EJERCICIO 1 - PRÁCTICO 4

PROBLEMAS DE COINCIDENCIAS

ANA GEORGINA FLESIA - PATRICIA KISBYE.

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 ($i = 1,\ldots,n$).

Pregunta: ¿Cuál es la distribución de la variable número de éxitos? ¿Cual es su esperanza y su varianza?

2 Problemas de coincidencias en la permutación de n objetos

2.1 Notación

Supongamos tener el caso de la permutación de los n primeros números naturales.Denotamos por K la variable que cuenta el número de éxitos, es decir el número de coincidencias entre el valor de un número y la posición que ocupa. Notar que el éxito o fracaso en una determinada posición influye en los resultados siguientes por lo tanto los resultados no son independientes.

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 posición.
- $E = \bigcup_{i=1}^{n} E_i$ evento al menos una coincidencia.
- $E_{i_1,i_2,\ldots,i_m} = E_{i_1} \cap \cdots \cap E_{i_m}$: evento en el cual ocurren éxitos en las posiciones $i_1, \ldots, i_m$.

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

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

Por lo cual, la distribución de la variable K esta relacionada con la probabilidad de los eventos $A_{i_1,i_2,\ldots,i_m}$.

2.2 Cálculo de la probabilidad de por lo menos un éxito, $P(E)$ y ningún éxito, p_n

Ya vimos en Probabilidad y Estadística que

$$P(E_i \cup E_j) = P(E_i) + P(E_j) - P(E_i \cap E_j)$$

Y que

$$P(E_i \cup E_j \cup E_k) = P(E_i) + P(E_j) + P(E_k) - P(E_i \cap E_j) - P(E_i \cap E_k) - P(E_k \cap E_j) + P(E_i \cap E_j \cap E_k)$$

Por inducción sobre el número de conjuntos a unir, sale que si

$$S_r = \sum_{1 \leq < i_1 < i_2, \cdots < i_r} P(E_{i_1, i_2, \ldots, i_r})$$

entonces

$$P(E) = \sum_{r=1}^{n} (-1)^{r-1} S_r = S_1 - S_2 + S_3 \cdots + (-1)^{n-1} S_n$$

Observamos que cada S_r tiene $\binom{n}{r}$ términos, esto es, el número de subconjuntos de indices de tamaño r. Observemos también que dado que las intersecciones son siempre entre eventos distintos,

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

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

$$P(E_{i_1, i_2, \ldots, i_r}) = \frac{(n-r)!}{n!}$$

que no depende de los indices sino de el número de intersecciones r, y por lo tanto

$$\begin{aligned} S_r &= \sum_{1 \leq < i_1 < i_2, \cdots < i_r} P(E_{i_1, i_2, \ldots, i_r}) \\ &= \binom{n}{r} \frac{(n-r)!}{n!} \\ &= \frac{n!}{r!(n-r)!} \frac{(n-r)!}{n!} \\ &= \frac{1}{r!} \end{aligned}$$

y

$$\begin{aligned} P(E) &= \sum_{r=1}^{n} (-1)^{r-1} S_r = S_1 - S_2 + S_3 \cdots + (-1)^{n-1} S_n \\ &= \sum_{r=1}^{n} \frac{(-1)^{r-1}}{r!} \end{aligned}$$

Con esta información podemos calcular p_n , la probabilidad de no tener ninguna coincidencia, ya que $P(E) = 1 - p_n$

$$p_n = 1 - \sum_{r=1}^{n} \frac{(-1)^{r-1}}{r!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

2.3 Cálculo de las probabilidades $P(E_i)$ y $P(A_i)$

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

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

Ahora, la probabilidad de que haya exactamente r coincidencias y que estas ocurran en r posiciones fijas $\{i_1, i_2, \ldots, i_r\}$ es la misma para todas las especificaciones, depende solo de r el numero de coincidencias, no donde se realizan. Por lo cual podemos pensar primero en el siguiente evento, "que haya exactamente

r coincidencias y que estén en los primeros r lugares". Este evento es equivalente al evento " que no haya ninguna coincidencia en los últimos $n-r$ lugares". Y la fracción de permutaciones que cumplen esa condición es $(n-r)!p_{n-r}$, por lo cual

$$\begin{aligned} P(A_{i_1,i_2,\ldots,i_r}) &= P(A_{1,2,\ldots,r}) \\ &= \frac{(n-r)!p_{n-r}}{n!} \\ &= \frac{(n-r)!}{n!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} \end{aligned}$$

Tambien es importante observar que los eventos $A_{i_1,i_2,\ldots,i_r}$ son eventos disjuntos si las especificaciones $\{i_1,i_2,\ldots,i_r\}$ son diferentes, y el numero de tales especificaciones es $\binom{n}{r}$.

Por lo cual,

$$\begin{aligned} P^n(r) = P(\text{exactamente r coincidencias}) &= \binom{n}{r}P(A_{1,2,\ldots,r}) \\ &= \binom{n}{r}\frac{(n-r)!}{n!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} \\ &= \frac{n!}{r!(n-r)!}\frac{(n-r)!}{n!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} \\ &= \frac{1}{r!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} \end{aligned}$$

2.4 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 r éxitos tiende a $\frac{e^{-1}}{r!}$:

$$\lim_{n\to\infty}P^n(r) = \frac{1}{r!}\lim_{n\to\infty}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} = \frac{e^{-1}}{r!}.$$

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

3 Resolución del ejercicio

Consideremos la variable K_n, el número de coincidencias obtenidas en una baraja de n cartas. Entonces, la variable K_n tiene distribución discreta

$$P(K_n = m) = P^n(m) = \frac{1}{m!}\sum_{h=0}^{n-m}\frac{(-1)^h}{h!}$$

y su esperanza y varianza son

$$
\begin{aligned}
E(K) &= \sum_{m=0}^{n} mP(K=m) \\
&= \sum_{m=1}^{n} m\frac{1}{m!}\sum_{h=0}^{n-m}\frac{(-1)^h}{h!} \\
&= \sum_{m=1}^{n} \frac{1}{(m-1)!}\sum_{h=0}^{(n-1)-(m-1)}\frac{(-1)^h}{h!} \\
&= \sum_{m=1}^{n} \frac{1}{(m-1)!}\sum_{h=0}^{((n-1)-(m-1)}\frac{(-1)^h}{h!} \\
&= \sum_{k=0}^{n-1} \frac{1}{k!}\sum_{h=0}^{n-1-k}\frac{(-1)^h}{h!} \\
&= \sum_{k=0}^{n-1} P(K_{(n-1)}=k) \\
&= 1
\end{aligned}
$$

donde $K_{(n-1)}$ es el numero de coincidencias en una baraja con $(n-1)$ cartas.

$$
\begin{aligned}
E(K^2) &= \sum_{m=0}^{n} m^2\frac{1}{m!}\sum_{h=0}^{n-m}\frac{(-1)^h}{h!} = \sum_{m=1}^{n}\frac{m}{(m-1)!}\sum_{h=0}^{(n-m)}\frac{(-1)^h}{h!} \\
&= \sum_{m=1}^{n} \frac{m-1+1}{(m-1)!}\sum_{h=0}^{(n-m)}\frac{(-1)^h}{h!} \\
&= \sum_{m=2}^{n} \frac{1}{(m-2)!}\sum_{h=0}^{(n-2)-(m-2)}\frac{(-1)^h}{h!} + \sum_{m=1}^{n}\frac{1}{(m-1)!}\sum_{h=0}^{(n-1)-(m-1)}\frac{(-1)^h}{h!} \\
&= \sum_{m=0}^{n-2} P^{n-2}(m) + \sum_{m=0}^{n-1} P^{n-1}(m) \\
&= 1+1=2
\end{aligned}
$$

Luego

$$
Var(K) = E(K^2) - (E(K))^2 = 2-1 = 1.
$$

Por lo cual, el número de coincidencias tiene la esperanza y varianza de una distribución de Poisson $\mathcal{P}(1)$, y su distribución converge a una distribución $\mathcal{P}(1)$ cuando el número de cartas tiene a infinito.