1 Combinatoria

En muchos de los ejemplos que hemos visto hasta ahora, ha sido necesario determinar el número de maneras en que un evento puede ocurrir, así, conociendo la probabilidad de cada una de esas ocurrencias y los axiomas de la función probabilidad podemos calcular la probabilidad del evento.

Existen muchas técnicas, basadas en la ley de probabilidad total para contar todas las posibilidades. Es usual describir este problema en términos de cómo se pueden elegir `bolas distinguibles` de una caja asumiendo que la elección es `aleatoria`, es decir, cada bola en la caja es igualmente probable que se elija.

Alternativamente, el problema puede ser expresado en términos de cómo se pueden insertar `bolas indistinguibles` en cajas distinguibles.

Esta primera forma se llama **problema de selección** y la última forma se llama **problema de asignación**. Formularemos nuestra notas en términos del problema de selección.

En el contexto del problema de selección, una vez que se ha elegido una de las bolas distinguibles de la caja y antes de la selección de la bola siguiente, se debe tomar una decisión sobre qué se debe hacerse con la primera bola elegida. Si se vuelve a colocar en la caja, se dice que la selección se hace con reemplazo [1]. En este caso, la misma bola puede ser elegida una segunda vez, luego una tercera vez, y una cuarta vez, y así sucesivamente. La segunda posibilidad es que la bola sea puesta a un lado y nunca vuelva a la caja. En este caso, la selección se hace sin reemplazo [2].

Un último punto se refiere al `orden` en que se seleccionan las bolas. En algunos casos, este orden es importante. Por ejemplo, puede ser necesario saber si la bola negra fue elegida antes o después de la bola blanca. Cuando el orden es importante, el término es de **permutación**. En otros casos, todo lo que se necesita saber es que se eligió una bola negra y una bola blanca y no el orden en que fueron elegidos. Esto se conoce como una **combinación**.

1.1 Principio del conteo

Ocasionalmente, se encuentran espacios muestrales para los cuales los puntos del espacio muestral son igualmente probables. Si este es el caso y si el espacio muestral S contiene n puntos, entonces, puesto que la probabilidad total en el espacio muestral es 1, cada punto tiene una probabilidad $1/n$. Si denotamos los puntos mutuamente exclusivos en A por $a_i, i = 1, 2, 3, \ldots, n$, entonces la probabilidad de un evento A, es la suma de las probabilidades de los puntos del espacio muestral en A. Es decir

$$\mathbb{P}(A) = \sum_{a_i \in A} \mathbb{P}(a_i) = \sum_{a_i \in A} \frac{1}{n}$$

así

$$\mathbb{P}(A) = \frac{\text{Número de puntos de A}}{n} = \frac{\text{Número de puntos de A}}{\text{Número de puntos de S}}$$

Con el fin de considerar problemas que conducen a espacios muestrales con puntos igualmente probables, debemos considerar algunas técnicas para contar conjuntos de puntos. Estas técnicas proporcionan algunos problemas desafiantes.

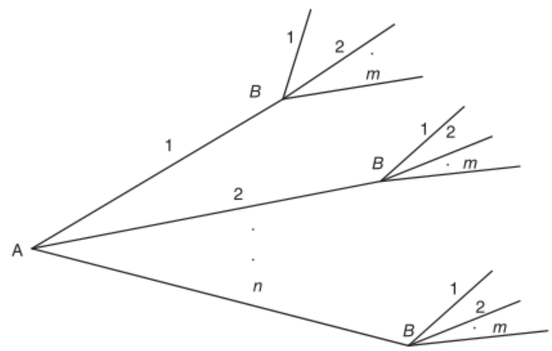
Se advierte que se debe tener cuidado al concluir que sólo porque un espacio muestral tiene n puntos, cada punto tiene probabilidad $1/n$. Por ejemplo, un viaje en avión se completa de forma segura o no, se

[1] with replacement
[2] without replacement

hace evidente que estos eventos no tienen cada uno una probabilidad 1/2! Las técnicas de conteo consideradas en este caso se basan en dos principios de conteo fundamentales relativos a eventos mutuamente exclusivos A y B

- Principio 1 Si los eventos A y B pueden ocurrir en n y m maneras respectivamente, entonces A y B pueden ocurrir juntos de $n \cdot m$ maneras.
- Principio 2 Si los eventos A y B pueden ocurrir en n y m maneras respectivamente, entonces A o B pueden ocurrir juntos de $n + m$ maneras.

El Principio 1 se establece fácilmente ya que A puede ocurrir de n maneras y luego debe ser seguida por cada manera en que B puede ocurrir. Un diagrama de árbol, mostrado en la siguiente figura ilustra el resultado. El Principio 2 simplemente usa la palabra o en un sentido exclusivo.



1.2 Permutaciones

Una disposición (arreglo) de ítems, también llamada una secuencia ordenada de ítems , se dice que es una permutación de los ítems. Una secuencia ordenada de n ítems se llama una n-permutación.

Con n ítems distintos, hay $n!$ permutaciones posibles. Este es el número de diferentes maneras en que los n ítems distintos pueden ser colocados, ordenados .

Ejemplo 1.1 Si se tienen tres letras diferentes A, B y C hay $3! = 6$ permutaciones (arreglos, disposiciones), a saber

$$ABC,\ ACB,\ BAC,\ BCA,\ CAB,\ CBA$$

Debes observar que no sólo hay tres elementos (items) distintos A, B y C, sino que también hay tres lugares diferentes en los que pueden colocarse, es decir, la primera, segunda y tercera posición de una cadena de tres letras. Por esta razón, se considera que ABC es una permutación diferente de ACB.

Ahora consideremos el número de permutaciones que pueden generarse a partir de elementos que no son todos distintos.

Ejemplo 1.2 Considere la palabra OXO. Esta vez no todas las letras son distintas. Hay dos O y una X y sólo se pueden encontrar tres arreglos diferentes, a saber

$$OXO,\ XOO,\ OOX$$

Puesto que el carácter O aparece dos veces, el número $n! = 3!$ obtenido con $n = 3$ caracteres distintos debe ser dividido por 2!. En este caso, tenemos $3!/2! = 3$ arreglos.

Ejemplo 1.3 Considere la palabra PUPPY. Si los cinco caracteres fueran distintos, el número de permutaciones que se obtendría seria 5!. Sin embargo, como el carácter P aparece tres veces en PUPPY, el 5! debe ser dividido por 3!, para obtener $5!/3! = 5 \times 4 = 20$. El número de diferentes maneras en que los caracteres de la palabra PUPPY se pueden colocar o disponer es por lo tanto 20.

Ejemplo 1.4 Consideremos la palabra NEEDED, que contiene tres E, dos D y una N. El número de diferentes maneras en que los caracteres en esta palabra puede ser dispuestos es

$$\frac{6!}{2! \times 3!} = 60$$

Por lo tanto, el número de permutaciones que se pueden obtener de n ítems, no todos los cuales son distintos, puede obtenerse asumiendo primero que son distintos (lo que nos da $n!$) y luego, para cada elemento múltiple se divide por el factorial de su multiplicidad.

1.3 Permutaciones con reemplazos

En el caso de permutaciones con reemplazo, nuestro objetivo es contar el número de formas en que se pueden seleccionar k bolas entre n bolas distinguibles. Después de cada bola es elegida y sus características son registradas, se sustituye en la caja y se elige la bola siguiente.

Una forma alternativa de afirmar esto es decir que después de cada bola es seleccionada, la bola es puesta a un lado y una bola que es idéntica a ella toma su lugar en la caja. De esta manera es posible que la misma bola sea elegida muchas veces.

Si $k = 1$(se elige una bola), entonces el número de permutaciones posibles es n, ya que cualquiera de las n bolas puede ser elegida. Si $k = 2$, entonces cualquiera de las n bolas puede ser elegida como la primera bola y luego reemplazada en la caja. Para cada una de estas n elecciones de la primera bola, la siguiente bola elegida puede ser también cualquiera de las n bolas distinguibles y por lo tanto el número de permutaciones cuando $k = 2$ es $n \times n = n^2$. Este razonamiento puede continuar y demostrar que el número de permutaciones obtenidas para cualquier k es n^k.

Ejemplo 1.5 El número de códigos de cuatro dígitos que se pueden obtener utilizando el sistema de números decimales (con diez dígitos de 0 a 9 inclusive) es $10^4 = 10.000$. Estos son los códigos que van desde 0000 hasta 9999.

Supongamos ahora que hay n_1 maneras de elegir un primer elemento (ítem) y n_2 formas de elegir un segundo elemento(ítem). Entonces el número de pares ordenados distintos es igual a $n_1 n_2$.

En términos de bolas distinguibles en cajas, esto puede ser visto como el número de maneras en que una bola puede ser elegida de una primera caja que contiene n_1 bolas distinguibles y una segunda bola elegida de una segunda caja que contiene n_2 bolas distinguibles. La extensión a más de dos ítems diferentes es inmediata.

Si hay n_i formas de elegir un ítem i, para $i = 1, 2, \ldots, k$, entonces el número de k-tuplas ordenadas distintas es igual a $n_1 n_2 \ldots n_k$.

Ejemplo 1.6 Supongamos que una camisa se puede elegir entre $n_1 = 12$ camisas diferentes y una corbata de $n_2 = 20$ corbatas diferentes. Entonces el número de combinaciones de camisas y corbatas es $n_1 n_2 = 240$.

1.4 Permutaciones sin reemplazos

Pasemos ahora al problema de contar el número de diferentes permutaciones obtenidas al seleccionar k bolas de entre n bolas distinguibles en el caso de que una vez que se elija una bola en particular,esta no se devuelve a la caja. Esto es igual al número de secuencias ordenadas de k objetos distinguibles, es decir, el número de k-permutaciones, que se pueden obtener a partir de n objetos distinguibles.

Denotamos a este número como $P(n,k)$ y asignamos los valores $P(n,0) = 1, n = 0,1,\ldots$, por convención.

Decimos que $P(n,k)$ es el número de permutaciones de n objetos tomados k a la vez.

Si $k = 1$, entonces cualquier bola puede ser elegida y el número total de permutaciones simplemente n. Si $k = 2$, entonces cualquiera de las n bolas puede ser elegida como la primera, para cada una de estas n elecciones posibles, quedan $n - 1$ bolas en la caja, cualquiera de las cuales puede ser elegida como la segunda. Así, el número total de permutaciones para $k = 2$ es igual a $n(n-1)$. Con k = 3, hay n posibilidades diferentes para la primera bola, $n - 1$ para la segunda y $n - 2$ para la tercero, lo que da $P(n,3) = n(n-1)(n-2)$.

Podemos ahora generalizar esto a cualquier arbitrario $k \leq n$ para obtener

$$P(n,k) = n(n-1)(n-2)\ldots(n-k+1) = \frac{n!}{(n-k)!}.$$

Existen varias relaciones entre los valores de $P(n,k)$ para diferentes valores de los parámetros n y k.

Por ejemplo,

$$P(n,k) = nP(n-1,k-1), \quad k = 1,2,\ldots,n, \quad n = 1,2,\ldots$$

Podemos explicar esta fórmula de la siguiente manera: La primera bola elegida puede ser cualquiera de las n bolas distinguibles. Esto deja $n-1$ bolas a partir de las cuales $k-1$ debe ser elegidas, el número de maneras en que $k-1$ bolas se pueden elegir de $n-1$ es igual a $P(n-1,k-1)$.

Ejemplo 1.7 Sea $n = 4, k = 3$ y distinguimos las cuatro bolas por medio de las letras A, B, C y D. Tenemos $P(4,3) = 4 \times 3 \times 2 = 24$.

Esas 24 posibilidades, están dadas por

```
ABC, ABD, ACB, ACD, ADB, ADC,
BAC, BAD, BCA, BCD, BDA, BDC,
CAB, CAD, CBA, CBD, CDA, CDB,
DAB, DAC, DBA, DBC, DCA, DCB.
```

Debes observar que distinguimos entre ABC y ACB, ya que el orden es importante, pero no incluimos la misma letra más de una vez (por ejemplo, AAB no está presente), esto es por que una vez que las letras son usadas no pueden ser utilizadas de nuevo.

Ejemplo 1.8 Supongamos que deseamos encontrar el número de formas en que un código de cuatro dígitos, en el que todos los dígitos son diferentes, puede ser seleccionado. Dado que trabajamos con diez dígitos (0 a 9), el primer dígito puede ser cualquiera de los diez, el segundo cualquiera de los nueve restantes, y así sucesivamente. Continuando de esta manera, vemos que el número total de permutaciones es $10 \times 9 \times 8 \times 7$ que naturalmente es el mismo valor cuando se calcula directamente con la fórmula

$$P(10,4) = \frac{10!}{(10-4)!} = 5040.$$

Los problemas como estos se formulan frecuentemente como preguntas de probabilidad,por ejemplo, para encontrar la probabilidad de que un código de cuatro dígitos generado al azar tenga todos los dígitos diferentes. Como acabamos de ver, el número total de códigos de cuatro dígitos que tienen todos los dígitos diferentes es 5.040, mientras que el número total de números de cuatro dígitos es 10.000. Si se supone que todas las opciones de código de cuatro dígitos son igualmente probables, entonces la probabilidad de obtener uno en el que todos los dígitos son diferentes es 5.040/10.000 o aproximadamente 0.5.

Tales preguntas nos obligan a encontrar el tamaño del espacio muestral de permutaciones sin reemplazo y compararlo con el tamaño del espacio de muestra con reemplazo. Obviamente, esto sólo funciona correctamente cuando todas las opciones son igualmente probables (selección aleatoria).

Ejemplo 1.9 Si los passwords pueden consistir de 6 letras, encuentra la probabilidad que aleatoriamente un password, pueda ser elegido, que no tenga letras repetidas.

En este caso el número posible de passwords es 26^6 y el número casos favorables es $\binom{26}{6}$ y la probabilidad de escoger palabras no repetidas es $\frac{\binom{26}{6}}{26^6}$ y esto cálculo puede hacerse usando R

```
> prod(26:21)/26^6

[1] 0.5366045
```

Ejemplo 1.10 Supongamos que r palomas que regresan de una carrera tienen la misma probabilidad de entrar en cualquiera de n nidos de acogida, donde $n > r$. Nos gustaría saber la probabilidad de que todas las palomas terminen en nidos diferentes?.

Sea k_1 el nido donde entra la primera paloma, k_2 donde entra la segunda paloma, y así sucesivamente. Tenemos $1 \leq k_i \leq n$, para $i = 1, 2, \ldots, r$. El número total de posibilidades disponibles para las palomas es igual a $N = n^r$, ya que la paloma número 1 puede volar a cualquiera de los n nidos y lo mismo para la paloma número 2 y así sucesivamente. Esta es la situación de las permutaciones con reemplazo.

Decimos que cada una de estas N opciones posibles es igualmente probable. Por lo tanto, existen n^r disposiciones distintas y equiprobables de las palomas en los n nidos. Estos n^r eventos son los eventos elementales que constituyen el espacio muestral del experimento.

Sea A el evento en que todas las palomas terminan en nidos diferentes. Este evento ocurre si todos los $k_i, i = 1, 2, \ldots, k_r$ son distintos: cada paloma vuela en un nido diferente. Esta es la situación de la permutación sin reemplazo.

La primera paloma puede entrar en cualquiera de los n nidos, la segunda puede entrar en cualquiera de los $n - 1$ nidos restantes y así sucesivamente. El número de posibles formas en que las r palomas se pueden disponer en los n nidos de manera que ninguno comparte un nido es dado por $n(n-1)\ldots(n-r+1)$. Este es el número de salidas que resultan en el evento A y dado que hay un total de n^r resultados equiprobables en el espacio muestral, la probabilidad de A debe ser

$$\mathbb{P}(A) = \frac{n(n-1)\cdots(n-r+1)}{n^r} = \frac{n!/(n-r)!}{n^r} = \frac{\text{conteo sin reemplazo}}{\text{conteo con reemplazo}}.$$

Ejemplo 1.11 Considere el `problema del cumpleaños`, el de determinar la probabilidad de que entre un grupo de k personas al menos dos tengan el mismo cumpleaños (dia y mes solamente).

Para encontrar la probabilidad de que al menos dos personas tengan el mismo cumpleaños, es más fácil calcular el complemento de este evento, la probabilidad de que no haya dos individuos que tengan el mismo cumpleaños. Si asumimos que un año tiene 365 dias, entonces el número de permutaciones con reemplazo es de 365^k . Este es el tamaõ del espacio muestral desde el cual se toman todas las posibles permutaciones de cumpleaños. El número de permutaciones sin reemplazo es $365!/(365-k)!$. Este es el número de salidas en el espacio muestral en el que no hay dos cumpleaños en la permutación que sean los mismos. La probabilidad de no encontrar entre k personas dos con el mismo cumpleaños esta dado por la siguiente proporción entonces

$$
\begin{aligned}
\frac{365!/(365-k)!}{365^k} &= \frac{365}{365}\times\frac{364}{365}\times\frac{363}{365}\times\cdots\times\frac{365-k+1}{365} \\
&= \left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\cdots\left(1-\frac{k-1}{365}\right)
\end{aligned}
$$

Se puede notar que en una clase con 23 estudiantes, la probabilidad de que dos estudiantes no tengan el mismo cumpleaños es menos del 50%.

Ejemplo 1.12 Resolvamos los siguiente

- En una fiesta de cinco estudiantes, calcular la probabilidad, que al menos dos tengan el mismo (dia/mes) de cumpleaños, asumiendo que el año tiene 365 dias.
- La probabilidad que dos estudiantes en una clase tengan el mismo cumpleaños es al menos el 75%. ¿ Cuál es el mínimo tamaño de la clase?.

En efecto

$$
\begin{aligned}
\text{P (todos tiene diferentes cumpleaños)} &= \frac{\binom{36}{5}}{365^5}, \\
\text{P(Al menos dos tiene el mismo cumpleaños)} &= 1-\frac{\binom{36}{5}}{365^5}.
\end{aligned}
$$

En R

```
> k <- 5
> prod(365:(365-k+1))/365^k     # todos diferentes

[1] 0.9728644

> 1 -prod(365:(365-k+1))/365^k  # al menos dos cumplen el mismo dia

[1] 0.02713557
```

P(al menos 2 en el mismo k) $= 1-\frac{\binom{365}{k}}{365^k}$. Escogemos un k tal que $1-\frac{\binom{365}{k}}{365^k} \geq 0.75$. En efecto k se encuentra entre 30 y 40. Podemos inferir que k está muy cerca de 30.

```
> k<-30
> 1-prod(365:(365-k+1))/365^k  # al menos dos cumplen el mismo dia

[1] 0.7063162

> k<-31
> 1-prod(365:(365-k+1))/365^k  # al menos dos cumplen el mismo dia

[1] 0.7304546

> k<-32
> 1-prod(365:(365-k+1))/365^k  # al menos dos cumplen el mismo dia

[1] 0.7533475
```

El mínimo k es 32. Es necesario una clase de 32 o más para estar asegurar al 75%, de que dos estudiantes tienen el mismo dia de cumpleaños.

1.5 Combinaciones sin reemplazos

Sea al caso en el que las bolas se seleccionan sin reemplazo, pero el orden en que se seleccionan no tiene importancia. Lo único que importa es que se han elegido algunas bolas y el punto en el que se seleccionaron no es de interés.

Por ejemplo, podemos haber elegido una bola verde, una bola roja y una bola negra, pero no sabemos o nos preocupamos cuál de estas tres fue elegida primera, segunda o tercera. Sea $C(n,k)$ que denota el número de maneras en que se pueden seleccionar k bolas sin reemplazo e independientemente de su orden de una caja que contiene n bolas. Este símbolo se dice que es el número de combinaciones de n elementos tomados k a la vez no teniendo en cuenta el orden.

Como vimos anteriormente, cualquier colección de k ítems distintos puede ser colocada en $k!$ diferentes permutaciones. Por ejemplo, con $k = 3$ y usando las letras ABC, tenemos $3! = 6$ permutaciones.

```
ABC   ACB   BAC   BCA   CAB  CBA
```

Estas son todas diferentes permutaciones porque el orden es importante. Si el orden no es importante, entonces la única información que necesitamos es que haya un A, un B y un C. Todas las 3! permutaciones producen entonces una sola combinación. Esto nos proporciona un medio para relacionar permutaciones sin reemplazo a combinaciones sin reemplazo.

Dado que cualquier secuencia de k ítems puede ser dispuesta en $k!$ permutaciones, se deduce que debe haber $k!$ veces más permutaciones (sin reemplazo) de k ítems distintos de lo que hay en combinaciones (sin reemplazo), ya que el orden importa en permutaciones pero no en combinaciones. En otras palabras, debemos tener

$$k!C(n,k) = P(n,k) \text{ para } n = 0,1,2,\ldots \text{ y } k = 0,1,\ldots,n.$$

Esto conduce a

$$C(n,k) = \frac{P(n,k)}{k!} = \frac{n!}{k!(n-k)!} \text{ para } n = 0,1,2,\ldots \text{ y } k = 0,1,\ldots,n.$$

Ejemplo 1.13 Sean A, B, C, D y E cinco elementos distinguibles de los cuales dos deben ser escogidos (lo cual implica, sin reemplazo, y no importa cuál viene primero: es decir, una combinación). El número de formas en que estos dos pueden ser elegidos está dado por $C(5,2) = 5!/(2!3!) = 10$. Estos son

```
AB    AC   AD   AE
      BC   BD   BE
           CD   CE
                DE
```

Si las elecciones se hacen de forma aleatoria, entonces la probabilidad de obtener una de estas diez combinaciones es $1/10$. Calculemos la probabilidad de elegir exactamente uno de (A, B) y uno de (C, D, E). Este es el evento que contiene las salidad $\{AC, AD, AE, BC, BD, BE\}$. El número de formas en que tal combinación puede aparecer es el producto del número de maneras en que podemos elegir uno de dos (es decir $C(2,1)$) por el número de maneras que podemos elegir uno de los tres restantes (es decir $C(3,1)$). La probabilidad de tal elección es entonces

$$\frac{C(2,1) \times C(3,1)}{C(5,2)} = \frac{2!/(1!1!) \times 3!/(1!2!)}{5!/(2!3!)} = \frac{2 \times 3}{10} = 0.6.$$

$C(n,k)$ se denomina coeficiente binomial ya que es el coeficiente del término $p^k q^{n-k}$ en la expansión del binomio $(p+q)^n$. Formas alternativas para escribir este coeficiente binomial son

$$C(n,k) = C_k^n = \binom{n}{k}.$$

Para calcular los coeficientes binomiales, notamos que

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{k}\left(\frac{(n-1)!}{(k-1)!(n-k)!}\right)$$
$$= \frac{n}{k}\left(\frac{(n-1)!}{(k-1)!([n-1]-[k-1])!}\right) = \frac{n}{k}\binom{n-1}{k-1}$$

Lo que conduce a

$$\binom{n}{k} = \frac{n}{k} \times \frac{n-1}{k-1} \times \cdots \times \frac{n-(k-1)}{k-(k-1)} = \prod_{j=1}^{k} \frac{n-j+1}{j}.$$

y esta expresión computacionalmente es más eficiente que formar factoriales y tomar razones.

Ejemplo 1.14 Calculemos el número de maneras de escoger 5 numeros de 40 y la probabilidad de que eso ocurra.

```
> prod(40:36)/prod(5:1)

[1] 658008

> 1/choose(40,5)

[1] 1.519738e-06
```

Los coeficientes binomiales poseen un gran número de propiedades interesantes, de las cuales sólo cuatro se enumeran a continuación. Para $0 \leq j \leq k \leq n$ se cumple

- $\binom{n}{k} = \binom{n}{n-k}$.
- $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$.
- $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.
- $\binom{n}{k}\binom{k}{j} = \binom{n}{j}\binom{n-j}{k-j}$.

Consideremos ahora el problema de colocar k bolas distinguibles en n cajas diferentes de tal manera que el número de bolas en la caja i sea k_i para $i = 1, 2, \ldots, n$.

Asumimos que $\sum_{i=1}^{n} k_i = k$, de modo que cada bola se pone en una de las cajas y de manera que no queda ninguna. El número de combinaciones obtenidas al seleccionar las primeras k_1 bolas para la caja 1 es $C(k, k_1)$. Esto deja $k - k_1$ bolas, de las cuales k_2 son escogidas y colocadas en la caja 2. El número de combinaciones obtenidas al seleccionar estas k_2 bolas de las $k - k_1$ bolas es $C(k - k_1, k_2)$, de modo que el total obtenido de las dos primeras cajas es $C(k, k_1) \times C(k - k_1, k_2)$.

Esto deja ahora $k - k_1 - k_2$ de los cuales k_3 deben ser elegidos y puesto en la caja 3. Continuando de esta manera, vemos que el número total de combinaciones está dado por

$$C(k,k_1)\times C(k-k_1,k_2)\times C(k-k_1-k_2,k_3)\times\cdots\times C\left(k-\sum_{i=1}^{n-1}k_i,k_n\right),$$

donde $C(k-\sum_{i=1}^{n-1}k_i,k_n) = C(k_n,k_n)$. Sustituyendo en la fórmula para los números de combinación tenemos

$$\frac{k!}{k_1!(k-k_1)!}\times\frac{(k-k_1)!}{k_2!(k-k_1-k_2)!}\times\frac{(k-k_1-k_2)!}{k_2!(k-k_1-k_2-k_3)!}\times\cdots$$
$$=\frac{k!}{k_1!k_2!\cdots k_n!}\equiv\binom{k}{k_1,k_2,\ldots k_n}.$$

Estos son los llamados coeficientes multinomiales. Cuando $k=k_1+k_2$ tenemos

$$\binom{k}{k_1,k_2}=\binom{k}{k_1}=\binom{k}{k_2}.$$

Ejemplo 1.15 Calculemos el número de maneras en que cinco cartas pueden ser repartidas desde una baraja de 52 cartas. Puesto que no importa en qué orden se reparten las cartas, el problema es con combinaciones en lugar de permutaciones, por lo que la respuesta es $C(52,5)=2,598,960$. Sea K el evento de que hay exactamente tres reyes entre las cartas seleccionadas. El número de salidas que resultan en el evento K es el producto de $C(4,3)$ y $C(48,2)$ ya que tres reyes deben ser sacados de cuatro cartas y dos cartas de los 48 restantes. La probabilidad de obtener exactamente tres reyes es

$$\frac{C(4,3)\times C(48,2)}{C(52,5)}=0.001736.$$

Ejemplo 1.16 Durante el control de calidad de un lote de 144 artefactos, 12 son elegidos al azar y inspeccionados. Si alguno de los 12 está defectuoso, el lote es rechazado, de lo contrario se acepta. Deseamos calcular la probabilidad de que un lote que contenga 10 artefactos defectuosos sea aceptado.

El número de formas en que el inspector puede elegir los 12 elementos para la inspección es el número de combinaciones de 144 elementos tomados 12 a la vez y por lo tanto es igual a

$$N=C(144,12)=\frac{144!}{12!132!}.$$

Se nos dice que todos son equiprobables,ya que el inspector elige los 12 aparatos al azar.

Sea A el evento de que el lote es aceptado, es decir, ninguno de los 10 artefactos defectuosos aparece en la muestra de 12 escogida por el inspector. Esto significa que los 12 artefactos seleccionados pertenecen a los 134 buenos. El número de formas en que esto puede suceder, denotado por $N(A)$ es

$$N(A)=C(134,12)=\frac{134!}{12!122!}$$

Se sigue entonces

$$\mathbb{P}(A)=\frac{N(A)}{N}=\frac{134!12!132!}{12!122!144!}=0.4066.$$

1.6 Combinaciones con reemplazo

Sólo queda determinar el número de combinaciones posibles, cuando se seleccionan k bolas con reemplazo de una caja que contiene n bolas distinguibles. Se puede demostrar que esto es idéntico al problema de contar el número de combinaciones cuando se seleccionan k bolas sin sustitución de una caja que contiene un total de $n+k-1$ bolas distinguibles, es decir

$$\frac{(n+k-1)!}{(n-1)!k!}.$$

Es útil describir diferentes escenarios que resultan en el mismo número de combinaciones. Consideremos, por ejemplo, el que hemos estado utilizando hasta ahora, el de contar el número de diferentes combinaciones posibles al seleccionar k bolas de una caja que contiene n bolas distinguibles.

Como ejemplo, veamos el caso que $n = 4$ y diferenciamos las bolas con las letras A, B, C y D. Si elegimos k = 3 entonces la fórmula nos dice que hay $(4+3-1)!/(3! \times 3!) = 20$ combinaciones diferentes. Estas son

```
AAA   AAB   AAC   AAD  ABB  ABC  ABD   ACC  ACD   ADD
                       BBB  BBC  BBD   BCC  BCD   BDD
                                       CCC  CCD   CDD
                                                  DDD.
```

Debes de Observar que, aunque las cuatro bolas se distinguen, la misma bola puede aparecer más de una vez en una combinación dada. Esto se debe a que después de que una bola ha sido elegida, esta se sustituye y se puede elegir una vez más. Observa también que no incluimos combinaciones como BAA o CBA porque con combinaciones, el orden no es importante y estas son equivalentes a AAB y ABC, respectivamente.

Un segundo escenario equivalente cambia las cosas. Se puede demostrar que contar el número de combinaciones en la selección de k bolas de una caja de n bolas distinguibles es equivalente a contar el número de combinaciones obtenidas al distribuir k bolas indistinguibles entre n cajas distinguibles.

Este es el problema de asignación mencionado anteriormente. Considérese el ejemplo con $n = 4$ cajas y $k = 3$ bolas y que las cuatro cajas distinguibles sean representadas por un vector de cuatro componentes enteras, que se distinguen de acuerdo con su posición en el vector. Las 20 combinaciones diferentes se dan entonces como

```
(0 0 0 3) (1 0 0 2) (2 0 0 1) (3 0 0 0)
(0 0 1 2) (1 0 1 1) (2 0 1 0)
(0 0 2 1) (1 0 2 0) (2 1 0 0)
(0 0 3 0) (1 1 0 1)
(0 1 0 2) (1 1 1 0)
(0 1 1 1) (1 2 0 0)
(0 1 2 0)
(0 2 0 1)
(0 2 1 0)
(0 3 0 0)
```

donde, por ejemplo, (1 0 0 2) indica que una bola está en la caja 1 y dos bolas están en la caja 4 y no somos capaces de distinguir entre las bolas. En este caso, necesitamos distinguir entre, por ejemplo, (1 0 0 2) y (2 0 0 1) ya que en el primer caso, la primera caja contiene una bola y la cuarta dos bolas, mientras que en la segunda, la primera caja contiene dos bolas y la cuarta sólo una bola.

Este último escenario es útil para determinar el número de estados en ciertos **modelos de redes de colas** y **cadenas de Markov**. El problema es determinar el número de formas en que k clientes idénticos pueden ser distribuidos entre n centros de cola diferentes. Esto es equivalente a que el número de vectores enteros de longitud n que satisfacen las restricciones

$$(k_1, k_2, \ldots, k_n) \text{ con } 0 \leq k_i \leq k \text{ y } \sum_{i=1}^{n} k_i = k$$

es dado por

$$\frac{(n+k-1)!}{(n-1)!k!}.$$

Resumimos las cuatro posibilidades a continuación, utilizando un ejemplo de cuatro elementos distinguibles, llamados A, B, C y D, tomados dos a la vez bajo las diversas posibilidades.

- Permutación con reemplazo

```
AA AB AC AD
BA BB BC BD
CA CB CC CD
DA DB DC DD
```

- Permutación sin reemplazo

```
   AB AC AD
BA    BC BD
CA CB    CD
DA DB DC
```

- Combinación con reemplazo

```
AA AB AC AD
   BB BC BD
      CC CD
         DD
```

- Combinación sin reemplazo

```
AB AC AD
   BC BD
      CD
```

1.7 Ensayos de Bernoulli

Vimos anteriormente que en un sólo lanzamiento de una moneda, la probabilidad de conseguir caras es la mitad. Ahora consideremos lo que sucede cuando esta moneda es lanzada n veces.

Un elemento del espacio muestral se puede escribir como $(s_1, s_2, \ldots, s_n)$, es decir, una cadena de longitud n en la que cada letra s_i es $s_i = C$ o $s_i = S, i = 1, 2, \ldots, n$. Se sigue que el tamaño del espacio muestral es 2^n y que todos los eventos elementales en este espacio son equiprobables ya que cada lanzamiento es independiente de todos los lanzamientos anteriores. Así, la probabilidad de obtener todas las caras es 2^{-n} y es la misma que la de obtener todos los sellos o exactamente k caras seguidas por $n-k$ sellos para $0 \leq k \leq n$.

Sea A el evento en la que hay exactamente k caras entre los n lanzamientos. El número de salidas en el evento A es $C(n,k)$, que es el número de maneras en que las k posicionesconteniendo C pueden seleccionarse en una cadena de longitud n. La probabilidad de obtener exactamente k caras en n lanzamientos de la moneda se da por

$$\binom{n}{k} 2^{-n}.$$

Ahora modificamos los resultados o el caso en el que la moneda no es una moneda justa. Supongamos que la probabilidad de obtener caras en un lanzamiento de la moneda es dada por p y la probabilidad de obtener sello es q donde $p + q = 1$. El espacio muestral es el mismo que antes, sólo cambian las probabilidades de los eventos elementales.

La probabilidad de lanzar n caras es $p \times p \times \cdots \times p = p^n$, la de lanzar n sellos es $q \times q \times \cdots \times q = q^n$ y la probabilidad de lanzar k caras seguidas por $n - k$ sellos es $p \times \cdots \times p \times q \times \cdots \times q = p^k q^{n-k}$.

La probabilidad de que cualquier resultado individual tenga k caras y $n - k$ sellos en algún orden también es dado por $p^k q^{n-k}$, porque los valores p y q pueden intercambiarse. La probabilidad de obtener exactamente k caras en n lanzamientos se da como

$$\binom{n}{k} p^k q^{n-k}.$$

Sustituyendo $p = q = 1/2$ da el mismo resultado anterior. Debemos notar que si sumamos sobre todos los valores posibles de k obtenemos desde el teorema binomial que debe ser el caso para una asignación de probabilidad adecuada.

$$1 = (p + q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k},$$

Secuencias de n repeticiones independientes de un experimento de probabilidad como este se denominan secuencias de Bernoulli. A menudo los resultados de cada ensayo, en lugar de ser denominados caras y sellos, se llaman éxito y fracaso, o bueno y defectuoso y así sucesivamente, dependiendo de la aplicación particular en cuestión.

Ejemplo 1.17 Un experimento tiene probabilidad de éxito igual a 0.7. Determinemos la probabilidad de tres éxitos y dos fracasos en una secuencia de cinco ensayos independientes del experimento.

Sea la probabilidad de éxito denotada por p y la probabilidad de fracaso por $q = 1 - p$. La probabilidad de obtener tres éxitos y dos fracasos es igual a $p^3 q^2 = 0,7^3 \times 0,3^2 = 0,03087$. El número de maneras en que esto puede suceder, esto es, el número de resultados con exactamente tres éxitos, es $C(5,3) = 10$. Así, la respuesta que buscamos es igual a 0.3087.

Ejemplo 1.18 Consideremos un canal de comunicación binario que envía palabras codificadas en n bits. La probabilidad de que un sólo bit sea recibido correctamente es p. La probabilidad de que se reciba eun error es $q = 1 - p$. Utilizando las funciones de corrección de código, es posible corregir tantos e errores por palabra enviada. Determinemos la probabilidad de que la transmisión de cualquier palabra sea exitosa, suponiendo que la transmisión de bits individuales es independiente.

Para resolver este problema, observe que, siempre y cuando el número de bits en error sea e o menos, la palabra se recibirá correctamente. La probabilidad de que la palabra sea recibida teniendo exactamente k bits en error se da como

$$\binom{n}{k} (1 - p)^k p^{n-k}$$

Por tanto la probabilidad que la palabra sea correctamente recibida es

$$\sum_{k=0}^{e} \binom{n}{k} (1 - p)^k p^{n-k}.$$

Obsérvese que en una secuencia de n ensayos independientes con probabilidad de éxito igual a p, la probabilidad de n_1 éxitos y n_2 fracasos, con $n = n_1 + n_2$ viene dado por

$$\binom{n}{n_1} p^{n_1} (1 - p)^{n-n_1} = \binom{n}{n_2} (1 - p)^{n_2} p^{n-n_2},$$

Los ensayos de Bernoulli aparecen a menudo en el contexto del `modelado de la confiabilidad`. Un sistema es representado por un conjunto de componentes independientes, cada uno de los cuales tiene su propia probabilidad de éxito. Los componentes pueden organizarse en series en las que la salida de un componente se convierte en la entrada para el siguiente. El éxito se consigue si se obtiene un resultado exitoso con cada uno de los componentes de la serie. Si alguno de los componentes falla, entonces todo el sistema falla.

Alternativamente, los componentes pueden estar dispuestos en paralelo y se logra un resultado satisfactorio si uno o más de los componentes logra completar correctamente su tarea. El propósito de disponer componentes en paralelo es a menudo aumentar la confiabilidad general del sistema: en el caso de que un componente que falla los otros siguen funcionando y un resultado exitoso para el sistema como un todo sigue siendo posible.

También son posibles las combinaciones de componentes en serie y en paralelo. Consideremos el caso de n componentes en serie, en las que la probabilidad de éxito de cada componente es p. La probabilidad de que la operación completa tenga éxito es igual a la probabilidad de que todos los componentes n funcionen correctamente y es igual a p^n. La probabilidad de que la operación general no tenga éxito es $1 - p^n$.

Ahora consideremos el caso cuando n componentes se colocan en paralelo y se logra el éxito si `al menos uno de los componentes` tiene éxito. En este caso, es más fácil calcular primero la probabilidad del evento que de la operación del sistema no sea un éxito y luego calcular la probabilidad del complemento de este evento. En general, en las preguntas relacionadas con la probabilidad, a menudo es `más fácil calcular la probabilidad de intersección de eventos que computar la unión de un conjunto de eventos`.

El uso de las técnicas de complementación y la ley de DeMorgan con frecuencia nos permiten elegir la opción más fácil. En la situación actual, se produce un fallo del sistema si todos los n componentes fallan (todos implican intersección, mientras que al menos uno implica unión). La probabilidad de fallo en un sistema de n componentes dispuestos en paralelo, en el que la probabilidad de fallo de cada componente es $1 - p$ es igual a $(1 - p)^n$. A partir de esto podemos calcular la probabilidad de éxito de todo el sistema: $1 - (1 - p)^n$.

Ejemplo 1.19 En entornos críticos, no es raro que varias computadoras realicen exactamente la misma secuencia de tareas de modo que si falla un equipo, la misión no se ve comprometida.

La misión continuará y se considerará exitosa, siempre y cuando por lo menos una de las computadoras ejecute sus tareas correctamente. Considere un sistema en el cual tres computadoras realizan la misma secuencia de seis tareas una después de la otra. Las pruebas de consistencia se integran en cada tarea.

Se supone que una computadora funciona correctamente si su estado satisface las pruebas de consistencia. Sea 0.04 la probabilidad de que una computadora no cumpla la prueba de consistencia durante cualquier tarea . Queremos calcular la probabilidad de que la misión sea exitosa.

Esta situación puede representarse como un modelo de fiabilidad que consta de tres partes en paralelo (las tres computadoras) y en la que cada parte contiene seis componentes en serie. Sea $p = 0,96$ la probabilidad de que una computadora satisfaga la prueba de consistencia en cualquier tarea dada. Entonces, la probabilidad de que una sola computadora complete satisfactoriamente las seis tareas es igual a $p^6 = 0.7828$, por otro lado, $p^{'} = 1 - p^6$ es la probabilidad de que no.

Dado que hay tres equipos en paralelo, la probabilidad de que los tres fallan es $(1 - p^{'})^3 = 0.01025$. Luego la probabilidad de una misión exitosa $1 - (1 - p^{'})^3 = 0.9897$.

El concepto de un ensayo de Bernoulli puede ser generalizado a partir de dos resultados posibles a muchos resultados posibles. Supongamos que en lugar de dos posibilidades (caras y sellos, 0 y 1 , éxito y fracaso) hay m posibilidades $h_1, h_2, \ldots h_m$ con probabilidades $p_1, p2, \ldots, p_m$ y $\sum_{j=1}^{m} p_j = 1$.

En cada ensayo, puede ocurrir una y sólo una de las m posibilidades. Esta vez el espacio muestral consta de n-tuplas, en el cual cada elemento en una de las m opciones $h_i, i = 1, 2, \ldots, m$. Supongamos por ejemplo que el número de ensayos es $n = 6$ y que el número posible de salidas de cada ensayo es $m = 4$. Denotemos las diferentes salidas por a, b, c y d. Algunos posibles elementos del espacio muestral son

```
(a, a, a, a, a, a), (c, a, d, d, a, b), (a, b, b, c, c, c), ....
```

Consideremos ahora esos los elementos particulares del espacio muestral para los cuales h_i se produce n_i veces, para $i = 1, 2, \ldots, m$. Debemos tener que $\sum_{j=1}^{m} n_j = n$. Por ejemplo, sea $n_1 = 2, n_2 = 2, n_3 = 0$ y n_4, algunas posibilidades son

```
(a, b, b, d, d, d), (b, a, d, d, d, b), (d, a, d, d, b, b), ....
```

El número total de eventos elementales que tienen la propiedad de que h_i ocurre n_i veces, sujeto a las condiciones anteriores, viene dado por el coeficiente multinomial

$$\binom{n}{n_1, n_2, \ldots, n_m} = \frac{n!}{n_1! n_2! \ldots n_m!}.$$

En general, el número de formas de colocar h_1 en n_1 ranuras, h_2 en n_2 ranuras y así sucesivamente, con $\sum_{j=1}^{m} n_j = m$ es dado por el coeficiente multinomial. La probabilidad de ocurrencia de algún evento elemental es dado por $p_1^{n_1}, p_2^{n_2}, \cdots p_m^{n_m}$ y así la probabilidad que h_i ocurre n_1 veces, h_2 ocurre n_2 veces y así es dado por

$$p(n_1, n_2, \ldots, n_m) = \frac{n!}{n_1! n_2! \ldots n_m!} p_1^{n_1}, p_2^{n_2}, \cdots p_m^{n_m}.$$

Ejemplo 1.20 Una clase introductoria de nivel de postgrado en probabilidad aplicada tiene 10 estudiantes de pregrado, 30 estudiantes de maestría y 20 estudiantes de doctorado. Durante las horas de oficina del profesor, los estudiantes llegan de forma independiente a buscar respuestas. ¿Cuál es la probabilidad de que, de los últimos ocho estudiantes que visitan al profesor, dos fueran estudiantes de pregrado, cinco fueran de maestría y uno fuera de doctorado?

Asumimos que la probabilidad de que un estudiante de pregrado busque al profesor durante las horas de oficina es $10/60$, la de un estudiante de maestría $30/60$ y la de un estudiante de doctorado $20/60$. En este caso la respuesta requerida es dada por

$$\binom{8}{2, 5, 1} \left(\frac{10}{60}\right)^2 \left(\frac{30}{60}\right)^5 \left(\frac{20}{60}\right) = \frac{7}{144}.$$

2 Referencias

1. Probability, Stochastic Processes, and Queueing Theory:The Mathematics of Computer Performance Modeling Nelson, Randolph chapter 3.

2. Lecture 3: Learning to count; Binomial Distribution, Introduction to Probability and Statistics Karl Border 2016.

3. Introduction to Probability Dimitri P. Bertsekas and John N. Tsitsiklis Athena Scientific , Belmont , Massachusetts 2008 chapter 1 **Counting**.