Cálculo de Probabilidades I — Febrero 2014 — Primera semana

Ejercicio 1. Un jugador toma, de una en una, cartas de una baraja española (con un total de cuarenta cartas). Sea N el número de extracción en la que ha obtenido el cuarto as.

(a) Determinar la función de probabilidad de N y hallar su moda.

(b) Calcular $\mathrm{E}[N]$ y $\mathrm{Var}(N)$.

Ejercicio 2. La variable aleatoria X tiene distribución de Poisson de parámetro $\lambda > 0$. Condicionada por el valor $X = n$, con $n \geq 0$, la variable aleatoria Y tiene distribución binomial de parámetros n y p, con $0 < p < 1$. (Cuando $X = 0$, se entiende que $Y = 0$.)

(a) Determinar la función de probabilidad de Y y la la función de probabilidad de X condicionada por $Y = m$, para $m \geq 0$.

(b) Dar la expresión de la recta de regresión de X sobre Y. Calcular $\mathrm{Cov}(X, Y)$ y el coeficiente de correlación entre X e Y.

(c) Probar que las variables Y y $X - Y$ son independientes.

Solución

Ejercicio 1.
(a) Dado un entero $4 \leq k \leq 40$ se tiene que

$$\begin{aligned} \mathrm{P}\{N = k\} &= \frac{\binom{4}{3} \cdot \binom{36}{k-4}}{\binom{40}{k-1}} \cdot \frac{1}{41-k} \\ &= \frac{1}{6 \cdot \binom{40}{4}} \cdot (k-1)(k-2)(k-3), \end{aligned}$$

dado que, para que sea $N = k$, hay que obtener tres ases en las $k-1$ primeras extracciones (y $k-4$ cartas de las 36 que no son ases) y un as en la extracción k-ésima, cuando quedan $41-k$ cartas en la baraja.

También se puede razonar de la siguiente manera. Al extraer las 40 cartas nos interesa la posición de los cuatro ases. Hay $\binom{40}{4}$ posiciones posibles. Las que hacen que $N = k$ son las que sitúan tres ases en los $k-1$ primeros lugares, con $\binom{k-1}{3}$ posibilidades, y un as en el k-ésimo lugar. Se tiene pues

$$\mathrm{P}\{N = k\} = \frac{\binom{k-1}{3}}{\binom{40}{4}}.$$

Se observa que $\mathrm{P}\{N = k\}$ es una función creciente de $k \geq 4$, luego alcanza el máximo en $k = 40$ siendo $\mathrm{P}\{N = 40\} = \frac{1}{10}$.
(b) El valor esperado de N es

$$\begin{aligned} \mathrm{E}[N] &= \frac{1}{6 \cdot \binom{40}{4}} \sum_{k=4}^{40} k(k-1)(k-2)(k-3) \\ &= \frac{4}{\binom{40}{4}} \sum_{k=4}^{40} \binom{k}{4}. \end{aligned}$$

Se tiene (ver ecuación (I.13) del libro de texto)

$$\sum_{k=4}^{40} \binom{k}{4} = \binom{41}{5}, \quad \text{luego} \quad \mathrm{E}[N] = \frac{164}{5} = 32{,}8.$$

También se puede razonar directamente. Nótese que $\sum_{k=4}^{40} \binom{k}{4}$ es el coeficiente del término de grado 4 en el polinomio

$$Q(x) = 1 + (1+x) + (1+x)^2 + \cdots + (1+x)^{40}.$$

Puesto que

$$Q(x) = \frac{(1+x)^{41} - 1}{x},$$

resulta que $\sum_{k=4}^{40} \binom{k}{4}$ es el coeficiente del término de grado 5 de $(1+x)^{41}$, esto es,

$$\sum_{k=4}^{40} \binom{k}{4} = \binom{41}{5}.$$

Para calcular la varianza, se hace

$$\begin{aligned}
\mathrm{E}[N(N+1)] &= \frac{1}{6 \cdot \binom{40}{4}} \sum_{k=4}^{40} (k+1)k(k-1)(k-2)(k-3) \\
&= \frac{20}{\binom{40}{4}} \sum_{j=5}^{41} \binom{j}{5} \\
&= 20 \cdot \frac{\binom{42}{6}}{\binom{40}{4}} = 1\,148,
\end{aligned}$$

donde se ha hecho el cambio de índice $j = k+1$. Así, $\mathrm{E}[N^2] = 1\,115{,}2$. Se obtiene finalmente $\mathrm{Var}(N) = 39{,}36$ y una desviación típica $\sigma(N) \simeq 6{,}27$.

Ejercicio 2.
(a) Fijado $m \geq 0$ se tiene

$$\begin{aligned}
\mathrm{P}\{Y = m\} &= \sum_{n=m}^{\infty} \mathrm{P}\{Y = m \mid X = n\} \cdot \mathrm{P}\{X = n\} \\
&= \sum_{n=m}^{\infty} \binom{n}{m} p^m (1-p)^{n-m} \cdot e^{-\lambda} \frac{\lambda^n}{n!} \\
&= e^{-\lambda} \frac{(\lambda p)^m}{m!} \sum_{n=m}^{\infty} \frac{(\lambda(1-p))^{n-m}}{(n-m)!} \\
&= e^{-\lambda p} \frac{(\lambda p)^m}{m!}.
\end{aligned}$$

Se concluye que Y tiene distribución de Poisson de parámetro λp.

Dados $0 \leq m \leq n$ se tiene

$$\begin{aligned}
\mathrm{P}\{X = n \mid Y = m\} &= \frac{\mathrm{P}\{Y = m \mid X = n\} \cdot \mathrm{P}\{X = n\}}{\mathrm{P}\{Y = m\}} \\
&= \frac{\binom{n}{m} p^m (1-p)^{n-m} \cdot e^{-\lambda} \frac{\lambda^n}{n!}}{e^{-\lambda p} \frac{(\lambda p)^m}{m!}} \\
&= e^{-\lambda(1-p)} \frac{(\lambda(1-p))^{n-m}}{(n-m)!}.
\end{aligned}$$

Por tanto, condicionada por $Y = m$, X se distribuye como una Poisson de parámetro $\lambda(1-p)$ más m unidades; informalmente, $m + \mathcal{P}(\lambda(1-p))$.

(b) Se deduce que $\mathrm{E}[X \mid Y = m] = m + \lambda(1-p)$, que es por tanto también la expresión de la recta de regresión de X sobre Y, es decir, $n = m+\lambda(1-p)$. Nótese que, por construcción, la recta de regresión de Y sobre X tiene ecuación $m = np$.

La pendiente de la recta de regresión de X sobre Y es igual a $\mathrm{Cov}(X,Y)/\mathrm{Var}(Y) = 1$. Dado que $\mathrm{Var}(Y) = \lambda p$, obtenemos $\mathrm{Cov}(X,Y) = \lambda p$. Por otro lado, el coeficiente de correlación ρ verifica

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \frac{\lambda p}{\sqrt{\lambda \cdot \lambda p}},$$

luego $\rho = \sqrt{p}$.

(c) Condicionado por $X = n$, se tiene que Y tiene distribución $B(n,p)$ y por tanto $Y - X = n - X$ tiene distribución $B(n, 1-p)$. Razonando como en el apartado (a) se llega entonces a que $Y - X$ tiene distribución $\mathcal{P}(\lambda(1-p))$. Se tiene, para $n, m \geq 0$,

$$\begin{aligned}
\mathrm{P}\{X - Y = n, Y = m\} &= \mathrm{P}\{X = n+m, Y = m\} \\
&= \mathrm{P}\{Y = m \mid X = n+m\} \cdot \mathrm{P}\{X = n+m\} \\
&= \binom{n+m}{m} p^m (1-p)^n \cdot e^{-\lambda} \frac{\lambda^{n+m}}{(n+m)!} \\
&= e^{-\lambda(1-p)} \frac{(\lambda(1-p))^n}{n!} \cdot e^{-\lambda p} \frac{(\lambda p)^m}{m!} \\
&= \mathrm{P}\{X - Y = n\} \cdot \mathrm{P}\{Y = m\}.
\end{aligned}$$

Se concluye que Y y $X - Y$ son independientes.