1. IV Premios Jorge Juan

1.1. **Cálculo de Probabilidades.**

1.- Se considera un mercado de nk consumidores, donde n y k son dos números naturales. Cada uno de ellos tiene que adquirir un único bien, existiendo en el mercado n proveedores. Cada consumidor elige completamente al azar el proveedor al cual comprará el objeto en cuestión.

Sea X_i el número de consumidores que compran el objeto al i-ésimo proveedor $(i = 1, 2, ..., n)$, y sea Y el número de proveedores que no han vendido ningún objeto. Se pregunta:

i) ¿Cuál es la distribución de probabilidad de X_i?.

ii) Sea $i \neq j$. A partir de la expresión que resulta de elevar al cuadrado la identidad $X_1+X_2+...+X_n = nk$, obtener las esperanzas $E(X_iX_j)$ y, a partir de ellas, calcular las covarianzas $\mathrm{cov}(X_i, X_j)$. Deducir, también, el coeficiente de correlación de X_i y X_j, e interpretar el caso $n = 2$.

iii) Calcular $E(Y)$ y $V(Y)$. Para ello resultará útil considerar las variables

$$Z_i := \begin{cases} 1, & \text{si el proveedor } i \text{ no vende ningún objeto} \\ 0, & \text{en caso contrario.} \end{cases}$$

iv) Encontrar los límites $\lim_{n\to\infty} E(Y/n)$ y $\lim_{n\to\infty} V(Y/n)$.

2.- Una urna contiene a bolas blancas y b bolas negras. Dos individuos, 1 y 2, practican el juego siguiente: comenzando el jugador 1, ellos van sacando alternativamente una bola (sin restituirla a la urna), hasta que uno de ellos saca una bola blanca, en cuyo momento este jugador es proclamado ganador del juego.

Deducir el valor de la suma

$$S := 1 + \frac{b}{a+b-1} + \frac{b(b-1)}{(a+b-1)(a+b-2)} + ... + \frac{b!}{(a+b-1)(a+b-2)...a}$$

Solución del Ejercicio 1:
i) $X_i \sim BIN(nk, \frac{1}{n})$.
ii) Puesto que

$$n^2k^2 = (X_1 + X_2 + ... + X_n)^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_iX_j,$$

se tendrá

$$\begin{aligned} E(n^2k^2) &= n^2k^2 = \sum_{i=1}^{n} E(X_i^2) + 2\sum_{i<j} E(X_iX_j) \\ &= nE(X_1^2) + 2\frac{n(n-1)}{2}E(X_1X_2). \end{aligned}$$

Despejando:

$$\begin{aligned} E(X_1X_2) &= \frac{1}{n(n-1)}\left\{n^2k^2 - n\left[k\left(1-\frac{1}{n}\right) + k^2\right]\right\} \\ &= k^2 - \frac{k}{n}. \end{aligned}$$

Por lo tanto

$$\text{cov}(X_i, X_j) = E(X_iX_j) - E(X_i)E(X_j) = k^2 - \frac{k}{n} - k^2 = -\frac{k}{n}.$$

Obtengamos, ahora, el coeficiente de correlación:

$$\rho(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{V(X_i)V(X_j)}} = \frac{-k/n}{k(1-\frac{1}{n})} = \frac{1}{1-n}.$$

Para $n = 2$ se obtiene $\rho(X_1, X_2) = -1$, lo cual es absolútamnete lógico en la medida que existe una relación funcional lineal entre ambas variables: $X_2 = 2k - X_1$.

iii) A la vista de la definición de las variables Z_i, se tiene

$$Y = \sum_{i=1}^{n} Z_i.$$

Las variables Z_i siguen una distribución de Bernoulli con parámetro ("probabilidad de éxito"):

$$\begin{aligned} E(Z_i) &= P(Z_i = 1) \\ &= P(X_i = 0) = (1 - \frac{1}{n})^{nk}, \end{aligned}$$

por lo que

$$E(Y) = \sum_{i=1}^{n} E(Z_i) = nE(Z_1) = n(1 - \frac{1}{n})^{nk}.$$

De otro lado
$$V(Y) = V\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=1}^{n} V(Z_i) + 2\sum_{i<j} \text{cov}(Z_i, Z_j).$$
Por una parte $V(Z_i) = (1-\frac{1}{n})^{nk} - (1-\frac{1}{n})^{2nk}$. Por otra
$$\begin{aligned} E(Z_iZ_j) &= P(Z_i = 1 \wedge Z_j = 1) \\ &= P(X_i = 0 \wedge X_j = 0) = (1-\frac{2}{n})^{nk}. \end{aligned}$$
En base a este último resultado:
$$\text{cov}(Z_i, Z_j) = (1-\frac{2}{n})^{nk} - (1-\frac{1}{n})^{2nk}.$$
Combinando los últimos resultados:
$$\begin{aligned} V(Y) &= n\left[(1-\frac{1}{n})^{nk} - (1-\frac{1}{n})^{2nk}\right] + \frac{2n(n-1)}{2}\left[(1-\frac{2}{n})^{nk} - (1-\frac{1}{n})^{2nk}\right] \\ &= n(1-\frac{1}{n})^{nk} - n^2(1-\frac{1}{n})^{2nk} + (n^2-n)(1-\frac{2}{n})^{nk}. \end{aligned}$$
iv)
$$E(Y/n) = (1-\frac{1}{n})^{nk} \Longrightarrow \lim_{n\to\infty} E(Y/n) = \exp(-k).$$
Análogamente,
$$\begin{aligned} \lim_{n\to\infty} V(Y/n) &= \lim_{n\to\infty}\{\frac{1}{n}(1-\frac{1}{n})^{nk} - (1-\frac{1}{n})^{2nk} + (1-\frac{2}{n})^{nk} - \frac{1}{n}(1-\frac{2}{n})^{nk}\} \\ &= 0 - \exp(-2k) + \exp(-2k) - 0 = 0. \end{aligned}$$

Solución al Ejercicio 2:

Denetaremos por A_k el suceso "gana 1 en la extracción k-ésima". Ouesto que las extraciones efectuadas por 1 son las impares, se tendrá $P(A_{2p}) = 0$, $\forall p \in \mathbb{N}$. De otro lado, $P(A_1) = \frac{a}{a+b}$, y para que gane 1 al efectuar la extracción $2p+1$, tiene que salir bola negra en las primeras *2p* extracciones, y blanca en la *2p+1*, con lo que
$$\begin{aligned} P(A_{2p+1}) &= \frac{b}{a+b}\cdot\frac{b-1}{a+b-1}\cdot....\cdot\frac{b-(2p-1)}{a+b-(2p-1)}\cdot\frac{a}{a+b-2p} \\ &= \frac{a}{a+b}\cdot\frac{b}{a+b-1}\cdot\frac{b-1}{a+b-2}....\cdot\frac{b-(2p-2)}{a+b-(2p-1)}\cdot\frac{b-(2p-1)}{a+b-2p}. \end{aligned}$$
Esta fórmula es válida siempre que $b > 2p-1$. Si $b \leq 2p-1$, se tendrá $P(A_{2p+1}) = 0$. Entonces,
$$P(\text{ganar el jugador 1}) = \sum_{0\leq p<\frac{1}{2}(1+b)} P(A_{2p+1}).$$

Si por B_k representamos el suceso "gana 2 en la extracción k-ésima", se verificará $P(B_{2p+1}) = 0$, $\forall p \in \mathbb{N}$, y

$$\begin{aligned} P(B_{2p}) &= \frac{b}{a+b} \cdot \frac{b-1}{a+b-1} \cdot .... \cdot \frac{b-(2p-2)}{a+b-(2p-2)} \cdot \frac{a}{a+b-(2p-1)} = \\ &= \frac{a}{a+b} \cdot \frac{b}{a+b-1} \cdot \frac{b-1}{a+b-2} .... \cdot \frac{b-(2p-2)}{a+b-(2p-1)}. \end{aligned}$$

Como

$$P(A_1) + \sum_{1 \leq p < \frac{1}{2}(1+b)} \{P(B_{2p}) + P(A_{2p+1})\} = 1,$$

se deduce

$$\begin{aligned} 1 &= \frac{a}{a+b} + \left\{ \frac{a}{a+b} \cdot \frac{b}{a+b-1} + \frac{a}{a+b} \cdot \frac{b}{a+b-1} \cdot \frac{b-1}{a+b-2} \right\} + ... \\ &= \frac{a}{a+b} \left\{ 1 + \frac{b}{a+b-1} + \frac{b}{a+b-1} \cdot \frac{b-1}{a+b-2} + ... \right\} = S \cdot \frac{a}{a+b}. \end{aligned}$$

Por lo tanto

$$S = \frac{a+b}{a}.$$