UNA FUNCIÓN PROBABILÍSTICA VOTIVA

Reproduce ***n*** veces un fenómeno aleatorio bipolar, con sólo dos resultados posibles, A o B. Cada realización unitaria específica es unívoca, con probabilidades respectivas p_i, q_i, siendo así

$$p_i + q_i = 1, i = 1,.., n$$

La probabilidad $P(h,k)$ de los resultados globales respectivos, esto es, de h resultados A y de k resultados B, independientemente de su orden de aparición, tiene la expresión multilineal

$$P(h,k) = \Sigma\, p_{i_1} p_{i_2} ... p_{i_h} q_{j_1} q_{j_2} ... q_{j_k} \quad \text{(I)}$$

siendo $\overline{A}_i \equiv \{p_{i_1} p_{i_2} ... p_{i_h}\}$ y $\overline{B}_j \equiv \{q_{j_1} q_{j_2} ... q_{j_k}\}$ subconjuntos parciales disjuntos del conjunto $\overline{N} \equiv \{1,2,...n\}$, formados de todas las maneras posibles, cumpliéndose, pues, las relaciones

$$\begin{cases} h + k = n \\ \overline{A}_i \cup \overline{B}_j = \overline{N} \end{cases}$$

El número de sumandos en (I) es, por tanto,

$$\nu = \binom{n}{h} = \binom{n}{k}$$

La correspondiente función dual generatriz, homogénea de grado n, en las variables x,y, se escribe

$$F(x,y) = \Sigma P(h,k) x^h y^k = \prod_1^n (p_i x + q_i y) \quad \text{(II)}$$

Conocidos los $2n$ valores p_i, q_i , realmente sólo n valores independientes, la programación de la determinación numérica computacional de todas las probabilidades $P(h,k)$, no presenta dificultad en una operativa recurrente, utilizando el último miembro de (II).

En cuanto a la aplicación simplificada de esta formulación en un entorno votivo dual, resultan admisibles algunas aproximaciones.

En primer lugar, la distribución de la población total n en tres agrupaciones distintas: cada una de los inclinados decididamente a uno de los resultados A o B, y los dudosos, esto es,

$$n = a + b + d$$

En segundo lugar, serían admisibles las siguientes asignaciones respectivas, comunes para todos los miembros de cada una de las poblaciones a y b,

$$\begin{cases} p_i(a) = 1, q_i(a) = 0 \\ p_j(b) = 0, q_j(b) = 1 \end{cases}$$

luego una expresión simplificada de (II) sería

$$F(x,y) = \Sigma P(h,k) x^h y^k = x^a y^b \prod_1^d (p_i x + q_i y) \quad \text{(III)}$$

En la práctica estadística, una estimación de los parámetros a,b, así como de valores medios de los $2d$ parámetros $p_i(a)$,$q_j(b)$, requeriría la realización de encuesta sobre una muestra adecuadamente equilibrada. La nueva aproximación, derivada de la simplificación asociada a valores medios, p y q, de las probabilidades elementales para el grupo dudoso, resulta

$$F(x,y) = \Sigma P(h,k)x^h y^k = x^a y^b \prod_1^d (px+qy)^d, \quad \begin{cases} 0 < p < 1 \\ 0 < q < 1 \end{cases}, \; p+q=1 \quad \text{(IV)}$$

y la probabilidad del suceso (*h*,*k*) admite las formas alternativas

$$P(h,k) = \binom{d}{h-a} p^{h-a} q^{k-b} = \binom{d}{k-b} p^{h-a} q^{k-b}, \quad \begin{cases} h \geq a, d \geq h-a \\ k \geq b, d \geq k-b \end{cases} \quad \text{(V)}$$

y también, la expresión simétrica en los parámetros, aplicando la forma combinatoria factorial,

$$P(h,k) = \frac{d!}{(h-a)!(k-b)!} p^{h-a} q^{k-b} \quad \text{(VI)}$$

Para valores suficientes de cada uno de los tres parámetros *h*-*a*,*k*-*b*,*d*, como pudiera ser ciertamente el caso en una votación popular, resulta aplicable la aproximación factorial de Stirling,

$$P(h,k) = \frac{d^d e^{-d}\sqrt{2\pi d}}{(h-a)^{h-a} e^{-h+a}\sqrt{2\pi(h-a)}.(k-b)^{k-b} e^{-k+b}\sqrt{2\pi(k-b)}} p^{h-a} q^{k-b} =$$

$$= \frac{d^{d+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}(h-a)^{h-a+\frac{1}{2}}(k-b)^{k-b+\frac{1}{2}}} p^{h-a} q^{k-b} \quad \text{(VII)}$$

Centrándonos ya en la probabilidad del suceso *h*=*k*, la probabilidad es cero si *n* es un número impar, y si *n* es un número par, esto es, con *h*=*k*=*n*/2, resulta

$$P\left(\frac{n}{2},\frac{n}{2}\right) = \frac{d^{d+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\left(\frac{n}{2}-a\right)^{\frac{n+1}{2}-a}\left(\frac{n}{2}-b\right)^{\frac{n+1}{2}-b}} p^{\frac{n}{2}-a} q^{\frac{n}{2}-b} \quad \text{(VIII)}$$

Una exigencia previa y en cualquier caso necesaria, del suceso *h*=*k*, son las desigualdades

$$a,b \leq \frac{n}{2},$$

pues la no verificación de una de ellas implica obviamente una probabilidad nula del mismo.

Resulta admisible, todavía, considerar una cota superior de (VIII) cumpliéndose *a fortiori*, además de las dos anteriores, las dos condiciones siguientes, y con la exigencia de la paridad adicional del parámetro, también número entero, *d*=*rn*,

$$\begin{cases} a = b = \dfrac{n-d}{2} = \dfrac{n(1-r)}{2} & \dfrac{2}{n} \leq r \leq 1 \\ p = q = \dfrac{1}{2} \end{cases} \quad \text{(IX)}$$

Luego, como decimos, para valores pares y suficientemente significantes de *n* y de *rn* se obtiene finalmente,

$$P^{+}\left(\frac{n}{2},\frac{n}{2}\right) \leq \frac{(rn)^{rn+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\left(\frac{rn}{2}\right)^{rn+1}}\left(\frac{1}{2}\right)^{rn} = \left(\frac{2}{\pi rn}\right)^{\frac{1}{2}} \quad \text{(X)}$$

Observamos que en (IX) se ha descartado el valor *d*=0, que implica también *r*=0, en cuyo caso no es aplicable la aproximación de Stirling en (VII), debiendo retraernos a la expresión original (III), que tiene la forma simplificada

$$F(x,y) = x^{\frac{n}{2}} y^{\frac{n}{2}}, \qquad \text{(XI)}$$

con el resultado obvio $P^+(n/2,n/2)=1$, cumpliéndose aquí, únicamente, la primera de las dos condiciones impuestas en (IX), mientras la segunda es inoperante por la inexistencia de sujetos de asignación de las probabilidades $p,q=1/2$ fijadas.
Si el número de dudosos, d, no se anula, se han figurado en (IX) las dos cotas teóricas, inferior y superior, del parámetro r. No obstante, son éstas, ambas condiciones, el número de dudosos 2 o n, ciertamente poco asumibles en una realidad sociológica, con valores significantes de n. Asimismo, como en el caso d=0, no resulta todavía aplicable la aproximación (X), en el caso de la cota inferior, para la factorial $d!=2$. En los supuestos con valores muy limitados de la población d, como son estos dos casos concretos d=0,2, no tiene significación la sustitución de la expresión general (VI) por la fórmula simplificada (X).

Finalmente, sustituyendo n=3030 en (X), resulta la tabla siguiente para distintos valores de r que pudieran considerarse fracciones consistentes de la población dudosa respecto de la población efectiva votante.

$r \rightarrow$	0,2	0,4	0,6	0,8
$P^+ \rightarrow$	3,24%	2,29%	1,87%	1,62%

Los valores de r seleccionados verifican la condición de la paridad del parámetro d. Se excluyen valores r < 0,2, y también valores r > 0,8, que, como se ha anticipado, no se consideran consistentes con una realidad sociológica. En particular, para r =1 se obtiene el valor P^+=1,45%, que ha sido indicado por comentaristas como una probabilidad producida, en una determinada situación de encrucijada política. Obviamente no han tenido representación en el análisis, ni participación alguna en los resultados finales los no votantes y votantes correspondientes a abstenciones y/o votos nulos.

En resumen, la representación (r,P^+) de los valores $P^+(n/2,n/2)$ es la de una función discontinua puntual, de manera que para valores impares de n resultan puntos situados en el eje coordenado P^+=0, mientras para valores pares de n resultan puntos situados en la rama superior de tipo hiperbólico de la cúbica (X), asintótica con el eje P^+=0. Esta disposición está limitada en sentido creciente por el punto (r=1,P^+=1,45%), correspondiente a la situación en la que todos los votantes serían en origen dudosos y con ambas probabilidades equivalentes, p,q=1/2. El valor 1,45 %, con carácter de probabilidad máxima en este supuesto, es así, paradójicamente, el mínimo de los correspondientes a las restantes probabilidades máximas para valores inferiores de r, como se han indicado en la tabla.
Asimismo, la disposición puntual sobre la rama asintótica estará limitada en sentido decreciente por el valor inferior de r que permita la aplicación de la aproximación de Stirling a la factorial $(rn/2)!$, en cuyos casos debemos retraernos a la expresión (VI), cumpliéndose de nuevo las condiciones (IX), esto es,

$$P^+\left(\frac{n}{2},\frac{n}{2}\right) = \frac{(rn)!}{\left[\left(\frac{rn}{2}\right)!\right]^2 2^{rn}},$$

aunque su aplicación a los correspondientes enteros pares muy limitados de $d=rn$ no resulta significante en el contexto votivo de referencia.
Por último, como ya se ha mencionado, la representación funcional de probabilidades máximas indicada, se completa con el punto (r=0,P^+=1) en el supuesto d=0, en el que no existen votantes dudosos y se cumple además la relación $a = b = n/2$.