Número de éxitos para n variables Bernoulli no independientes

Pedro Pury

1. Introducción

Se trata de resolver el siguiente problema:

Se baraja un conjunto de $n = 100$ cartas (numeradas consecutivamente del 1 al 100) y se extrae del mazo una carta por vez. Consideramos que ocurre un "éxito" si la i-ésima carta extraída es aquella cuyo número es i ($i = 1, \dots, n$).

Pregunta: ¿Cuál es la distribución de la variable número de éxitos?

Resolveremos este problema considerando un mazo de n cartas, y estudiaremos además la distribución de la variable número de éxitos cuando el número de cartas n tiende a infinito.

Notemos que el resultado éxito o fracaso para una determinada carta, influye en el resultado de las siguientes, por lo tanto los resultados no son independientes.

1.1. Notación

Usaremos la siguiente notación para denotar los eventos en los cuales ocurre algún éxito:

- E_i: evento en el cual ocurre un éxito en la i-ésima carta.
- $E_{i_1,i_2,\dots,i_m} = E_{i_1} \cap \cdots \cap E_{i_m}$: evento en el cual ocurren éxitos en las cartas $i_1, \dots, i_m$.

Notemos que estos eventos no excluyen la posibilidad que haya éxitos en otras cartas. Por lo tanto usaremos otra notación para especificar que hay éxito sólo en determinadas cartas:

- A_i: evento en el cual sólo ocurre un éxito en la i-ésima carta.
- $A_{i_1,i_2,\dots,i_m} = A_{i_1} \cap \cdots \cap A_{i_m}$: evento en el cual ocurren éxitos sólo en las cartas $i_1, \dots, i_m$.

2. Cálculo de las probabilidades $P(E_i)$ y $P(A_i)$

Dado que E_i es el evento en el cual la i-ésima carta es un éxito, independientemente de lo que ocurra en las demás, se tiene que las $n-1$ cartas restantes pueden ocupar cualquier lugar. Por lo tanto

$$P(E_i) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

En cambio, si consideramos A_i, debemos tener en cuenta que ninguna otra carta salvo la i-ésima puede resultar en éxito. Notemos que el evento A_i puede describirse en términos de uniones e intersecciones de eventos como

$$A_i = E_i \cap \left(\bigcup_{j \neq i} E_j \right)^c ,$$

y por lo tanto su probabilidad de ocurrencia es igual a

$$\begin{aligned} P(A_i) &= P\left(E_i \cap \left(\bigcup_{j \neq i} E_j \right)^c \right) \\ &= P(E_i) - \sum_{i \neq j = 1}^{n} P(E_i \cap E_j) + \sum_{j<k_{\neq i}}^{n} P(E_i \cap E_j \cap E_k) + \cdots + (-1)^{n-1} P(E_1 \cap E_2 \cap \cdots \cap E_n) \end{aligned}$$

Calculando cada término de esta suma, tenemos que

$$P(E_i \cap E_j) = \frac{(n-2)!}{n!},$$

ya que debemos contar todas las permutaciones posibles de las $n-2$ cartas distintas de i y j. Así siguiendo tenemos

$$\begin{aligned} P(E_i \cap E_j \cap E_j) &= \frac{(n-3)!}{n!} \\ P(E_i \cap E_j \cap E_j \cap E_l) &= \frac{(n-4)!}{n!} \\ &\vdots \\ P(E_1 \cap E_2 \cap \cdots \cap E_n) &= \frac{1}{n!} \end{aligned}$$

Por último,

$$\begin{aligned} P(A_i) &= \frac{(n-1)!}{n!} - (n-1)\frac{(n-2)!}{n!} + \binom{n-1}{2}\frac{(n-3)!}{n!} - \cdots + (-1)^{n-1}\frac{1}{(n-1)!} \\ &= \sum_{j=0}^{n-1} \binom{n-1}{j} \frac{(n-1-j)!}{n!} (-1)^j \\ &= \sum_{j=0}^{n-1} (-1)^j \frac{1}{nj!} \end{aligned}$$

Ahora, la probabilidad P_1 de tener éxito en exactamente una carta es igual a

$$P_1 = P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{j=0}^{n-1} (-1)^j \frac{1}{j!}.$$

Notemos que si $n \to \infty$, entonces $P_1 = e^{-1}$.

3. Cálculo de la probabilidad de ocurrencia de más de un éxito

Consideramos ahora el evento $E_{i_1,\ldots,i_m}$ en el cual se tiene éxito simultáneamente en las cartas $i_1,\ldots,i_m$, independientemente del resto. Notemos que

$$E_{i_1,\ldots,i_m} = E_{i_1} \cap \cdots \cap E_{i_m}$$

por lo que, según ya se ha calculado antes

$$P\left(E_{i_1,\ldots,i_m}\right) = \frac{(n-m)!}{n!}.$$

Si ahora consideramos el evento $A_{i_1,\ldots,i_m}$ en el cual sólo se tienen éxitos en estas m cartas, tenemos que

$$\begin{aligned}
P\left(A_{i_1,\ldots,i_m}\right) &= P\left(E_{i_1,\ldots,i_m} \cap \left(\bigcup_{j\neq i_1,\ldots,i_m} E_j\right)^c\right) \\
&= P\left(E_{i_1,\ldots,i_m}\right) - \sum_{j\neq i_1,\ldots,i_m} P\left(E_{i_1,\ldots,i_m}\cap E_j\right) + \\
&\quad + \sum_{j<k} P\left(E_{i_1,\ldots,i_m}\cap E_j \cap E_k\right) + \cdots + (-1)^{n-m} P\left(E_1\cap E_2\cap\cdots\cap E_n\right) \\
&= \frac{(n-m)!}{n!} - (n-m)\frac{(n-(m+1))!}{n!} + \binom{n-m}{2}\frac{(n-(m+2))!}{n!} \\
&\quad + \cdots + (-1)^{n-m}\frac{1}{n!} \\
&= \sum_{h=0}^{n-m} (-1)^h \binom{n-m}{h}\frac{(n-m-h)!}{n!} \\
&= \sum_{h=0}^{n-m} (-1)^h \frac{(n-m)!}{h!\,n!}
\end{aligned}$$

Finalmente se obtiene que

$$P\left(A_{i_1,\ldots,i_m}\right) = \frac{1}{m!\binom{n}{m}} \sum_{h=0}^{n-m} \frac{(-1)^h}{h!}.$$

De este modo, la probabilidad P_m de tener exactamente m éxitos se calcula multiplicando por el número de veces que se pueden elegir m cartas distintas, es decir

$$P_m = \binom{n}{m} P\left(A_{i_1,\ldots,i_m}\right) = \frac{1}{m!}\sum_{h=0}^{n-m}\frac{(-1)^h}{h!}.$$

4. La distribución de Poisson $\mathcal{P}(1)$

De las fórmulas obtenidas anteriormente, podemos ver que si $n \to \infty$ entonces la probabilidad de tener exactamente m éxitos tiende a $\frac{e^{-1}}{m!}$:

$$\lim_{n\to\infty} P_m = \frac{e^{-1}}{m!}.$$

Esto responde a una distribución de Poisson de parámetro $\lambda = 1$.