MA-34A Probab. y Procesos Estocásticos **30 de Agosto, 1999**

Pauta Control No. 1

Prof. Cátedra: M. Kiwi, A. Moreira *Prof. Auxiliar: J. Barrera, E. Moreno, M. Saez*

PROBLEMA 1:

(i).- Llamemos A_i, $i = 1, ..., 6$, a los eventos "el sujeto es adherente del candidato i", para los i respectivos. Si suponemos que la encuesta estuvo bien hecha, entonces $\mathbb{P}(A_i) = \frac{f_i}{100}$, donde f_i es el porcentaje de apoyo obtenido por cada candidato. Llamemos ahora E al evento "vemos una letra al azar del apellido del candidato, y resulta ser una A". Lo que se nos pide calcular, entonces, es $\mathbb{P}(A_3|E)$. Es una aplicación típica del teorema de Bayes:

$$\mathbb{P}(A_3|E) = \frac{\mathbb{P}(E|A_3)\,\mathbb{P}(A_3)}{\sum_{i=1}^{6}\mathbb{P}(E|A_i)\,\mathbb{P}(A_i)} = \frac{\mathbb{P}(E|A_3)\,\frac{f_3}{100}}{\sum_{i=1}^{6}\mathbb{P}(E|A_i)\,\frac{f_i}{100}} = \frac{\mathbb{P}(E|A_3)\,f_3}{\sum_{i=1}^{6}\mathbb{P}(E|A_i)\,f_i}.$$

Falta determinar $\mathbb{P}(E|A_i)$; como la letra se escoge al azar, esta probabilidad corresponderá a la fracción que la cantidad de letras A representa dentro de las letras del apellido respectivo. Por lo tanto, $\mathbb{P}(E|A_1) = \frac{0}{4}$, $\mathbb{P}(E|A_2) = \frac{0}{6}$, $\mathbb{P}(E|A_3) = \frac{1}{5}$, $\mathbb{P}(E|A_4) = \frac{2}{7}$, $\mathbb{P}(E|A_5) = \frac{1}{5}$, $\mathbb{P}(E|A_6) = \frac{1}{5}$, y se tiene

$$\mathbb{P}(A_3|E) = \frac{\frac{1}{5}f_3}{\frac{1}{5}f_3 + \frac{2}{7}f_4 + \frac{1}{5}f_5 + \frac{1}{5}f_6} = \frac{f_3}{f_3 + \frac{10}{7}f_4 + f_5 + f_6} = \frac{36}{36+10+36+7} = \frac{36}{89}.$$

(ii).- Llamemos A, B, C, a los eventos de dejar fuera a todos los argentinos, brasileños y chilenos, respectivamente. Lo que se nos está pidiendo, entonces, es $\mathbb{P}(A \cup B \cup C)$. Estos eventos no son disjuntos, pues es posible dejar fuera simultáneamente a dos nacionalidades. Usando el principio de inclusión/exclusión, tenemos

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

Notemos de inmediato que $A \cap B \cap C$ significa dejar fuera a las tres nacionalidades; esto es imposible, por lo que $\mathbb{P}(A \cap B \cap C) = 0$. Ahora bien, para escoger 4 personas entre 12, existen $\binom{12}{4}$ posibilidades. Si se deja fuera a una nacionalidad, las posibilidades son $\binom{8}{4}$; por lo tanto $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \binom{8}{4}/\binom{12}{4}$. Por otro lado, si se quiere dejar fuera simultáneamente a dos nacionalidades, sólo habrá una forma de hacerlo (escogiendo a las 4 personas de la nacionalidad restante). Por lo tanto, $\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = 1/\binom{12}{4}$. Reemplazando arriba, tenemos

$$\mathbb{P}(A \cup B \cup C) = 3\frac{\binom{8}{4}}{\binom{12}{4}} - 3\frac{1}{\binom{12}{4}} = 3\frac{\binom{8}{4}-1}{\binom{12}{4}} = \frac{23}{55}.$$

PROBLEMA 2:

(i).- Este experimento es equivalente a lanzar dos dados, cada uno con 365 caras. El modelo adecuado viene dado por $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, donde $\Omega = \{(i,j) : i,j \in 1, \ldots, 365\}$. Hay muchos casos posibles de pares de eventos que cumplen con las propiedades pedidas; veremos algunos ejemplos para cada caso.

(i.1).- Dos casos triviales de eventos independientes son $E = F = \Omega$ y $E = F = \emptyset$. Otro ejemplo es el caso en que E y F corresponden a los eventos "la primera persona nació un primero de enero", y "la segunda persona nació un primero de enero", pues en ese caso $\mathbb{P}(E) = \mathbb{P}(F) = \frac{1}{365}$, y $\mathbb{P}(E \cap F) = \frac{1}{365^2}$. En general, dados dos subconjuntos E', F' en $\{1,\ldots,365\}$, los eventos E y F que corresponden a "el resultado para la primera persona está en E'", y "el resultado para la segunda persona está en F'", respectivamente, son independientes, pues $\mathbb{P}(E) = \frac{|E'|}{365}$, $\mathbb{P}(F) = \frac{|F'|}{365}$, y $\mathbb{P}(E \cap F) = \frac{|E'|\cdot|F'|}{365^2}$.

(i.2).- Cualquier par de eventos disjuntos de probabilidad no nula nos dará un ejemplo, pues en ese caso tenemos

$$\mathbb{P}(F|E) \;=\; \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)} \;=\; \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(E)} \;=\; 0 \;<\; \mathbb{P}(F)\;.$$

Un ejemplo concreto en este caso sería E y F correspondientes a los eventos "tienen cumpleaños en días distintos", y "tienen cumpleaños el mismo día," respectivamente. Claramente $\mathbb{P}(F|E) = 0$, mientras que $\mathbb{P}(F) = \frac{1}{365}$.

(i.3).- Cualquier par de eventos E y F tales que $E \subseteq F$ y $\mathbb{P}(F) < 1$ servirá de ejemplo, pues en este caso, como $E \cap F = E$, se tiene

$$\mathbb{P}(F|E) \;=\; \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)} \;=\; \frac{\mathbb{P}(E)}{\mathbb{P}(E)} \;=\; 1 \;>\; \mathbb{P}(F)\;.$$

Un ejemplo concreto sería el caso en que los eventos E y F corresponden a "la primera persona tiene cumpleaños en enero", y "la primera persona tiene cumpleaños en enero o febrero", respectivamente (en este caso, $\mathbb{P}(F|E) = 1$, $\mathbb{P}(F) = \frac{59}{365}$).

Se puede observar también que si E acarrea información negativa acerca de F, entonces E acarrea información positiva acerca de F^c. En efecto, si recordamos que E es la unión disjunta de $E \cap F$ y $E \cap F^c$, vemos que $\mathbb{P}(F^c|E) = \frac{\mathbb{P}(E\cap F^c)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E)-\mathbb{P}(E\cap F)}{\mathbb{P}(E)} = 1 - \frac{\mathbb{P}(E\cap F)}{\mathbb{P}(E)} > 1 - \mathbb{P}(F) = \mathbb{P}(F^c)$. De una manera semejante se puede ver que F acarrea información positiva acerca de E^c. Luego, basándonos en uno de los ejemplos de (i.2) se construyen ejemplos de (i.3) y viceversa.

(ii.1).- Recordemos primero que, en toda serie convergente, la cola de la serie tiende a cero (es decir, si $\sum_{n=1}^{\infty} a_n < \infty$, entonces $\lim_{m\to\infty} \sum_{n=m}^{\infty} a_n = 0$). Por lo tanto, en el caso presente se tiene que $\lim_{m\to\infty} \sum_{n=m}^{\infty} \mathbb{P}(C_n) = 0$. Ocupando además la desigualdad de Boole, obtenemos

$$0 \;\le\; \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n=m}^{\infty} C_n\right) \;\le\; \lim_{m\to\infty} \sum_{n=m}^{\infty} \mathbb{P}(C_n) \;=\; 0\,,$$

y concluimos.

(ii.2).- Definamos $D_m = \bigcup_{n=m}^{\infty} C_m$; en la parte anterior se probó que $\lim_{m\to\infty} \mathbb{P}(D_m) = 0$. Notemos ahora, además, que D_m es una familia decreciente; en efecto, a medida que m crece, D_m representa la unión de *menos* conjuntos. Se puede aplicar por lo tanto la propiedad de continuidad de $\mathbb{P}$, que nos dice que $\mathbb{P}(\lim_{m\to\infty} D_m) = \lim_{m\to\infty} \mathbb{P}(D_m) = 0$. Finalmente, vemos que

$$\mathbb{P}\left(\limsup_{n\to\infty} C_n\right) \;=\; \mathbb{P}\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} C_n\right) \;=\; \mathbb{P}\left(\bigcap_{m=1}^{\infty} D_m\right) \;=\; \mathbb{P}\left(\lim_{m\to\infty} D_m\right) \;=\; 0\,.$$

PROBLEMA 3:

(i).- El número mínimo de corridas de letras A es 1, lo que ocurre cuando todas las letras A's son contiguas. El número máximo de corridas ocurre cuando se pueden separar todas las letras A's en corridas de largo 1, lo que requiere que hayan al menos $a-1$ letras B's. Como $b \geq a-1$, lo anterior siempre se cumple, luego el número máximo de corridas de A's es a.

En conclusión el número de corridas posibles de letras A es $\{1,\ldots,a\}$.

(ii).- El conjunto de arreglos en que el número de letras A's y B's es a y b respectivamente es

$$\Omega \quad = \quad \{(x_1,\ldots,x_{a+b}) \,:\, x_i \in \{A,B\}, |\{x_i \,:\, x_i = A\}| = a\}\,.$$

Cada $\omega \in \Omega$ puede identificarse unívocamente con un subconjunto $S_\omega \subseteq \{1,\ldots,a+b\}$ de tamaño a, específicamente con

$$S_\omega \quad = \quad \{i \in \{1,\ldots,a+b\} \,:\, \omega_i = A\}\,.$$

Pero, la cantidad de subconjuntos de $\{1,\ldots,a+b\}$ de tamaño a, es $\binom{a+b}{a}$, luego $|\Omega| \;=\; \binom{a+b}{a}$.

(iii).- Distribuir a letras A en n corridas es equivalente a distribuir a bolas indistinguibles en n urnas distinguibles de forma que ninguna quede vacía (donde la cantidad de bolas en la i-ésima urna representa el largo de la i-ésima corrida). Luego, hay $\displaystyle\binom{a-1}{n-1}$ distribuciones distintas posibles.

(iv).- El modelo probabilista adecuado para estudiar el problema planteado es $(\Omega,\mathcal{F},\mathbb{P})$ donde Ω es como en (ii), $\mathcal{F} = \mathcal{P}(\Omega)$ (puesto que Ω es finito), y $\mathbb{P}$ es la medida sobre el espacio medible $(\Omega,\mathcal{F})$ que lo hace equiprobable.

Nos piden calcular $\mathbb{P}(E)$, donde $E = \{\omega \in \Omega \,:\, |\{i \,:\, \omega_i = A\}| = n, |\{i \,:\, \omega_i = B\}| = m\}$.

Como $\mathbb{P}(E) = |E|/|\Omega|$, y por (ii) sabemos que $|\Omega| = \binom{a+b}{a}$, sólo nos queda calcular $|E|$. Para ello observar que si $m = n \pm 1$, entonces hay una sola forma de juntar n corridas de A's y m corridas de B's de manera de obtener n y m corridas de A's y B's respectivamente (si $m = n+1$ hay que alternar las corridas de A's y B's comenzando con una corrida de B's, y si $n = m+1$ hay que comenzar con una corrida de A's). Luego, por el principio generalizado del conteo, $|E|$ es igual al producto del número de n corridas de A's posibles y el número de m corridas de B's posibles, i.e., por (iii),

$$|E| \quad = \quad \binom{a-1}{n-1}\binom{b-1}{m-1}\,.$$

Sigue que

$$\mathbb{P}(E) \quad = \quad \frac{\binom{a-1}{n-1}\binom{b-1}{m-1}}{\binom{a+b}{a}}\,.$$

(v).- Del punto anterior, se concluye que la probabilidad de observar $n = 5$ corridas de A's y $m = 6$

corridas de B's con $a = 5$ letras A's y $b = 11$ letras B's es

$$\frac{\binom{5-1}{5-1}\binom{11-1}{6-1}}{\binom{5+11}{5}} = \frac{\binom{10}{5}}{\binom{16}{5}} = \frac{3}{52} < 0.06\,,$$

i.e., es un evento poco probable, lo que indica que probablemente la anomalía no aparece en forma aleatoria y que la dolencia tiende a espaciarse en el tiempo.

También del punto anterior, se concluye que la probabilidad de observar $n = 1$ corridas de A's y $m = 2$ corridas de B's con $a = 5$ letras A's y $b = 11$ letras B's es

$$\frac{\binom{5-1}{1-1}\binom{11-1}{2-1}}{\binom{5+11}{5}} = \frac{\binom{10}{1}}{\binom{16}{5}} = \frac{5}{14 \cdot 13 \cdot 12} < 0.01\,,$$

i.e., es un evento muy raro que indicaría que las apariciones de la dolencia tienden a aglomerarse en el tiempo.