$$p = \frac{V}{V_0} \qquad \text{y} \qquad q = \frac{V'}{V_0} \qquad (28)$$

Así pues, $p + q = 1$, como en (6). ¿Cuál es entonces la probabilidad $P(n)$ en el conjunto de que n de las N moléculas se hallen en el volumen V (mientras las restantes $n' = N - n$ se encuentran en V')? La respuesta la da la distribución binómica (14). En particular, si $V = V'$ de modo que $p = q = \frac{1}{2}$, hemos resuelto explícitamente el problema de la Sección 1.1, en donde deseábamos hallar la probabilidad de que n de N moléculas se encontrasen en la mitad izquierda de la caja.

Lanzamiento de monedas o dados

Consideremos el lanzamiento de un grupo de N monedas, cuyo comportamiento puede considerarse estadísticamente independiente. Llamemos p a la probabilidad de que una moneda cualquiera salga cara y q la de que salga cruz. Por simetría, podemos suponer que $p = q = \frac{1}{2}$. ¿Cuál es entonces la probabilidad $P(n)$ de que n de las N monedas salgan caras?

El lanzamiento de una serie de N dados es semejante. De nuevo pueden considerarse estadísticamente independientes. Llamemos p a la probabilidad de sacar un "6" y $q = 1 - p$ a la de que no salga. Como un dado tiene seis caras, podemos suponer que $p = \frac{1}{6}$ por simetría, mientras que $q = 1 - p = \frac{5}{6}$. ¿Cuál es entonces la probabilidad $P(n)$ de que n de los N dados muestre un "6"? Esta pregunta se responde una vez más mediante la distribución binómica (14).

2.4 *Valores medios*

Supóngase que una variable u de algún sistema puede tomar uno cualquiera de los α valores posibles

$$u_1, u_2, \ldots, u_\alpha$$

con probabilidades respectivas

$$P_1, P_2, \ldots, P_\alpha$$

Esto significa que, en un conjunto de $\mathcal{N}$ sistemas semejantes (en donde $\mathcal{N} \to \infty$), la variable u toma el valor particular u_r en un número $\mathcal{N}_r = \mathcal{N} P_r$ de dichos sistemas.

La especificación de las probabilidades P_r para todos los α valores posibles de u_r constituye la descripción estadística más completa del sistema. Es conveniente también, sin embargo, definir parámetros que caracterizan de un modo menos detallado la distribución de los posibles valores de u en el conjunto. Estos parámetros son ciertos valores *medios*. Su noción es muy familiar. Por ejemplo, el resultado de un examen realizado a un grupo de alumnos se describe completamente (si no interesa identificar a cada alumno individualmente) especificando el número de alumnos que ha recibido cada calificación entre las posibles que se pueden obtener en el examen. Pero el resultado puede caracterizarse también de un modo menos detallado calculando la nota media de los alumnos. Esto se hace según convenio multiplicando cada nota posible por el número de alumnos que la han recibido, sumando todos los productos así hallados y dividiendo esta suma por el número total de alumnos. De modo semejante, el valor medio de u en el conjunto se define multiplicando u_r por el número $\mathcal{N}_r$ de sistemas para todos los α valores posibles de la variable u y dividiendo entonces esta suma por el número total $\mathcal{N}$ de sistemas en el conjunto. El *valor medio de u* (o *media del conjunto* de u), que llamaremos $\bar{u}$ viene *definido* así por

$$\bar{u} \equiv \frac{\mathcal{N}_1 u_1 + \mathcal{N}_2 u_2 + \cdots + \mathcal{N}_\alpha u_\alpha}{\mathcal{N}} = \frac{\sum_{r=1}^{\alpha} \mathcal{N}_r u_r}{\mathcal{N}}. \qquad (29)$$

Pero como $\mathcal{N}_r/\mathcal{N} \equiv P_r$ es la probabilidad de que se presente el valor u_r, la definición (29) se reduce simplemente a [14]

$$\boxed{\bar{u} \equiv \sum_{r=1}^{\alpha} P_r u_r} \tag{30}$$

De modo semejante, si $f(u)$ es una función cualquiera de u, el *valor medio* (o *media del conjunto*) de f se define por la expresión

$$\boxed{\bar{f}(u) \equiv \sum_{r=1}^{\alpha} P_r f(u_r)} \tag{31}$$

Esta definición implica que los valores medios tienen algunas propiedades muy sencillas. Por ejemplo, si $f(u)$ y $g(u)$ son dos funciones de u,

$$\overline{f+g} \equiv \sum_{r=1}^{\alpha} P_r[f(u_r) + g(u_r)] = \sum_{r=1}^{\alpha} P_r f(u_r) + \sum_{r=1}^{\alpha} P_r g(u_r)$$

o

$$\boxed{\overline{f+g} = \bar{f} + \bar{g}} \tag{32}$$

Este resultado muestra con mucha generalidad que el valor medio de una suma de términos es igual a la suma de los valores medio de estos términos. Así pues, las operaciones sucesivas de realizar una suma y tomar una media dan el mismo resultado sin importar el orden en que se llevan a cabo [15]. De modo semejante si c es una constante

$$\overline{cf} = \sum_{r=1}^{\alpha} P_r[cf(u_r)] = c\sum_{r=1}^{\alpha} P_r f(u_r)$$

o

$$\boxed{\overline{cf} = c\bar{f}} \tag{33}$$

Así pues, las operaciones de multiplicar por una constante y tomar una media pueden llevarse a cabo en cualquier orden sin que influya sobre el resultado. Si $f = 1$, la relación (33) equivale a

[14] El valor medio $\bar{u}$ es dependiente del tiempo si el conjunto lo es, es decir, si alguna de las probabilidades P_r depende del tiempo. Obsérvese también que el *valor medio* o *media del conjunto* $\bar{u}$ es una media calculada sobre todos los sistemas del conjunto en un instante particular. Ordinariamente es diferente de la media en el tiempo definida en (1.6) para un sistema aislado, excepto en el caso especial de conjuntos independientes del tiempo, en donde la media en el tiempo se toma en un intervalo de tiempo muy largo.

[15] En lenguaje matemático se diría que estas operaciones "conmutan".

la afirmación obvia de que el valor medio de una constante es simplemente igual a dicha constante.

Ejemplo

Considérese un sistema de 4 spines en el que $p = q = \frac{1}{2}$. El número de momentos que señalan hacia arriba puede ser $n = 0, 1, 2, 3, 4$. Estos números aparecen con probabilidades $P(n)$ que se deducen inmediatamente de (16) y que ya fueron calculadas con gran sencillez en (1.4*a*). Como se indicaba en la Fig. 2,6, estas probabilidades son, respectivamente,

$$P(n) = \tfrac{1}{16}, \tfrac{4}{16}, \tfrac{6}{16}, \tfrac{4}{16}, \tfrac{1}{16}$$

El número medio de momentos señalando hacia arriba es entonces

$$\begin{aligned}\bar{n} &= \sum_{n=0}^{4} P(n)n \\ &= (\tfrac{1}{16} \times 0) + (\tfrac{4}{16} \times 1) + (\tfrac{6}{16} \times 2) \\ &\quad + (\tfrac{4}{16} \times 3) + (\tfrac{1}{16} \times 4) \\ &= 2\end{aligned}$$

Obsérvese que este resultado es simplemente igual a $Np = 4 \times \frac{1}{2}$.

Como $p = q$, no hay ningún sentido preferido en el espacio. El número medio de momentos señalando hacia abajo debe, por tanto, ser igual al número medio de momentos señalando hacia arriba, es decir,

$$\overline{n'} = \bar{n} = 2$$

Este resultado se sigue también de la relación (32), que nos permite escribir

$$\overline{n'} = \overline{N - n} = \bar{N} - \bar{n} = 4 - 2 = 2$$

Como no existe ningún sentido preferido en el espacio, el momento magnético medio debe ser nulo. Ciertamente se tiene

$$\bar{m} = \overline{n - n'} = \bar{n} - \overline{n'} = 2 - 2 = 0$$

El valor de $\bar{m}$ podía haberse calculado también directamente, utilizando la probabilidad $P'(m)$ de que m tome sus valores posibles $m = -4, -2, 0, 2, 4$. Así se tiene, por definición

$$\begin{aligned}\bar{m} &= \sum_{m} P'(m)m \\ &= [\tfrac{1}{16} \times (-4)] + [\tfrac{4}{16} \times (-2)] \\ &\quad + [\tfrac{6}{16} \times 0] + [\tfrac{4}{16} \times 2] + [\tfrac{1}{16} \times 4] \\ &= 0\end{aligned}$$

Con frecuencia es importante otra propiedad de los valores medios. Supongamos que nos estamos refiriendo a dos variables u y v, que pueden tomar los valores

$$u_1, u_2, \ldots, u_\alpha$$
$$v_1, v_2, \ldots, v_\beta$$

respectivamente. Llamemos P_r a la probabilidad de que u tome el valor u_r y P_s a la de que v adquiera el valor v_s. Si la probabilidad de que u tome cualquiera de sus valores es independiente de los que adquiere v (es decir, si las variables u y v son estadísticamente independientes), entonces la probabilidad compuesta P_{rs} de que u tome el valor u_r y v el valor v_s es, según (5), simplemente igual a

$$P_{rs} = P_r P_s \tag{34}$$

Supongamos ahora que $f(u)$ es una función cualquiera de u, mientras que $g(v)$ es otra función cualquiera de v. Entonces el valor medio del producto fg viene, según la definición (31), dado con toda generalidad por

$$\overline{f(u)g(v)} \equiv \sum_{r=1}^{\alpha} \sum_{s=1}^{\beta} P_{rs} f(u_r) g(v_s) \tag{35}$$

en donde la suma se extiende a todos los valores posibles u_r y v_s de las variables. Si éstas son estadísticamente independientes de modo que (34) es válida, la ecuación (35) se reduce a

$$\begin{aligned} \overline{fg} &= \sum_r \sum_s P_r P_s f(u_r) g(v_s) \\ &= \sum_r \sum_s [P_r f(u_r)][P_s g(v_s)] \\ &= \Big[\sum_r P_r f(u_r)\Big]\Big[\sum_s P_s g(v_s)\Big] \end{aligned}$$

Pero el primero de los factores de la derecha es simplemente el valor medio de f, mientras que el segundo es el valor medio de g. Se llega, pues, al resultado de que

$$\boxed{\begin{array}{c}\text{si } u \text{ y } v \text{ son estadísticamente independientes} \\ \overline{fg} = \bar{f}\,\bar{g}\end{array}} \tag{36}$$

es decir, la media de un producto es entonces igual simplemente al producto de las medias.

Dispersión

Supongamos que una variable u toma sus valores posibles u_r con las probabilidades respectivas P_r. Algunas características generales de la distribución de probabilidad pueden obtenerse entonces mediante el empleo de ciertos parámetros de interés. Uno de ellos es simplemente el mismo valor medio de u, es decir, la magnitud $\bar{u}$ definida en (30). Este parámetro indica el valor central de u alrededor del cual se distribuyen los diversos valores u_r. Puede ser conveniente entonces medir los valores posibles de $\bar{u}$ respecto a su valor medio escribiendo

$$\Delta u \equiv u - \bar{u} \tag{37}$$

en donde Δu es la desviación de u respecto al valor medio $\bar{u}$. Obsérvese que el valor medio de esta desviación se anula. Ciertamente, utilizando la propiedad (32),

$$\overline{\Delta u} = (\overline{u - \bar{u}}) = \bar{u} - \bar{u} = 0. \tag{38}$$

También es útil definir un parámetro que mida la amplitud con que se extienden todos los valores posibles de u alrededor de su valor medio $\bar{u}$. El valor medio de Δu no proporciona por sí mismo esta medida, puesto que Δu es, en valor medio, tantas veces positivo como negativo de modo que su valor medio se anula de acuerdo con (38). Por otra parte, la cantidad $(\Delta u)^2$ no puede ser nunca negativa. Su valor medio, definido por

$$\overline{(\Delta u)^2} \equiv \sum_{r=1}^{\alpha} P_r(\Delta u_r)^2 \equiv \sum_{r=1}^{\alpha} P_r(u_r - \bar{u})^2 \tag{39}$$

se denomina *dispersión* (o *varianza*) de u y tampoco puede ser negativo, puesto que cada término de la suma (39) es no negativo.[16] Así pues,

$$\overline{(\Delta u)^2} \geq 0 \tag{40}$$

La dispersión sólo puede anularse si *todos* los valores de u_r son iguales a $\bar{u}$; y aumenta progresivamente cuando estos valores tienen una apreciable probabilidad de presentarse lejos de $\bar{u}$. La dispersión, por lo tanto, proporciona una medida conveniente del orden de la repartición de los valores adquiridos por u.

La dispersión $\overline{(\Delta u)^2}$ es una magnitud que tiene las dimensiones del cuadro de u. Una medida lineal de la repartición de los valores posibles de u viene dada por la raíz cuadrada de la dispersión, es decir, por la magnitud

$$\boxed{\underline{\Delta} u \equiv [\overline{(\Delta u)^2}]^{1/2}} \tag{41}$$

que tiene las mismas dimensiones que u y que se denomina *desviación standard* de u. La definición (39) muestra que incluso unos pocos valores de u que aparezcan con una probabilidad apreciable alejados de $\bar{u}$ tendrán una contribución importante en $\underline{\Delta} u$. La mayoría de los valores de u deben presentarse por tanto dentro de un intervalo del orden de $\underline{\Delta} u$ alrededor de su valor medio $\bar{u}$.

---

[16] Obsérvese que $\overline{(\Delta u)^2}$ es diferente de $(\overline{\Delta u})^2$, es decir, hay una gran diferencia en elevar al cuadrado primero y luego tomar la media y realizar las operaciones en orden inverso.

Ejemplo

Volvamos al ejemplo anterior de cuatro spines cuando $p = q = \frac{1}{2}$. Como $\bar{n} = 2$, la dispersión de n es, por definición

$$\overline{(\Delta n)^2} \equiv \sum_n P(n)(n-2)^2$$

$$= [\tfrac{1}{16} \times (-2)^2] + [\tfrac{4}{16} \times (-1)^2] + [\tfrac{6}{16} \times (0)^2] + [\tfrac{4}{16} \times (1)^2] + [\tfrac{1}{16} \times (2)^2]$$

$$= 1$$

De aquí que la desviación standard de n sea

$$\Delta n = \sqrt{1} = 1$$

Análogamente se puede calcular la dispersión del momento magnético. Como $\bar{m} = 0$, se tiene por definición de modo que

$$\overline{(\Delta m)^2} \equiv \sum_m P'(m)(m-0)^2$$

$$= [\tfrac{1}{16} \times (-4)^2] + [\tfrac{4}{16} \times (-2)^2] + [\tfrac{6}{16} \times (0)^2] + [\tfrac{4}{16} \times (2)^2] + [\tfrac{1}{16} \times (4)^2] = 4$$

de modo que

$$\Delta m = \sqrt{4} = 2$$

Comprobemos que los resultados anteriores concuerdan. Como $\bar{m} = 0$, mientras que $\bar{n} = \bar{n}' = 2$, se tiene para todos los valores de m o n

$$\Delta m = m = n - n' = n - (4 - n)$$

$$= 2n - 4 = 2(n-2)$$

o bien $\Delta m = 2(n - \bar{n}) = 2\,\Delta n$

De aquí que $\overline{(\Delta m)^2} = 4\overline{(\Delta n)^2}$

De acuerdo con lo que hemos encontrado mediante el cálculo explícito.

Un conocimiento de las probabilidades P_r para todos los valores u_r da una información estadística completa sobre la distribución de los valores de u en el conjunto. Por otra parte, el conocimiento de algunos valores medios como $\bar{u}$ y $\overline{(\Delta u)^2}$ proporciona sólo una información parcial de las características de esta distribución y no es suficiente para determinar las probabilidades P_r sin ambigüedad. Estos valores medios, sin embargo, pueden calcularse a veces con gran sencillez sin un conocimiento explícito de las probabilidades, incluso en los casos en que el cálculo real de estas probabilidades, sería una tarea difícil. Aclararemos estos detalles en la sección siguiente.

2.5 *Cálculo de los valores medios para un sistema de spines*

Consideremos un gas ideal de N spines $\frac{1}{2}$. El hecho de que estos spines sean estadísticamente independientes nos permite calcular diversos valores medios en condiciones muy generales de un modo sencillo. El cálculo puede llevarse a cabo sin necesidad de evaluar ninguna probabilidad como la $P(n)$ obtenida en (14).

Empecemos entonces nuestra investigación por una magnitud físicamente interesante de este sistema de spines, su momento magnético total M en el sentido hacia arriba. Llamemos μ_i al componente hacia arriba del spin i. El momento magnético total M es entonces simplemente igual a la suma de los momentos magnéticos de todos los spines de modo que

$$M = \mu_1 + \mu_2 + \cdots + \mu_N$$

o, más abreviadamente,

$$M = \sum_{i=1}^{N} \mu_i \tag{42}$$

Deseamos calcular el valor medio y la dispersión de este momento magnético total.

Para calcular el valor medio de M, necesitamos únicamente tomar los valores medios de ambos miembros de (42). La propiedad general del promedio (32), que nos permite intercambiar el orden de promediar y sumar, nos conduce inmediatamente al resultado

$$\overline{M} = \overline{\sum_{i=1}^{N} \mu_i} = \sum_{i=1}^{N} \overline{\mu_i} \tag{43}$$

Pero la probabilidad de que cualquier momento magnético tenga una orientación dada (hacia arriba o hacia abajo) es la misma para cada momento; de aquí que el momento magnético medio sea el mismo para cada spin (es decir, $\bar{\mu}_1 = \bar{\mu}_2 = \ldots = \bar{\mu}_N$) y puede llamarse sencillamente $\bar{\mu}$. La suma de (43) se compone por tanto, de N términos iguales, de modo que (43) se reduce simplemente a

$$\boxed{\overline{M} = N\bar{\mu}} \tag{44}$$

Este resultado es casi evidente por sí mismo; afirma simplemente que el momento magnético total medio de N spines es N veces mayor que el momento medio de un spin.

Calculemos ahora la dispersión de M, es decir, la magnitud $\overline{(\Delta M)^2}$ en donde

$$\Delta M \equiv M - \overline{M} \tag{45}$$

Restando (43) de (42) se tiene

$$M - \overline{M} = \sum_{i=1}^{N} (\mu_i - \bar{\mu})$$

o bien

$$\Delta M = \sum_{i=1}^{N} \Delta\mu_i \tag{46}$$

en donde

$$\Delta\mu_i \equiv \mu_i - \bar{\mu} \tag{47}$$

Para hallar $(\Delta M)^2$ necesitamos únicamente multiplicar (46) por sí mismo. Así pues,

$$\begin{aligned}(\Delta M)^2 &= (\Delta\mu_1 + \Delta\mu_2 + \cdots + \Delta\mu_N)\,(\Delta\mu_1 + \Delta\mu_2 + \cdots + \Delta\mu_N)\\ &= [(\Delta\mu_1)^2 + (\Delta\mu_2)^2 + (\Delta\mu_3)^2 + \cdots + (\Delta\mu_N)^2]\\ &\quad + [\Delta\mu_1\,\Delta\mu_2 + \Delta\mu_1\,\Delta\mu_3 + \cdots + \Delta\mu_1\,\Delta\mu_N\\ &\qquad + \Delta\mu_2\,\Delta\mu_1 + \Delta\mu_2\,\Delta\mu_3 + \cdots + \Delta\mu_N\,\Delta\mu_{N-1}]\end{aligned}$$

o bien

$$(\Delta M)^2 = \sum_{i=1}^{N} (\Delta\mu_i)^2 + \sum_{\substack{i=1\\ i\neq j}}^{N} \sum_{j=1}^{N} (\Delta\mu_i)(\Delta\mu_j) \tag{48}$$

El primer término del segundo miembro, representa todos los términos al cuadrado que surgen de los términos de la suma (46) que se multiplican por sí mismos; el segundo término representa todos los términos cruzados que surgen al multiplicar los términos *diferentes* de la suma (46). Tomando el valor medio de (48) y usando de nuevo la propiedad (32) que nos permite intercambiar el orden de sumar y promediar resulta:

$$\overline{(\Delta M)^2} = \sum_{i=1}^{N} \overline{(\Delta\mu_i)^2} + \sum_{\substack{i=1\\ i\neq j}}^{N} \sum_{j=1}^{N} \overline{(\Delta\mu_i)(\Delta\mu_j)} \tag{49}$$

Todos los productos de la segunda suma, en que $i \neq j$, se refieren a spines diferentes. Pero como éstos son estadísticamente independientes, la propiedad (36) implica que el valor medio de cada producto es simplemente igual al producto de los valores medios de sus factores. Así pues,

para $i \neq j$

$$\overline{(\Delta\mu_i)(\Delta\mu_j)} = (\overline{\Delta\mu_i})(\overline{\Delta\mu_j}) = 0 \tag{50}$$

puesto que,

$$\overline{\Delta\mu_i} = \overline{\mu_i} - \overline{\mu} = 0$$

En resumen, cada término cruzado de (49) se anula al promediar, ya que es negativo o positivo con la misma frecuencia. De aquí que (50) se reduzca simplemente a una suma de términos al cuadrado (ninguno de los cuales puede ser negativo):

$$\overline{(\Delta M)^2} = \sum_{i=1}^{N} \overline{(\Delta\mu_i)^2} \tag{51}$$

El razonamiento se hace ahora idéntico al que seguía a la ecuación (43). La probabilidad de que cualquier momento tenga una orientación dada cualquiera, es la misma para cada momento; de aquí que la dispersión $\overline{(\Delta\mu_i)^2}$ es la misma para cada spin [es decir, $\overline{(\Delta\mu_1)^2} = \overline{(\Delta\mu_2)^2} = \ldots = \overline{(\Delta\mu_N)^2}$] y puede representarse simplemente por $\overline{(\Delta\mu)^2}$. La suma en (51) se compone, pues, de N términos iguales y se reduce simplemente a

$$\boxed{\overline{(\Delta M)^2} = N\overline{(\Delta\mu)^2}} \tag{52}$$

Esta relación afirma que la dispersión del momento magnético total es sencillamente N veces mayor que la dispersión del momento magnético de un spin individual. En correspondencia, (52) implica también que

$$\boxed{\underline{\Delta} M = \sqrt{N}\,\underline{\Delta}\mu} \tag{53}$$

en donde $\quad \underline{\Delta} M \equiv [\overline{(\Delta M)^2}]^{1/2} \qquad$ y $\qquad \underline{\Delta}\mu \equiv [\overline{(\Delta\mu)^2}]^{1/2}$

son, de acuerdo con la definición general (41), las desviaciones standard del momento magnético total y del momento magnético por spin, respectivamente.

Las relaciones (44) y (53) muestran explícitamente cómo dependen $\overline{M}$ y $\underline{\Delta} M$ del número total N de spines del sistema. Cuando $\bar{\mu} \neq 0$, el momento magnético total $\overline{M}$ aumenta proporcionalmente a N. La desviación standard $\underline{\Delta} M$ (que mide la anchura de la distribución de valores de M alrededor de su valor medio $\overline{M}$) aumenta también cuando N crece, pero sólo proporcionalmente a $N^{1/2}$. De aquí que la magnitud *relativa* de $\underline{\Delta} M$ comparada con $\overline{M}$ decrece proporcionalmente a $N^{-1/2}$; realmente (44) y (53) implican que

$$\text{para } \bar{\mu} \neq 0 \qquad \frac{\underline{\Delta} M}{\overline{M}} = \frac{1}{\sqrt{N}}\left(\frac{\underline{\Delta}\mu}{\bar{\mu}}\right) \tag{54}$$

La figura 2.8 muestra esta tendencia característica.

Obsérvese que los resultados (44) y (53) son muy generales. Dependen únicamente de la relación aditiva (43) y del hecho de que los spines sean estadísticamente independientes. Todas nuestras consideraciones permanecerían, pues, igualmente válidas aunque los componentes de μ_i de cada momento magnético pudieran adquirir muchos valores posibles. (Este sería el caso si el spin de cada partícula fuera mayor de $\frac{1}{2}$, de modo que pudiera presentar más de dos orientaciones posibles en el espacio.)

Sistema de partículas con spin 1/2

Los resultados anteriores se aplican fácilmente al caso especial corriente en que cada partícula tiene spin $\frac{1}{2}$. Como es normal, supongamos que sus momentos magnéticos tienen entonces la probabilidad p de señalar hacia arriba de modo que $\mu_i = \mu_0$ y la probabilidad $q = 1 - p$ de señalar hacia abajo, de modo que $\mu_i = -\mu_0$. Su momento medio en sentido hacia arriba es, pues,

$$\bar{\mu} \equiv p\mu_0 + q(-\mu_0) = (p - q)\mu_0 = (2p - 1)\mu_0 \tag{55}$$

Como comprobación notemos que, en el caso simétrico en que $p = q$, $\bar{\mu} = 0$, como era de esperar.

La dispersión del momento magnético de un spin viene dada por

$$\overline{(\Delta\mu)^2} \equiv \overline{(\mu - \bar{\mu})^2} \equiv p(\mu_0 - \bar{\mu})^2 + q(-\mu_0 - \bar{\mu})^2 \tag{56}$$

Pero $\qquad \mu_0 - \bar{\mu} = \mu_0 - (2p - 1)\mu_0 = 2\mu_0(1 - p) = 2\mu_0 q$

y $\qquad \mu_0 + \bar{\mu} = \mu_0 + (2p - 1)\mu_0 = 2\mu_0 p$

Así pues, (56) se reduce a

$$\overline{(\Delta\mu)^2} = p(2\mu_0 q)^2 + q(2\mu_0 p)^2 = 4\mu_0{}^2 pq(q + p)$$

o bien

$$\overline{(\Delta\mu)^2} = 4pq\mu_0{}^2 \tag{57}$$

puesto que $p + q = 1$.

Las relaciones (44) y (52) llevan, por tanto, a los resultados

$$\boxed{\bar{M} = N(p - q)\mu_0} \tag{58}$$

y

$$\boxed{\overline{(\Delta M)^2} = 4Npq\,\mu_0{}^2} \tag{59}$$

La desviación standard de M es, por consiguiente,

$$\boxed{\Delta M = 2\sqrt{Npq}\,\mu_0} \tag{60}$$

Si escribimos $M = m\mu_0$, de modo que el número entero $m = M/\mu_0$ exprese el momento magnético total en unidades de μ_0, los resultados (58) a (60) pueden expresarse también en la forma

$$\bar{m} = N(p - q) = N(2p - 1) \tag{61}$$

m = n - n'

$$\overline{(\Delta m)^2} = 4Npq \tag{62}$$

$$\Delta m = 2\sqrt{Npq} \tag{63}$$

Estas relaciones contienen una cantidad apreciable de información sobre la distribución de los valores posibles de M o m en el conjunto de sistemas de spines. Así sabemos que únicamente se presentan con probabilidad apreciable aquellos valores de m que caen cerca de $\overline{m}$ y no difieren de éste en una cantidad mucho mayor que Δm. La fig. 2.8 nos proporciona una muestra específica.

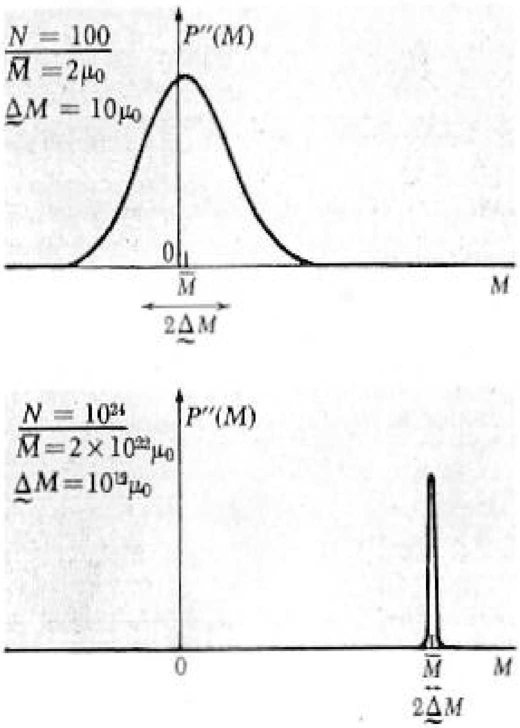


Fig. 2.9 Probabilidad $P''(M)$ de que el momento magnético total de un sistema de spines tenga un valor M cuando $N = 100$ y cuando $N = 10^{24}$. El campo magnético es tal que $p = 0{,}51$ y $q = 0{,}49$. Los gráficos indican la curva envolvente de los valores posibles de $P''(M)$, pero no se han dibujado a la misma escala.

Ejemplo

Supóngase que, en presencia de cierto campo magnético aplicado $\mathbf{B}$, el momento magnético de cada spin tiene una probabilidad $p = 0{,}51$ de orientarse paralelamente a $\mathbf{B}$, y una probabilidad de $q = 1 - p = 0{,}49$ de orientarse antiparalelamente a $\mathbf{B}$. El momento magnético total medio de un sistema de N spines es entonces,

$$\overline{M} = 0{,}02N\mu_0$$

La desviación estandard de su momento magnético total viene dado por (60), de modo que

$$\Delta M = 2\sqrt{Npq}\,\mu_0 \simeq \sqrt{N}\mu_0$$

de aquí que $\quad \dfrac{\Delta M}{\overline{M}} \simeq \dfrac{\sqrt{N}\mu_0}{0{,}02N\mu_0} = \dfrac{50}{\sqrt{N}}$

Consideremos primero un caso en el que el número total de spines es relativamente pequeño. Por ejemplo, supongamos que $N = 100$. Entonces

$$\frac{\Delta M}{\overline{M}} \simeq \frac{50}{\sqrt{100}} = 5$$

de modo que $\Delta M > \overline{M}$. Los posibles valores de M están entonces muy esparcidos. Realmente, es muy probable que existan valores de M que difieran ampliamente de $\overline{M}$ e incluso sean de signo contrario (Véase Figura 2.9).

Por otra parte, consideremos el caso de un sistema macroscópico de spines en el que N es del orden del número de Avogadro, o sea, $N = 10^{24}$. Entonces

$$\frac{\Delta M}{\overline{M}} \simeq \frac{50}{\sqrt{10^{24}}} = 5 \times 10^{-11}$$

de modo que $\Delta M \lll \overline{M}$. Los valores posibles de M están entonces muy poco esparcidos relativamente alrededor de su valor medio. Si pretendemos medir el momento magnético total del sistema, mediremos casi siempre un valor muy próximo a $\overline{M}$. Realmente, a menos que nuestro método de medición sea suficientemente preciso para detectar diferencias de momentos magnéticos menores de una parte de 10^{10}, siempre mediríamos virtualmente un momento magnético igual a $\overline{M}$ sin advertir la existencia de fluctuaciones alrededor de este valor. Este ejemplo demuestra concretamente la conclusión general de que la magnitud *relativa* de las fluctuaciones tienden a ser muy pequeños en un sistema macroscópico compuesto por muchísimas partículas.

Distribución de moléculas en un gas ideal

Consideremos un gas ideal de N moléculas contenidas en una caja de volumen V_0. Estamos interesados en investigar el número n de moléculas que se hallan dentro de un subvolumen especificado V de esta caja (véase Fig. 2.10). Si el gas está en equilibrio, la probabilidad p de hallar una molécula en este volumen V es simplemente igual a

$$p = \frac{V}{V_0} \tag{64}$$

como se mencionó previamente en (28)

Es muy fácil de calcular el valor medio de n y su dispersión. Hemos señalado ya al final de la Sección 2.3 que el problema del gas ideal es análogo al del sistema de spines. (Ambos problemas son del tipo que conduce a una distribución binómica.) De aquí que podamos aplicar inmediatamente los resultados (61) y (62) para hallar la información deseada sobre n. Llamaremos n' al número de moléculas en el volumen restante $V_0 - V$ de la caja y sea $m \equiv n - n'$.

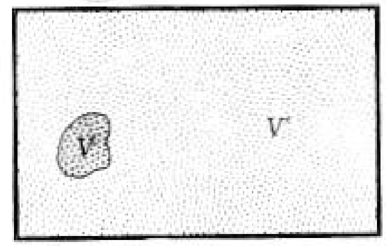


Fig. 2.10 Una caja de volumen V_0 contiene N moléculas de un gas ideal. En un instante dado cualquiera se encuentran situados cierto número de moléculas n en el subvolumen V, mientras que el número restante $n' = N - n$ está situado en el resto del volumen $V' = V_0 - V$.

Como se vio en (25), se tiene

$$n = \tfrac{1}{2}(N + m) \tag{65}$$

Utilizando el resultado (61) para $\overline{m}$, obtendremos entonces

$$\bar{n} = \tfrac{1}{2}(N + \overline{m}) = \tfrac{1}{2}N(1 + p - q)$$

o bien

$$\boxed{\bar{n} = Np} \tag{66}$$

puesto que, $q = 1 - p$. Además, a partir de (65) se obtiene la relación

$$\Delta n \equiv n - \bar{n} = \tfrac{1}{2}(N + m) - \tfrac{1}{2}(N + \overline{m}) = \tfrac{1}{2}[m - \overline{m}]$$

o bien

$$\Delta n = \tfrac{1}{2}\,\Delta m$$

De aquí que

$$\overline{(\Delta n)^2} = \tfrac{1}{4}\overline{(\Delta m)^2}$$

y (62) implica que[17]

$$\boxed{\overline{(\Delta n)^2} = Npq} \tag{67}$$

La desviación standard de n es entonces

$$\boxed{\underset{\sim}{\Delta} n = \sqrt{Npq}} \tag{68}$$

de modo que

$$\frac{\underset{\sim}{\Delta} n}{\bar{n}} = \frac{\sqrt{Npq}}{Np} = \left(\frac{q}{p}\right)^{1/2} \frac{1}{\sqrt{N}} \tag{69}$$

Estas relaciones muestran nuevamente que la desviación standard $\underset{\sim}{\Delta} n$ aumenta proporcionalmente a $N^{1/2}$. En correspondencia, el valor relativo $\underset{\sim}{\Delta} n/\bar{n}$ de la desviación standard *de*crece proporcionalmente a $N^{-1/2}$ y así resulta ser muy pequeña cuando N es grande. Estas consecuencias se comprenden bien en el caso especial del Capítulo 1 en donde considerábamos el número n de moléculas contenidas en una mitad de una caja. En este caso (64) implica que

[17] Las relaciones (66) y (67) podrían deducirse también directamente por los métodos de esta sección sin utilizar los resultados correspondientes de la magnitud m (véase el prob. 2.14).

$p = q = \frac{1}{2}$, de modo que (66) se reduce al resultado evidente

$$\bar{n} = \tfrac{1}{2}N$$

mientras que
$$\frac{\Delta n}{\bar{n}} = \frac{1}{\sqrt{N}}$$

Estas relaciones sitúan el estudio de las fluctuaciones de la Sección 1.1 sobre una base cuantitativa. El hecho de que el valor absoluto de las fluctuaciones (medidas por Δn) aumenta cuando crece N mientras que el valor relativo de las fluctuaciones (medidas por $\Delta n/n$) disminuye al crecer N se muestra explícitamente por los gráficos de las Figs. 1.5 y 1.6 para $N = 4$ y $N = 40$. Cuando la caja contiene alrededor de un mol de gas, N es del orden del número de Avogadro, de modo que $N \sim 10^{24}$. En este caso el valor relativo de las fluctuaciones $\Delta n/\bar{n} \sim 10^{-12}$ resulta tan pequeño que será despreciable casi siempre.

2.6 *Distribuciones continuas de probabilidad*

Consideremos un sistema ideal de spines compuesto por un gran número N de spines $\frac{1}{2}$. Existen muchos valores posibles del momento magnético total del sistema. Realmente, según (22) y (24)

$$M = m\mu_0 = (2n - N)\mu_0 \qquad \text{(70)}$$

$n\mu_0$ -

de modo que M puede adquirir cualquiera de los $(N + 1)$ valores posibles

$$M = -N\mu_0,\ -(N-2)\mu_0,\ -(N-4)\mu_0, \ldots, (N-2)\mu_0,\ N\mu_0 \qquad \text{(71)}$$

La probabilidad $P''(M)$ de que el momento magnético total adquiera un valor particular M es igual a la probabilidad de que se presente el valor correspondiente de m o n, es decir, a $P'(m)$ dado por (26) o $P(n)$ dado por (14).

Así pues
$$P''(M) = P'(m) = P(n)$$
en donde
$$m = \frac{M}{\mu_0} \qquad \text{y} \qquad n = \tfrac{1}{2}(N + m) \qquad \text{(72)}$$

Excepto en los casos en que M es cercano a sus valores extremos posibles $\pm N\mu_0$ [en donde $P''(M)$ es extraordinariamente pequeño], la probabilidad $P''(M)$ no varía apreciablemente al pasar de un posible valor de M al adyacente; es decir, $|P''(M + 2\mu_0) - P''(M)| \ll P''(M)$. La envolvente de los valores posibles de $P''(M)$ forma entonces una curva suave, como se indica en la Fig. 2.11. Así pues, es posible considerar a $P''(M)$ como una función que varía suavemente con la variable continua M, aunque únicamente tengan interés los valores discretos (71) de M.

Supóngase que μ_0 es despreciablemente pequeño comparado con el menor momento magnético de interés en cualquier medición macroscópica. El hecho de que M sólo pueda adquirir valores discretos separados por $2\mu_0$ es entonces inobservable dentro de la precisión de las observaciones realizadas. Así pues, M puede considerarse ciertamente como una variable continua. Además, se puede hablar con pleno significado de dM como un "infinitésimo macroscópico", es decir, una magnitud que es *macro*scópicamente muy pequeña aunque sea *micro*scópicamente grande. (En otras palabras, se supone que dM es despreciablemente pequeño comparado con el menor momento magnético de interés en un estudio macroscópico, aunque sea mucho mayor μ_0)[18]. Resulta interesante entonces la siguiente cuestión: ¿Cuál es la probabilidad de que el momento magnético total del sistema esté incluido dentro de un pequeño intervalo entre M y $M + dM$? El valor de esta probabilidad depende evidentemente del intervalo dM y debe tender a anularse cuando dM se hace despreciablemente pequeño. Se espera, por tanto, que esta probabilidad sea simplemente proporcional a dM de modo que pueda escribirse en la forma:

$$\left[\begin{array}{l}\text{Probabilidad de que el momento mag-}\\ \text{nético total está comprendido entre}\\ M \text{ y } M + dM.\end{array}\right] = \mathcal{P}(M)\,dM \qquad (73)$$

en donde $\mathcal{P}(M)$ es independiente del valor de dM [19]. La magnitud $\mathcal{P}(M)$ se denomina *densidad de probabilidad*; nos da una probabilidad real cuando se multiplica por el intervalo infinitesimal dM.

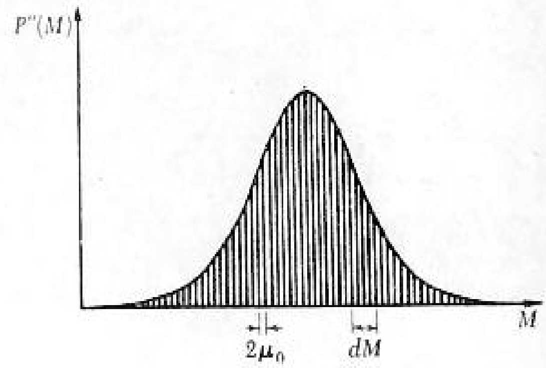


Fig. 2.11 La probabilidad $P''(M)$ de que el momento magnético total de un sistema de spines tenga un valor M en el caso en que el número de spines N es grande y el momento magnético de un spin μ_0 es pequeño.

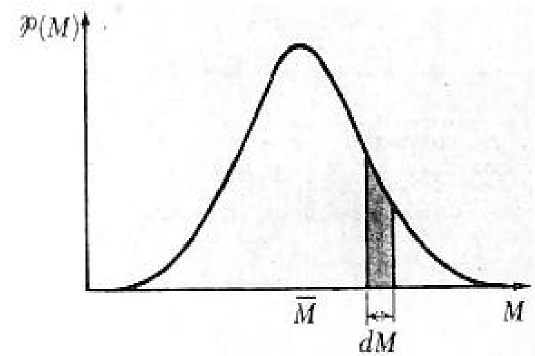


Fig. 2.12 La distribución de probabilidades de la Fig. 2.11 se muestra aquí expresada en función de la densidad de probabilidad $\mathcal{P}(M)$. $\mathcal{P}(M)\,dM$ [que es igual al área contenida bajo la curva en el intervalo pequeño entre M y $M + dM$] es la probabilidad de que el momento magnético caiga dentro del intervalo mencionado.

[18] Es interesante señalar que muchos de los infinitésimos utilizados en física son infinitésimos macroscópicos. Por ejemplo, en el estudio de la electricidad, se suele hablar de la carga Q que posee un cuerpo y de su incremento de carga dQ. Esta descripción diferencial es válida si se entiende que dQ ha de ser mucho mayor que la carga electrónica discreta e, aunque se suponga que es despreciablemente pequeño comparado con la carga Q misma.

[19] Como la probabilidad es una cierta función suave de dM, deberá poder expresarse cerca de cualquier valor de M como una serie de Taylor en potencias de dM cuando dM es pequeño. Así debe ser de la forma

$$\text{Probabilidad} = a_0 + a_1\,dM + a_2(dM)_2 + \ldots$$

en donde los coeficientes a_0, a_1, ... dependen de M. Aquí $a_0 = 0$, puesto que la probabilidad debe tender hacia cero cuando dM se hace muy pequeño; además los términos en que aparecen potencias superiores a dM, son despreciablemente pequeños comparados con el término principal que es proporcional a dM. De aquí se obtiene el resultado (73).

La probabilidad (73) se expresa explícitamente con facilidad en función de la probabilidad $P''(M)$ de que el momento magnético total adquiera el valor discreto particular M. Como (71) muestra que los valores posibles de M están separados por cantidades $2\mu_0$ y como $dM \gg 2\mu_0$, el intervalo entre M y $M + dM$ contiene $dM/(2\mu_0)$ valores posibles de M. Todos ellos se presentan con casi la misma probabilidad $P''(M)$ puesto que ésta varía muy lentamente en el pequeño intervalo en cuestión. De aquí que la probabilidad de que el momento total esté en el intervalo entre M y $M + dM$, se obtiene simplemente sumando $P''(M)$ respecto a todos los valores de M comprendidos en este intervalo, es decir, multiplicando el valor casi constante $P''(M)$ por $dM/(2\mu_0)$. Esta probabilidad es, pues, adecuadamente proporcional a dM y viene explícitamente dada por

$$\mathcal{P}(M)\, dM = P''(M)\,\frac{dM}{2\mu_0} \tag{74}$$

En la práctica, el cálculo real de $P''(M)$ puede ser laborioso si M/μ_0 es grande puesto que la distribución binómica (14) exige entonces el cálculo de factoriales grandes. Sin embargo, pueden soslayarse estas dificultades utilizando la aproximación gaussiana del Apéndice A.1.

Existen muchos problemas en los que una variable de interés, llamémosla u, es intrínsecamente continua. Por ejemplo, u puede designar el ángulo entre algún vector en un plano y una dirección fija; este ángulo puede tomar entonces cualquier valor en el dominio entre 0 y 2π. En el caso general u puede adquirir cualquier valor en el dominio $a_1 \leq u \leq a_2$. (Este dominio puede ser de extensión infinita, es decir, $a_1 \to -\infty$ o $a_2 \to \infty$ o ambos.) Pueden llegarse a conclusiones sobre la probabilidad de esta variable de una manera completamente análoga a la estudiada en el caso de M. Así pues, podemos enfocar nuestra atención sobre un intervalo infinitesimal entre u y $u + du$ y considerar la probabilidad de que la variable esté en dicho intervalo. Cuando du es suficientemente pequeño, esta probabilidad debe ser nuevamente proporcional a du, de modo que puede escribirse en la forma $\mathcal{P}(u)\, du$ en donde la magnitud $\mathcal{P}(u)$ es una *densidad de probabilidad* independiente del tamaño de du.

Las consideraciones sobre probabilidades en que interviene una variable continua u pueden reducirse fácilmente al caso más sencillo en que los valores posibles de las variables son discretos y pueden por tanto, contarse. Es necesario únicamente subdividir el dominio de los posibles valores de u en intervalos iguales arbitrariamente pequeños de tamaño fijo δu. Cada uno de estos intervalos puede entonces designarse por u_r y la probabilidad de que u esté comprendida en este intervalo por P_r o $P(u)$. Este procedimiento nos permite tratar con un conjunto enumerable de valores de la variable u, cada uno de los cuales corresponde a uno de los intervalos infinitesimales $r = 1, 2, 3, \ldots$ También resulta evidente que las relaciones en que intervienen probabilidades de variables discretas permanecen igualmente válidas para las probabilidades de variables continuas. Por ejemplo, las propiedades sencillas (32) y (33) de los valores medios son también aplicables si u es una variable continua.

Obsérvese que las sumas que intervienen al calcular las condiciones de normalización o los valores medios pueden expresarse como integrales si la variable es continua. Por ejemplo, la condición de normalización afirma que la suma de las probabilidades extendida a todos los valores posibles de la variable debe ser igual a la unidad; en símbolos

$$\sum_r P(u_r) = 1 \tag{75}$$

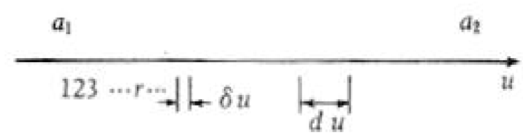


Fig. 2.13 Subdivisión del dominio de una variable continua u en un número que se puede contar de intervalos infinitesimales iguales de tamaño fijo δu. Cada intervalo de éstos se marca con un índice r que puede adquirir los valores 1, 2, 3, 4 ... Se muestra también el tamaño de un intervalo macroscópico infinitesimal du.

Sin embargo, si la variable es continua, se puede sumar primero en todos los intervalos discretos r para los que u_r está comprendida en el intervalo entre u y $u + du$; esto nos da la probabilidad $\mathcal{P}(u)\,du$ de que la variable caiga en este intervalo [20]. Se puede completar entonces la suma (75) sumando (es decir, integrando) todos los posibles intervalos du. Así pues, (75) es equivalente a

$$\int_{a_1}^{a_2} \mathcal{P}(u)\,du = 1 \tag{76}$$

que expresa la condición de normalización en función de la densidad de probabilidad $\mathcal{P}(u)$. Análogamente, la definición general (31) del valor medio de una función $f(u)$ de variables discretas viene dada por

$$\overline{f(u)} \equiv \sum_r P(u_r) f(u_r) \tag{77}$$

[20] Aquí ha de entenderse que el intervalo du es grande comparado con el intervalo arbitrariamente pequeño δu (de modo que $du \gg \delta u$), pero que es suficientemente pequeño de modo que $P(u_r)$ no varía apreciablemente dentro del intervalo du.

En una descripción continua se pueden nuevamente sumar en primer lugar todos los intervalos r en los que u_r está en el intervalo entre u y $u + du$; esto contribuye a la suma en una cantidad $\mathcal{P}(u)\,du\,f(u)$. Se puede entonces completar la misma suma integrando todos los intervalos posibles du. De aquí que (77) sea equivalente a la relación[21]

$$\overline{f(u)} = \int_{a_1}^{a_2} \mathcal{P}(u) f(u)\, du \tag{78}$$

Generalización en el caso de varias variables

La generalización de las notas anteriores en el caso de más de una variable es inmediata. Supóngase, por ejemplo, que se está tratando con dos variables continuas u y v. Entonces la probabilidad compuesta de que la variable u esté en el intervalo pequeño entre u y $u + du$ y que la variable v esté en el intervalo pequeño entre v y $v + dv$ es proporcional a la vez a du y dv y puede escribirse en la forma $\mathcal{P}(u,v)\,du\,dv$, en donde $\mathcal{P}(u,v)$ es la densidad de probabilidad independiente del tamaño de du y dv. Si se desea, la situación puede reducirse de nuevo a un caso de probabilidades discretas subdividiendo la variable u en intervalos fijos muy pequeños δu, cada uno de los cuales pueden señalarse mediante algún índice r y subdividiendo las variables v en intervalos fijos muy pequeños δv, cada uno de los cuales puede especificarse con algún otro índice s. En estos términos, la situación puede describirse entonces especificando la probabilidad P_{rs} de que las variables tengan valores comprendidos en una celdilla especificada con el par de índices r y s.

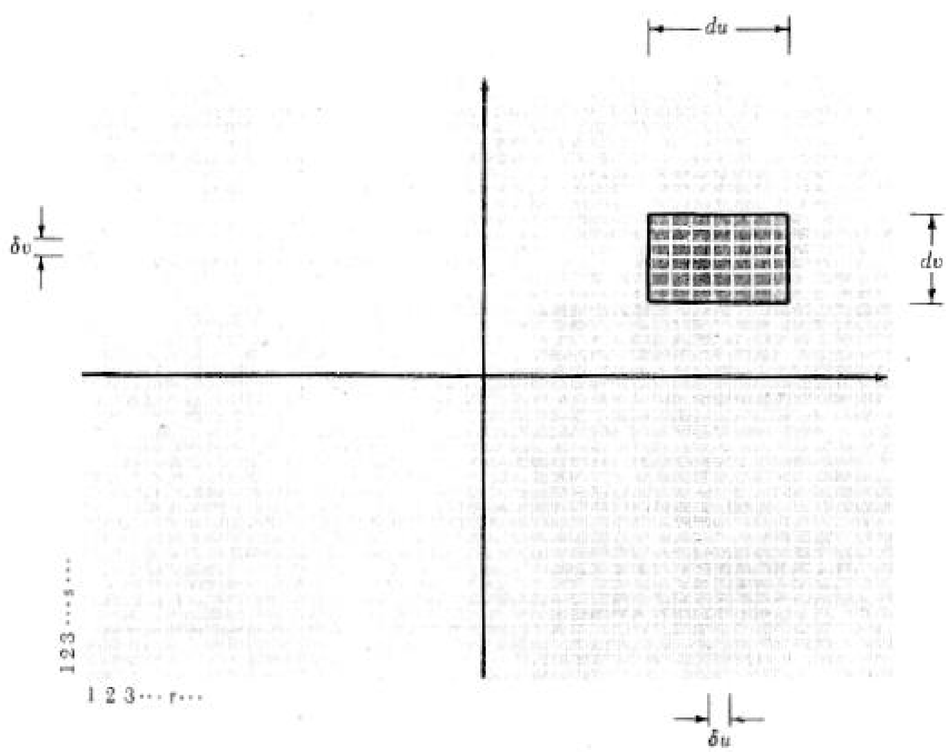


Fig. 2.14 Subdivisión de las variables continuas u y v en intervalos pequeños iguales de magnitud δu y δv señalados con las letras r y s, respectivamente. El plano uv está así subdividido en pequeñas celdas, que están marcadas con el par de índices r y s.

[21] Obsérvese que la *densidad de probabilidad* $\mathcal{P}(u)$ puede llegar a ser infinita para ciertos valores de u. Esto no supone ninguna dificultad, en tanto que cualquier integral $\int_{c_2}^{c_1} \mathcal{P}(u)\,du$ (que da la probabilidad de que el valor u esté en un intervalo arbitrario entre c_1 y c_2) permanezca finita.

Resumen de definiciones

Conjunto estadístico Conjunto compuesto por un número muy grande de sistemas mutuamente no interactivos que satisfacen las mismas condiciones que el sistema particular en consideración.

Conjunto independiente del tiempo Conjunto en el que el número de sistemas que presentan una propiedad particular cualquiera no varía con el tiempo.

Suceso. Resultado de un experimento u observación.

Probabilidad La probabilidad P_r de que se verifique un suceso r en un sistema se define respecto a un conjunto estadístico de $\mathcal{N}$ de dichos sistemas. Si $\mathcal{N}_r$ sistemas del conjunto presentan el suceso r, entonces

$$P_r \equiv \frac{\mathcal{N}_r}{\mathcal{N}} \text{ (en donde } \mathcal{N} \to \infty)$$

Independencia estadística Dos sucesos son estadísticamente independientes si la pesencia de uno de ellos no depende de la presencia o ausencia del otro.

Valor medio (o media del conjunto) El valor medio de u se designa por $\bar{u}$ y se define como

$$\bar{u} \equiv \sum_r P_r u_r$$

en donde la suma se extiende a todos los valores posibles u_r de la variable u y en donde P_r designa la probabilidad de que se presente el valor particular u_r.

Dispersión o varianza La dispersión de u se define como

$$\overline{(\Delta u)^2} \equiv \sum_r P_r(u_r - \bar{u})^2$$

Desviación standard La desviación standard de u se define como

$$\underset{\sim}{\Delta} u \equiv [\overline{(\Delta u)^2}]^{1/2}$$

Densidad de probabilidad La densidad de probabilidad $\mathcal{P}(u)$ se define por la propiedad de que $\mathcal{P}(u)\,du$ nos da la probabilidad de hallar la variable continua u en el intervalo comprendido entre u y $u + du$.

Relaciones importantes

Dados N sucesos estadísticamente independientes que tienen la probabilidad p de ocurrencia (y la probabilidad $q = 1 - p$ de no ocurrencia).

Probabilidad de que se presenten n de los N sucesos (distribución binómica):

$$P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \qquad \text{(i)}$$

Número de sucesos ocurridos: $\bar{n} = Np$ (ii)

Desviación standard de n: $\underset{\sim}{\Delta} n = \sqrt{Npq}$ (iii)

Sugerencias para lecturas complementarias

W. Weaver, *Lady Luck* (Anchor Books, Doubleday and Company, Inc., Garden City, N. Y., 1963). Una introducción elemental a los conceptos de probabilidad.

F. Mosteller, R. E. K. Rourke, y G. B. Thomas, *Probability and Statistics* (Addison-Wesley Publishing Company, Reading, Mass., 1961).

F. Reif, *Fundamentals of Statistical and Thermal Physics*, cap .1 (McGraw-Hill Book, Co., N. Y., 1965). El problema del camino aleatorio estudiado en este texto es análogo al problema del sistema ideal de spines, pero se trata con mayor extensión.

H. D. Young *Statistical Treatment of Experimental Data* (McGraw-Hill Book Company, N. Y., 1962). Estudio elemental de los métodos estadísticos, particularmente cuando se aplican a problemas de medidas experimentales.

W. Feller, *An Introduction to Probability Theory and its Applications*, 2.ª edición. (John Wiley and Sons, N. Y., 1959). Este texto sobre teoría de la probabilidad es más avanzado que los anteriores, pero estudia muchos ejemplos concretos.

Problemas

2.1 Problema elemental de dados

¿Cuál es la probabilidad de obtener un total de 6 puntos o menos con 3 dados?

2.2 Números aleatorios

Se escoge al azar un número entre 0 y 1. ¿Cuál es la probabilidad de que exactamente 5 de sus 10 primeros números decimales sean dígitos menores de 5?

2.3 Lanzamiento de dados

Suponer que las caras de un dado tienen la misma probabilidad de salir hacia arriba. Considérese un juego en el que intervienen 5 dados. Hallar la probabilidad de que salga el número "6",

a) en un solo dado

b) en un dado por lo menos

c) en dos dados precisamente.

2.4 Probabilidad de supervivencia

En el macabro juego de la *ruleta rusa* (*no* recomendado por el autor), se introduce un solo cartucho en una de las seis recámaras del cilindro de un revólver, dejando las otras cinco vacías. Se hace girar el cilindro, se apunta a la cabeza y se aprieta el gatillo. ¿Cuál es la probabilidad de estar vivo después de repetir el juego

a) una vez

b) dos veces

c) N veces?

d) ¿Cuál es la probabilidad de que se verifique el disparo cuando se repite el juego la Nsima vez?

2.5 Problema del camino aleatorio

Un hombre parte desde una farola en la mitad de una calle con pasos de igual longitud l. La probabilidad de que uno de sus pasos sea hacia la derecha es p, y $q = 1 - p$ de que sea hacia la izquierda. El hombre está tan borracho que su comportamiento en cada paso indica que no existen trazas en su memoria de lo que hizo en los pasos anteriores. Sus pasos son, pues, estadísticamente independientes. Suponer que el hombre ha dado N pasos.

a) ¿Cuál es la probabilidad $P(n)$ de que n de dichos pasos sean hacia la derecha y los restantes $n' = (N - n)$ hacia la izquierda?

b) ¿Cuál es la probabilidad $P'(m)$ de que el desplazamiento del hombre desde la farola sea igual a $m\,l$, en donde $m = n - n'$ es un número entero?

2.6 Probabilidad de volver al punto de partida

En el problema anterior suponer que $p = q$, de modo que cada paso tiene la misma probabilidad de ser hacia la derecha o hacia la izquierda. ¿Cuál es la probabilidad de que el hombre vuelva a la farola después de N pasos

a) si N es par

b) si N es impar?

2.7 Difusión unidimensional de un átomo

Considérese un alambre delgado de cobre estirado a lo largo del eje x. Un número pequeño de átomos de cobre, situados cerca de $x = 0$, se hacen radiactivos por bombardeo con partículas rápidas. Cuando se eleva la temperatura del alambre, los átomos se mueven con mayor facilidad. Cada átomo puede saltar entonces al nudo de la red adyacente situado a su izquierda (en el sentido $-x$) o bien al situado a su derecha (en el sentido $+x$). Los nudos de la red están separados una distancia l. Suponer que se ha de esperar un tiempo τ antes de que un átomo determinado salte a un nudo adyacente. Este tiempo τ es función rápidamente creciente de la temperatura absoluta del alambre. El proceso en virtud del cual los átomos se mueven mediante saltos sucesivos a los nudos adyacentes se conoce como *difusión*.

Suponer que la temperatura del alambre crece hasta alcanzar un valor elevado en un instante $t = 0$ y a continuación se mantiene a dicha temperatura.

a) Llamemos $\mathcal{P}(x)\,dx$ a la probabilidad de que un átomo radiactivo se halle a una distancia comprendida entre x y $x + dx$ después de un tiempo t. (Supondremos que $t \gg \tau$ para todos los tiempos t de interés físico, puesto que τ es muy corto cuando la temperatura del alambre es elevada.) Hacer un esquema que muestre el comportamiento de $\mathcal{P}(x)$ como función de x en los tres casos siguientes:

1) inmediatamente después del tiempo $t = 0$;

2) después de transcurrido un tiempo moderadamente largo;

3) después de transcurrido un tiempo extraordinariamente largo.

b) ¿Cuál es el desplazamiento medio $\bar{x}$ desde el origen de un átomo radioactivo al cabo de un tiempo t?

c) Hallar una expresión explícita para la desviación standard Δx del desplazamiento de un átomo radioactivo después de un tiempo t.

2.8 Cálculo de la dispersión

Utilizar las propiedades generales de los valores medios para demostrar que la dispersión de u puede calcularse por la relación general

$$\overline{(\Delta u)^2} \equiv \overline{(u - \bar{u})^2} = \overline{u^2} - \bar{u}^2 \qquad \text{(i)}$$

La última expresión de la igualdad proporciona un método sencillo para calcular la dispersión.

Demostrar también que (i) implica la desigualdad general

$$\overline{u^2} \geq \bar{u}^2 \tag{ii}$$

2.9 Valores medios para un spin aislado

El momento magnético de un spin $\frac{1}{2}$ es tal que su componente μ en sentido hacia arriba tiene la probabilidad p de ser igual a μ_0 y la probabilidad $q = 1 - p$ de ser igual a $-\mu_0$.

a) Calcular $\bar{\mu}$ y $\overline{\mu^2}$.

b) Utilizar la expresión (i) del Prob. 2.8 para calcular $\overline{(\Delta\mu)^2}$. Demostrar que el resultado está de acuerdo con la ecuación (57) del texto.

2.10 La desigualdad $\overline{u^2} \geq \bar{u}^2$

Suponer que la variable u puede adquirir los valores posibles u_r con probabilidades respectivas P_r.

a) Utilizando las definiciones de $\bar{u}$ y $\overline{u^2}$ y recordando el requisito de normalización de que $\sum_r P_r = 1$, demostrar que

$$\overline{u^2} - \bar{u}^2 = \frac{1}{2} \sum_r \sum_s P_r P_s (u_r - u_s)^2 \tag{i}$$

en donde cada suma se extiende a todos los valores posibles de la variable u.

b) Como ningún término de la suma (i) puede ser nunca negativo, demostrar que

$$\overline{u^2} \geq \bar{u}^2 \tag{ii}$$

en donde el signo igual se aplica únicamente al caso en que sólo un valor de u se presenta con probabilidad diferente de cero. El resultado (ii) está de acuerdo con el deducido en el Prob. 2.8.

2.11 La desigualdad $(\overline{u^n})^2 \leq \overline{u^{n+1}}\,\overline{u^{n-1}}$

El resultado (i) del problema anterior sugiere una generalización inmediata. Consideremos, pues, la expresión

$$\sum_r \sum_s P_r P_s u_r^m u_s^m (u_r - u_s)^2 \tag{i}$$

siendo m un entero cualquiera. Cuando m es par, esta expresión no puede ser nunca negativa; cuando m es impar, nunca puede ser negativa si los valores posibles de u son todos no negativos (o todos no positivos).

a) Realizando las multiplicaciones indicadas en (i) demostrar el siguiente resultado

$$(\overline{u^n})^2 \leq \overline{u^{n+1}}\,\overline{u^{n-1}} \tag{ii}$$

en donde $n \equiv m + 1$. Si n es impar, esta desigualdad es siempre válida; si n es par, es válida si los valores posibles de u son todos no negativos (o no positivos). El signo igual en (ii) se aplica únicamente en el caso en que se presenta un solo valor posible de u con probabilidad no nula.

b) Demostrar que (ii) implica, como caso especial, la desigualdad

$$\overline{\left(\frac{1}{u}\right)} \geq \frac{1}{\bar{u}} \tag{iii}$$

válida siempre que los valores posibles de u sean todos positivos (o todos negativos). El signo igual se aplica en el caso especial en que sólo se presente un valor de u con probabilidad no nula.

2.12 Método de inversión óptima

El caso práctico siguiente demuestra cómo las diferentes maneras de promediar la misma magnitud puede conducir a resultados significativamente diferentes. Supongamos que se desea invertir dinero comprando al principio de cada mes un cierto número de acciones de una compañía. El coste C_r por acción depende, naturalmente, del mes en cuestión y varía de un mes a otro de un modo más bien impredecible. Se sugieren por sí mismos dos métodos opcionales de hacer esta inversión regular: en el método *A*, se compra el mismo número s de acciones cada mes; en el método *B*, se compran acciones por la misma cantidad de dinero m cada mes. Después de N meses, se habría adquirido un número total S de acciones, habiéndose pagado una cantidad total de dinero M. El mejor método de inversión es evidentemente el que proporcione el mayor número de acciones por la menor cantidad de dinero, es decir, el que dé la razón S/M mayor.

a) Obtener una expresión para la razón S/M en el caso del método *A*.

b) Obtener una expresión para la razón S/M en el caso del método *B*.

c) Demostrar que el método *B* de inversión es el mejor, no importa cómo fluctúe el precio de las acciones de mes en mes. [*Sugerencia*: Hacer uso de la desigualdad (iii) del problema anterior.]

2.13 Sistemas de núcleos con spin 1

Consideremos un núcleo con spin 1 (es decir, con momento cinético de spin $\hbar$). Su componente μ del momento magnético sobre una dirección dada puede tener entonces *tres* valores posibles, a saber $+\mu_0$, 0, $-\mu_0$. Suponer que el núcleo no tiene simetría esférica, sino que es de forma elipsoidal. Como consecuencia tiende a orientarse preferentemente de modo que su eje mayor sea paralelo a una dirección dada en el sólido cristalino en el que está situado el núcleo. Existe así una probabilidad p de que $\mu = \mu_0$ y una probabilidad p de que $\mu = -\mu_0$; la probabilidad de que $\mu = 0$ es igual entonces a $1 - 2p$.

a) Calcular $\bar{\mu}$ y $\overline{\mu^2}$.

b) Calcular $\overline{(\Delta\mu)^2}$.

c) Suponer que el sólido en consideración contiene N de estos núcleos que interaccionan entre sí de un modo despreciable. Llamemos M a la componente total del momento magnético a lo largo de la dirección especificada de todos estos núcleos. Calcular $\bar{M}$ y su desviación standard ΔM en función de N, p y μ_0.

2.14 Cálculo directo de $\bar{n}$ y $\overline{(\Delta n)^2}$

Consideremos un sistema ideal de N spines $\frac{1}{2}$ idénticos. El número n de momentos magnéticos que señalan hacia arriba puede escribirse entonces en la forma

$$n = u_1 + u_2 + \cdots + u_N \tag{i}$$

en donde $u_i = 1$ si el momento magnético i señala hacia arriba y $u_i = 0$ si señala hacia abajo. Utilizar la expresión (i) y el hecho de que los spines son estadísticamente independientes para establecer los resultados siguientes.

a) Demostrar que $\bar{n} = N\,\bar{u}$.

b) Demostrar que $\overline{(\Delta n)^2} = N\,\overline{(\Delta u)^2}$.

c) Suponer que un momento magnético tiene la probabilidad p de señalar hacia arriba y la probabilidad $q = 1 - p$ de señalar hacia abajo. ¿Cuánto valen $\bar{u}$ y $\overline{(\Delta u)^2}$?

d) Calcular $\bar{n}$ y $\overline{(\Delta n)^2}$ y demostrar que sus resultados están de acuerdo con las relaciones (66) y (67) halladas en el texto mediante un método menos directo.

2.15 Fluctuaciones de densidad en un gas

Consideremos un gas ideal de N moléculas que está en equilibrio dentro de un recipiente de volumen V_0. Llamemos n al número de moléculas situadas dentro de cualquier subvolumen V del recipiente. La probabilidad p de que una molécula dada esté situada dentro del volumen V viene dada entonces por $p = V/V_0$.

a) ¿Cuál es el número medio $\bar{n}$ de moléculas situadas dentro de V? Expresar la respuesta en función de N, V_0 y V.

b) Hallar la desviación standard Δn en el número de moléculas situadas dentro del subvolumen V. De aquí, calcular $\Delta n/\bar{n}$ (expresando la respuesta en función de N, V_0 y V.

c) ¿A qué se reduce la respuesta de la parte (*b*) cuando $V \ll V_0$?

d) ¿Qué valor adquirirá la desviación standard Δn cuando $V \to V_0$? ¿Está de acuerdo la respuesta a la parte (*b*) con esta predicción?

2.16 Efecto "perdigón"

El filamento caliente de un tubo de vacío emite al azar electrones de carga e. En buena aproximación la emisión de un electrón cualquiera no afecta la probabilidad de emisión de cualquier otro electrón. Consideremos un intervalo pequeño de tiempo Δt. Existe entonces cierta probabilidad p de que un electrón sea emitido por el filamento durante este intervalo (y una probabilidad $q = 1 - p$ de que no sea emitido). Como Δt es muy pequeño, la probabilidad p de emisión durante este intervalo de tiempo es muy pequeña (es decir, $p \ll 1$) y la probabilidad de que se emita más de un electrón durante este tiempo es despreciable.

Consideremos un tiempo cualquiera t que sea mucho mayor que Δt. En este tiempo existen $N = t/\Delta t$ intervalos Δt posibles durante los cuales puede emitirse un electrón. La carga total emitida en el tiempo t puede escribirse en la forma

$$Q = q_1 + q_2 + q_3 + \cdots + q_N$$

en donde q_i designa la carga que se emite durante el intervalo i de tiempo Δt; así pues, $q_i = e$ si se emite un electrón, y $q_i = 0$ si no se emite.

a) ¿Cuál es la carga media $\bar{Q}$ emitida por el filamento durante el tiempo t?

b) ¿Cuál es la dispersión $\overline{(\Delta Q)^2}$ emitida por el filamento durante el tiempo t? Hacer uso del hecho de que $p \ll 1$ para simplificar la respuesta a esta cuestión.

c) La corriente I emitida durante el tiempo t es Q/t. Relacionar la dispersión $\overline{(\Delta I)^2}$ de la corriente con su valor medio $\bar{I}$, demostrando que

$$\overline{(\Delta I)^2} = \frac{e}{t}\bar{I}$$

d) El hecho de que la corriente medida durante cualquier intervalo t presente fluctuaciones (que son más pronunciadas cuanto más corto sea el intervalo de tiempo, es decir, cuanto menor sea el número total de electrones individuales que intervienen en el proceso de emisión) se denomina "*efecto perdigón*". Calcular la desviación standard ΔI de la corriente si su valor medio es $I = 1$ microampere y el tiempo de medida es 1 segundo.

2.17 Cálculo de un valor cuadrático medio

Una batería de fem total V está conectada a una resistencia R. Como resultado, se disipa en la misma una cantidad de potencia $P = V^2/R$. La batería se compone de N células individuales conectadas en serie de modo que V sea precisamente igual a la suma de las fem de todas ellas. La batería es algo usada, de modo que no todos sus componentes están en perfectas condiciones. Así pues, existe únicamente una probabilidad p de que la fem de cualquier célula individual tenga su valor normal v; y una probabilidad $(1 - p)$ que la fem de cualquier célula individual sea cero a causa de un cortocircuito interno. Las células individuales son estadísticamente independientes entre sí. Bajo estas condiciones calcular la potencia *media* $\bar{P}$ disipada en la resistencia. Expresar el resultado en función de N, v, p y R.

2.18 Estimación del error de una medición

Un hombre intenta determinar una distancia de 50 metros colocando una regla métrica 50 veces sucesivamente un extremo junto a otro. Este procedimiento va acompañado necesariamente por algún error. Así pues, el hombre no puede garantizar una distancia precisamente de un metro, entre las dos marcas de tiza que ha hecho cada vez que coloca la regla sobre el suelo. Sabe, sin embargo, que la distancia entre las dos marcas es igualmente probable que caiga entre 99,8 y 100,2 cm y que ciertamente no se saldría de estos límites. Después de repetir la operación 50 veces, el hombre habrá medido ciertamente una distancia media de 50 metros. Para estimar su error total calcular la desviación standard de su distancia medida.

✗2.19 Difusión de una molécula en un gas

Una molécula en un gas está libre para moverse en tres dimensiones. Llamemos $\mathbf{s}$ a su desplazamiento entre choques sucesivos con otras moléculas. Estos desplazamientos de la molécula, son, con buena aproximación, estadísticamente independientes. Además, como no existe ninguna dirección preferida en el espacio, una molécula tiene la misma probabilidad de moverse en una dirección o en la opuesta. Así pues, su desplazamiento medio será $\bar{\mathbf{s}} = 0$ (es decir, cada componente de este desplazamiento se anula en valor medio de modo que $\bar{s}_x = \bar{s}_y = \bar{s}_z = 0$).

El desplazamiento total $\mathbf{R}$ de la molécula después de N desplazamientos sucesivos puede escribirse entonces en la forma

$$\mathbf{R} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{s}_3 + \cdots + \mathbf{s}_N$$

i-ésimo

en donde $\mathbf{s}_i$ designa el desplazamiento de la molécula. Utilizar un razonamiento semejante al de la Sección 2.5 para contestar a las cuestiones siguientes:

a) ¿Cuál es el desplazamiento medio $\overline{\mathbf{R}}$ de la molécula después de N desplazamientos?

b) ¿Cuál es la desviación standard $\Delta R \equiv \left[\overline{(\mathbf{R} - \overline{\mathbf{R}})^2}\right]^{1/2}$ de este desplazamiento después de N choques? En particular, ¿cuál es ΔR si el *valor* de cada desplazamiento s tiene la misma longitud?

2.20 Distribución de desplazamientos de osciladores aleatorios

El desplazamiento x de un oscilador armónico simple en función del tiempo t viene dado por

$$x = A \cos (\omega t + \varphi)$$

en donde ω es la frecuencia angular del oscilador; A es su amplitud de oscilación y φ una constante arbitraria que puede tener cualquier valor en el intervalo $0 \leq \varphi \leq 2\pi$. Suponer que se observa un conjunto de osciladores que tienen todos la misma frecuencia ω y amplitud A, pero que tienen relaciones de fase al azar de modo que la probabilidad de que φ caiga en el intervalo entre φ y $\varphi + d\varphi$ viene dada simplemente por $d\varphi/2\pi$. Hallar la probabilidad $\mathcal{P}(x)\,dx$ de que el desplazamiento de un oscilador en cualquier instante dado t, se encuentre comprendido en el intervalo entre x y $x + dx$.