La distribución binomial

Si se consideran conjuntamente n repeticiones independientes de un suceso que presenta dos alternativas con probabilidades p y q, respectivamente, se obtiene una distribución con n+1 clases que se denomina distribución binomial (o serie binomial), ya que las frecuencias de las distintas clases se corresponden con los términos del desarrollo del binomio elevado a una potencia: $(p + q)^n$

Las situaciones de este tipo son muy frecuentes en Biología, y especialmente en Genética.

Ejemplo 1.- Un heterozigoto para cualquier gen (A,a) forma gametos con el alelo A (probabilidad p= 1/2) o el alelo a (probabilidad q= 1I2). Un heterocigoto para cuatro genes con dominancia completa que se transmiten independientemente, AaBbCcEe, formará 5 gametos diferentes en lo que se refiere al número de alelos dominantes (D) y recesivos (R) que contienen (cada clase se puede definir por un solo componente, por ejemplo el número de alelos dominantes: clase 0: 0D+4R; clase 1: 1D+3R; etc.).

Ejemplo 2.- Supongamos que la probabilidad de que se produzca una mutación en un gen concreto en una división bacteriana es $p= 10^{-6}$ (la probabilidad de que no se produzca la mutación es $q= 1-10^{-6}$). Si en un cultivo bacteriano se han producido 10^7 divisiones, se podrán producir desde 0 hasta 10^7 mutaciones (10^7+1 clases).

Ejemplo 3.- Un descendiente del cruzamiento entre dos heterozigotos para un gen con dominancia completa (Aa x Aa) podrá tener fenotipo A (probabilidad p= 3/4) o fenotipo a (probabilidad q= 1/4). Un grupo de 5 descendientes de ese cruzamiento podrá ser de 6 formas diferentes en lo que respecta al número de individuos con uno u otro fenotipo (clase 0= 0 individuos A + 5 a, clase 1= 1A+4a, clase 2= 2A+3a, clase 3= 3A+2a, clase 4= 4A+1a, clase 5= 5A+0a).

Podemos analizar este último ejemplo con más detalle. Si hacemos distinción entre los descendientes denominándoles 1º, 2º, 3º, 4º y 5º (por ejemplo, en función de su edad, o por el orden en que se analizan, etc.), vemos que, verdaderamente, hay 32 descendencias posibles de 5 individuos en las que cualquiera de ellos pueda tener uno de los fenotipos A o a. En la tabla de la derecha aparecen esas descendencias con la probabilidad de cada una de ellas. Para cada descendiente, considerado individualmente, la probabilidad de fenotipo A es 3/4 y la de fenotipo a es 1/4. Como los distintos individuos de una descendencia pueden considerarse sucesos independientes (un individuo puede ser A o a independientemente de como sean los demás), la probabilidad de una descendencia compuesta por i individuos A y n-i individuos a en un orden específico es $(3/4)^i$ x $(1/4)^{n-i}$ (vea conceptos básicos sobre probabilidad y recuerde que un número elevado a 0 = 1).

Las 32 descendencias indicadas en la tabla de la derecha constituyen 6 grupos de 1, 5, 10, 10, 5 y 1 combinaciones posibles que cumplen, respectivamente, las condiciones de las clases 0 a 5 descritas anteriormente. Los números 1, 5, 10, 10, 5 y 1, son los coeficientes del desarrollo del binomio $(p + q)^n$, para n=5. El triángulo de Pascal (o de Tartaglia) es un juego matemático en el que se representan y establecen relaciones entre los coeficientes binomiales para distintos valores de n. En general, en una serie binomial con n repeticiones (o n+1 clases) el número de combinaciones que cumplen la condición de la clase i es:

$$\binom{n}{i} = \frac{n!}{i! \times (n-i)!}$$

en donde n es el número de repeticiones (en este ejemplo es el número de descendientes, n= 5); i es es número de la clase (en este ejemplo i es el número de descendientes de fenotipo A y n-i el número de desdendientes de fenotipo a); ! es el símbolo de factorial (recuerde que n!=nx(n-1)x(n-2)x(n-3)x..., y que 0!=1).

Ahora, podemos calcular las probabilidades de las clases 0 a 5 (véase tabla a la derecha). Hay una sola combinación de descendientes que cumple las condiciones de la clase 0 (0A+5a), por lo que puede concluirse que la probabilidad de esta clase es $f_0= (3/4)^0 x (1/4)^5$. Como hay 5 combinaciones diferentes que cumplen las condiciones de la clase 1 (1A+4a), la probabilidad de esta clase será la suma de las probabilidades de esas cinco combinaciones. Como todas esas combinaciones tienen la misma probabilidad, la probabilidad de la clase 1 es $f_1= 5x(3/4)^1x(1/4)^4$. Siguiendo este razonamiento, las probabilidades de todas las clases son:

$f_0= 1x(3/4)^0x(1/4)^5$; $f_1= 5x(3/4)^1x(1/4)^4$; $f_2= 10x(3/4)^2x(1/4)^3$; $f_3= 10x(3/4)^3x(1/4)^2$; $f_4= 5x(3/4)^4x(1/4)^1$; $f_5= 1x(3/4)^5x(1/4)^0$

Efectivamente, las probabilidades (o frecuencias) de las distintas clases de esta distribución son los correspondientes términos del desarrollo del binomio: $(3/4 + 1/4)^5$.

En resumen, en una distribución binomial generada a partir de n repeticiones de un suceso con dos alternativas con probabilidades p y q, respectivamente, la frecuencia de la clase i es:

$$f_i = \frac{n!}{i! \times (n-i)!} \times p^i \times q^{n-i}$$

Descendientes 1º	2º	3º	4º	5º	Probabilidad	
a	a	a	a	a	$(3/4)^0$ x $(1/4)^5$	Clase 0 (0 A + 5 a) Probabilidad: $f_0= 1x(3/4)^0x(1/4)^5$
A	a	a	a	a	$(3/4)^1$ x $(1/4)^4$	Clase 1 (1 A + 4 a) Probabilidad: $f_1= 5x(3/4)^1x(1/4)^4$
a	A	a	a	a	$(3/4)^1$ x $(1/4)^4$	
a	a	A	a	a	$(3/4)^1$ x $(1/4)^4$	
a	a	a	A	a	$(3/4)^1$ x $(1/4)^4$	
a	a	a	a	A	$(3/4)^1$ x $(1/4)^4$	
A	A	a	a	a	$(3/4)^2$ x $(1/4)^3$	Clase 2 (2 A + 3 a) Probabilidad: $f_2= 10x(3/4)^2x(1/4)^3$
A	a	A	a	a	$(3/4)^2$ x $(1/4)^3$	
A	a	a	A	a	$(3/4)^2$ x $(1/4)^3$	
A	a	a	a	A	$(3/4)^2$ x $(1/4)^3$	
a	A	A	a	a	$(3/4)^2$ x $(1/4)^3$	
a	A	a	A	a	$(3/4)^2$ x $(1/4)^3$	
a	A	a	a	A	$(3/4)^2$ x $(1/4)^3$	
a	a	A	A	a	$(3/4)^2$ x $(1/4)^3$	
a	a	A	a	A	$(3/4)^2$ x $(1/4)^3$	
a	a	a	A	A	$(3/4)^2$ x $(1/4)^3$	
A	A	A	a	a	$(3/4)^3$ x $(1/4)^2$	Clase 3 (3 A + 2 a) Probabilidad: $f_3= 10x(3/4)^3x(1/4)^2$
A	A	a	A	a	$(3/4)^3$ x $(1/4)^2$	
A	A	a	a	A	$(3/4)^3$ x $(1/4)^2$	
A	a	A	A	a	$(3/4)^3$ x $(1/4)^2$	
A	a	A	a	A	$(3/4)^3$ x $(1/4)^2$	
A	a	a	A	A	$(3/4)^3$ x $(1/4)^2$	
a	A	A	A	a	$(3/4)^3$ x $(1/4)^2$	
a	A	A	a	A	$(3/4)^3$ x $(1/4)^2$	
a	A	a	A	A	$(3/4)^3$ x $(1/4)^2$	
a	a	A	A	A	$(3/4)^3$ x $(1/4)^2$	
a	A	A	A	A	$(3/4)^4$ x $(1/4)^1$	Clase 4 (4 A + 1 a) Probabilidad: $f_4= 5x(3/4)^4x(1/4)^1$
A	a	A	A	A	$(3/4)^4$ x $(1/4)^1$	
A	A	a	A	A	$(3/4)^4$ x $(1/4)^1$	
A	A	A	a	A	$(3/4)^4$ x $(1/4)^1$	
A	A	A	A	a	$(3/4)^4$ x $(1/4)^1$	
A	A	A	A	A	$(3/4)^5$ x $(1/4)^0$	Clase 5 (5 A + 0 a) Probabilidad: $f_5= 1x(3/4)^5x(1/4)^0$

La probabilidad de que el heterozigoto AaBbCcEe del ejemplo 1, forme un gameto con un alelo dominante y tres recesivos (1D+3R) es:

$$f_1 = \frac{4!}{1! \times 3!} \times \left(\frac{1}{2}\right)^1 \times \left(\frac{1}{2}\right)^3$$

Y la probabilidad de que en el cultivo bacteriano del ejemplo 2 no se hayan producido mutaciones es:

$$f_0 = \frac{10^7!}{0! \times 10^7!} \times \left(10^{-6}\right)^0 \times \left(1-10^{-6}\right)^{10^7}$$

(véase la distribución de Poisson para situaciones como la de este ejemplo).

Triángulo de Pascal (o de Tartaglia)

n		Suma
1	1 1	2
2	1 2 1	4
3	1 3 3 1	8
4	1 4 6 4 1	16
5	1 5 10 10 5 1	32
6	1 6 15 20 15 6 1	64
7	1 7 21 35 35 21 7 1	128
8	1 8 29 56 70 56 28 8 1	256
9	1 9 36 84 126 126 84 36 9 1	512
10	1 10 45 120 210 252 210 120 45 10 1	1024
...	....	...
n	$\binom{n}{0}$ $\binom{n}{1}$ $\binom{n}{2}$ ... $\binom{n}{i}$ ... $\binom{n}{n-2}$ $\binom{n}{n-1}$ $\binom{n}{n}$	2^n

Cada fila del triángulo empieza y termina con 1 y el resto de los números se generan sumando los dos numeros más próximos de la fila superior