DISTRIBUCIONES DE VARIABLES DISCRETAS

Jorge Dagnino S.[1]

- La distribución normal es una distribución teórica de variables aleatorias continuas. Hay distribuciones de muestreo de variables discretas de las cuales las más usadas en medicina son la binomial y la de Poisson.
- La binominal da la probabilidad de que un resultado específico ocurra dentro de un número dado de pruebas independientes. Por ejemplo, si la mortalidad de un cuadro es de 40%, permite calcular cuál es la probabilidad que 3 pacientes consecutivos sobrevivan.
- La de Posisson es apropiada para estudiar eventos de muy escasa ocurrencia.

En otro artículo se describió la distribución Normal como la principal de las distribuciones teóricas de variables aleatorias continuas. En este haremos lo mismo con dos distribuciones teóricas de variables discretas, completándolo con una somera descripción de algunas distribuciones de muestreo. Si bien la descripción adecuada supone el conocimiento del cálculo de probabilidades y su manejo matemático, intentaremos una descripción conceptual pensando que su conocimiento es útil pues aparecen con alguna frecuencia en la literatura médica. Vale la pena reiterar que son distribuciones teóricas de la variable aleatoria de interés y que su utilidad reside en que se puede comparar la distribución de los datos observados con ese modelo.

DISTRIBUCIÓN BINOMIAL

La distribución más simple para datos discretos de mayor uso en medicina es cuando existen solo dos posibilidades, por ejemplo vivos o muertos, donde estamos interesados en la proporción de una de ellas.

La distribución binomial da la probabilidad de que un resultado específico ocurra dentro de un número dado de pruebas independientes. En la Figura 1 se grafica la probabilidad de ocurrencia de una situación simple: lanzar al aire una moneda no sesgada donde la variable de interés es que salga una cara, una variable discreta, que sólo tiene dos posibilidades, sí o no, 1 ó 0. Con un lanzamiento, sólo hay dos resultados posibles y la probabilidad de tener o de no tener una cara es de 0,5. Con dos lanzamientos, la probabilidad de tener una cara sigue siendo de 0,5 mientras que la probabilidad de tener dos caras o ninguna cara es del 0,25. En el caso de 4 lanzamientos, la probabilidad de tener cuatro caras es sólo del 0,0625.

El nombre de distribución binomial viene de que la distribución en particular es determinada por dos parámetros. Uno es el número de pruebas independientes, designado por la letra n (número de veces que se lanza la moneda) y el otro es la probabilidad del evento en particular, designado por la letra π aunque con frecuencia aparece citado como p. Vale la pena subrayar que la distribución supone eventos **independientes** y que cada uno de ellos tiene la **misma** probabilidad de ocurrir.

Las probabilidades de ocurrencia con diferentes valores de n y de π pueden calcularse con la fórmula general de la distribución binomial o bien puede buscarse en tablas:

$$Prob\ (d\ eventos) = \frac{n!}{d!\ (n - D)!}\ \pi^{d}\ (1 - \pi)^{n - d}$$

El cálculo con más que unos pocos eventos puede ser complicado para hacerlo manualmente. Por ejemplo, la supervivencia de un shock séptico es de 40%. Si se toman 4 de tales pacientes, ¿cual es la probabilidad que 3 sobrevivan? En este ejemplo π es = 0,4 y n es = 3.

[1] Profesor Titular, División de Anestesiología, Pontificia Universidad Católica de Chile.

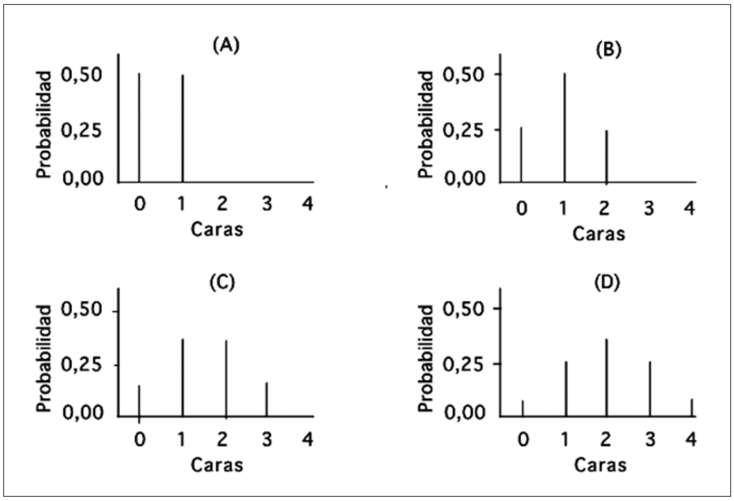

Figura 1. Probabilidades de tener caras en un determinado número de lanzamientos. El número de estos es 1 en (A), 2 en (B), 3 en (C), y 4 en (D).

$$Prob\ (3) = \frac{4!}{3\ (4-3)!}\ 0{,}4^3\ (1 - 0{,}4)^{4-3}$$

$$Prob\ (3) = \frac{24}{6} \times 0{,}064\ \times\ 0{,}6 = 0{,}15$$

Si el número es suficientemente grande (cuando nπ y n-nπ son ≥ 10) y como resultado del teorema del límite central, la distribución binomial se acerca mucho a una Normal, con la misma media y desviación estándar, lo que simplifica el análisis, permite calcular intervalos de confianza y se denomina la aproximación normal a la distribución binominal.

DISTRIBUCIÓN DE POISSON (Figura 2)

También se trata de una familia de distribuciones, pero a diferencia de la anterior, está determinada por un solo parámetro, la media, ya que la desviación estándar es igual a la media. Es apropiada para estudiar eventos de muy escasa ocurrencia por lo que también se la denomina como la distribución de los eventos raros. Da la probabilidad de que un resultado ocurra un número específico de veces cuando la cantidad de pruebas es grande y la probabilidad de ocurrencia es pequeña. Por ejemplo, siempre que en un medio cualquiera se encuentre un número elevado de partículas mezcladas al azar en dicho medio, el número de partículas en cada porción pequeña del medio, en relación al total del mismo, sigue una ley de Poisson: bacterias por cm^3 de aire, eritrocitos o granulocitos por cm^3 de sangre, conteos por minuto, etc. Puede usarse para calcular, por ejemplo, el número de camas de intensivo que necesita un hospital o el número de telefonistas que se necesitan para atender los llamados. Puede ser derivada matemáticamente como un caso particular de la distribución binomial, asumiéndose en este caso que el número de ensayos es indefinidamente grande mientras que la probabilidad de cada evento es muy rara, aproximándose a cero. La ventaja de usar esta distribución por sobre la binomial, cuando se cumplen las presunciones de que $n \geq 20$ y $p \leq 0{,}05$, es que es más fácil de calcular ya que se necesita conocer sólo un parámetro.

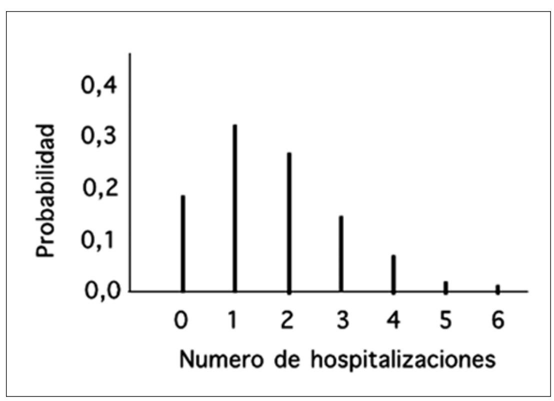

Figura 2. Aparece la distribución de Poisson cuando la media es de 1,69, en este caso el promedio de hospitalizaciones para un grupo de pacientes. Con sólo ese parámetro se calculan probabilidades del número de hospitalizaciones: de 0,18 que ninguna, de 0,31 que 2 veces o 0,006 para que se hospitalice 6 veces.

REFERENCIAS

1. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall, 1991.
2. Bland M. An Introduction to Medical Statistics. 3rd Ed, Oxford: OUP, 2006.
3. Dawson-Saunders B, Trapp RG. Bioestadística Médica. México D.F: Manual Moderno, 1993.
4. Fabregat López J, García Rojo B, Cook TM. Comments on a statistical note on the Poisson (binomial) distribution. Anaesthesia 2009; 64: 571-572.
5. Glantz SA. Primer of Biostatistics. 3a edición, New York: McGraw-Hill, 1992.
6. Pandit JJ. Gambling with ethics? A statistical note on the Poisson (binomial) distribution. Anaesthesia 2008; 63: 1163-1165.
7. Portney LG, Watkins MP. Foundations of Clinical Research. Applications to practice. 2nd ed. Upper Saddle River: Prentice-Hall, 2000.

Correspondencia a:
Dr. Jorge Dagnino S.
jdagnino@med.puc.cl