CAPÍTULO IV.

LA DISTRIBUCIÓN BINOMIAL

Si nos cae un ejercicio de este tipo…

… sería un regalo irrechazable.

NO DIGO MÁS

1. DISTRIBUCIONES DE PROBABILIDAD

1.1. Distribuciones de probabilidad: Media, varianza y desviación típica.

Cuando se analiza un fenómeno observable aparece una serie de resultados que han de ser tratados convenientemente, de manera que se puedan comprender mejor tanto los resultados como la característica objeto de estudio correspondiente a dicho fenómeno. Para este fin ya sabes realizar una primera descripción de los datos, histograma de frecuencias absolutas o relativas y polígono de frecuencias absolutas o acumuladas, y determinar sus parámetros: la media, varianza, desviación típica…

En ese caso, los propios resultados del experimento son numéricos como en el caso en el que se mide la velocidad de un vehículo, o la altura de un individuo. En cambio, en otras ocasiones, los resultados del experimento no proporcionan dicha información adecuadamente, como puede ser en el caso de los juegos de azar (ruleta, lotería, etc.).

En estos casos, se puede utilizar una variable aleatoria, que es una función que asigna un número real a cada resultado posible del espacio muestral del fenómeno bajo estudio. Por ejemplo, en los juegos de azar se puede asignar a cada resultado la ganancia o pérdida que produce en el jugador.

Las variables aleatorias se denotan mediante una letra mayúscula y pueden ser discretas (cuando pueden tomar un número finito o infinito numerable de valores) o continuas (cuando pueden tomar cualquier valor dentro de un rango).

En cuanto a las variables aleatorias discretas, éstas son las que pueden tomar un número finito o infinito numerable (como el conjunto N de los números naturales) de valores. Dado que la variable aleatoria puede tomar diferentes valores dependiendo de los resultados del experimento aleatorio al que está asociado, su valor no se podrá predecir de manera exacta. Así pues, para describir una variable aleatoria es necesario conocer su distribución de probabilidad.

Conocer la distribución de probabilidad de una variable aleatoria X discreta consiste en asignar una probabilidad a cada uno de los resultados posibles de dicha variable aleatoria. Es decir, se trata de saber calcular o asignar los valores $P[X = x]$, para todos los posibles valores x que puede tomar la variable aleatoria X.

Actividad resuelta

- *Se lanzan dos monedas y contamos el número de caras. La distribución de probabilidad es:*

Al hacer un diagrama en árbol calculamos las probabilidades:

Número de caras (x):	0	1	2
Probabilidad ($p(x)$):	1/4	1/2	1/4
Función de distribución $F(x)$:	1/4	3/4	4/4

Por un lado, tenemos la función $p(x)$, que es la probabilidad puntual o función de cuantía o función masa de probabilidad.

Por otro lado, podemos calcular lo que sería equivalente a las frecuencias acumuladas. La función $F(x)$, a la que se denomina función de distribución, nos indica la probabilidad de que $F(x) = P(X \leq x)$, es decir, calcula la probabilidad de que se tomen valores menores a x.

Autora: Raquel Caro
Revisores: Luis Carlos Vidal, Leticia González y Álvaro Valdés
Ilustraciones: Banco de Imágenes de INTEF

La tabla que hemos presentado es una **distribución de probabilidad**, donde hemos definido una función que asigna a la variable aleatoria x una probabilidad:

$$x \to p(x)$$

y es el resultado que nos ayudará a hacer predicciones sobre un experimento aleatorio.

También podemos representar la tabla anterior mediante un histograma para la función de cuantía, en el que las áreas de cada rectángulo son ahora probabilidades, en lugar de frecuencias relativas, y podemos representar con una línea la función de distribución.

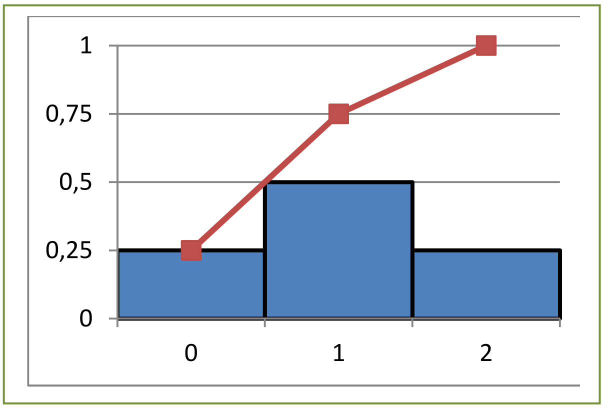
Observamos que siempre se verifican las siguientes propiedades.

Función de probabilidad o función de cuantía:

1) $p(x) \geq 0$
2) $\Sigma\, p(x) = 1$.

Función de distribución:

1) $0 \leq F(x) \leq 1$
2) $F(x)$ es una función creciente
3) $F(x_{Máximo}) = 1$.



Actividades propuestas

1. Se lanzan 3 monedas y contamos el número de caras que salen. Haz un diagrama en árbol. Escribe en una tabla la función de cuantía y la función de distribución. Representa la función de cuantía en un histograma y con una línea la función de distribución.

2. Se lanzan 2 dados y contamos el número de 5 que aparecen. Haz un diagrama en árbol, escribe en una tabla la función de cuantía y la función de distribución, y represéntalas gráficamente.

Actividad resuelta

- *Se lanzan dos dados. Por cada 5 que aparezca ganas 20 euros y pierdes 10 euros en caso contrario. ¿Te conviene ese juego? ¿Cuánto esperas ganar o perder en 36 jugadas?*

En cada lanzamiento puedes perder 10 euros o ganar 40 euros o ganar 20 euros. Esos son los valores de una variable aleatoria que podemos llamar ganancia, cuyas probabilidades calculamos, haciendo un diagrama de árbol, y escribimos en la siguiente tabla:

Ganancia (euros) (x):	–10	20	40
Probabilidad ($p(x)$):	25/36	10/36	1/36

Por tanto, en 36 jugadas "esperamos" perder 10 euros en 25 de ellas, ganar 20 euros en 10, y ganar 40 euros en una jugada.

Ahora la variable aleatoria, que es discreta, es la ganancia.

Podemos calcular la **media** o **esperanza matemática** $E(x)$ con la expresión:

$$E(x) = \sum_i x_i \cdot p(x_i)$$

La **esperanza matemática** es una media teórica, de ahí el nombre de esperanza. Indica que si repetimos el experimento varias veces se espera que la media de los valores obtenidos se aproxime a esta esperanza calculada.

Para distinguir la media de una distribución de frecuencias de la esperanza de una distribución de probabilidad, se suele utilizar para las frecuencias la letra m o $\bar{x}$, mientras que para la esperanza matemática se utiliza la letra griega μ (que se lee "mu") o $E(x)$.

Un juego es equitativo si la esperanza matemática de la ganancia es 0, es ventajoso si $E(x) > 0$, y es desventajoso si $E(x) < 0$.

En la actividad propuesta anteriormente calculamos la media o esperanza matemática:

Esperanza matemática = $E(x)$ = –10(25/36) + 20(10/36) + 40(1/36) = (–250 + 200 + 40)/36 = –10/36. Como $E(x) < 0$, el juego es desventajoso.

Esto sería como calcular lo que "esperas" perder en 36 jugadas.

Actividades propuestas

3. Se lanzan 3 monedas. Por jugar nos cobran 1 euro, y por cada cara que aparezca ganamos 1 euro. Escribe una distribución de probabilidad y representa el histograma. ¿Cuánto esperas ganar o perder en 100 lanzamientos?

4. Una persona apuesta 10 euros a un juego de tirar una moneda y que salga cara o cruz (o similar). Si gana se retira y deja de jugar. Si pierde, apuesta el doble, 20 euros. Si gana se retira. Si pierde apuesta el doble, 40 euros. Y así sucesivamente. Con esta estrategia siempre acaba ganando 10 euros, pero tiene un defecto, ¡que no lleve suficiente dinero para seguir jugando hasta ganar! Imagina que lleva 500 euros. A) Haz un diagrama de árbol y calcula todas las posibilidades y sus probabilidades. B) La distribución de probabilidad: Ganancia (x) $\rightarrow$ Probabilidad (x). C) ¿Es un juego ventajoso? ¿Y para nuestro jugador? D) Calcula la probabilidad de ganar 10 euros y la de perder 500 euros.

5. Lanzamos dos dados. Si apostamos al 7 y sale, recuperamos tres veces lo apostado. Si apostamos que sale menor que 7 y sale, recuperamos lo apostado, y lo mismo si apostamos que sale mayor que 7. ¿Cuál es la mejor estrategia?

Por ejemplo, asegurar un coche a todo riesgo es un juego desventajoso, pero nos asegura que no habrá pérdidas grandes.

Autora: Raquel Caro
Revisores: Luis Carlos Vidal, Leticia González y Álvaro Valdés
Ilustraciones: Banco de Imágenes de INTEF

Para saber si los valores son próximos a la media, ya sabes que se utiliza la **desviación típica**. Lo mismo en las distribuciones de probabilidad. Para medir esa dispersión se utiliza la expresión:

$$\sigma^2 = \sum_i (x_i - \mu)^2 \cdot p(x_i) = E(x^2) - E^2(x).$$

$$\sigma = \sqrt{E(x^2) - E^2(x)}$$

Ahora, cuando se refiere a distribuciones de probabilidad se utiliza la letra griega σ para indicar la desviación típica, y σ^2 para la varianza. Recuerda, con frecuencias utilizábamos s y s^2.

La desviación típica y la varianza son teóricas ya que se refieren a una distribución de probabilidad.

Las propiedades que verificaba la media y la desviación típica de las frecuencias se continúan verificando para la esperanza matemática y la desviación típica de las probabilidades.

DISTRIBUCIÓN DE PROBABILIDAD. Distribución y tipos de distribuciones. Estadística. Yasuri May

https://www.youtube.com/watch?v=AJt01UoVWZc

Actividades resueltas

- *Se lanza una moneda 3 veces y contamos el número de caras. Calcula la desviación típica de la distribución de probabilidad.*

Hacemos un diagrama de árbol y comprobamos que la distribución de probabilidad es:

Número de caras (x):	0	1	2	3
Probabilidad $p(x)$:	1/8	3·(1/8)	3·(1/8)	1/8

Completamos la tabla con las filas siguientes:

					Suma
$x \cdot p(x)$**:**	0	3/8	6/8	3/8	3/2
x^2**:**	0	1	4	9	
$x^2 \cdot p(x)$**:**	0	3/8	12/8	9/8	3

De donde deducimos que:

$$E(x) = \mu = 3/2.$$

$$E(x^2) = 3.$$

$$\sigma^2 = E(x^2) - E^2(x) = 3 - \left(\frac{3}{2}\right)^2 = 3 - \frac{9}{4} = \frac{3}{4}.$$

$$\sigma = \sqrt{\frac{3}{4}} = \sqrt{0'75} = 0'87$$

$$\sigma = \sqrt{\frac{3}{4}} = \sqrt{0.75} = 0.87 =$$

Autora: Raquel Caro

Revisores: Luis Carlos Vidal, Leticia González y Álvaro Valdés

Ilustraciones: Banco de Imágenes de INTEF

1.2. Distribución binomial

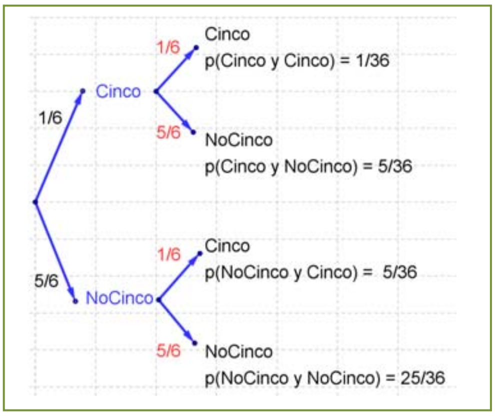
Este apartado está dedicado a describir y caracterizar matemáticamente algún modelo utilizado para variables aleatorias discretas que se repiten con frecuencia en las aplicaciones prácticas. Nos referimos al modelo de probabilidad discreto con más aplicaciones prácticas: la distribución binomial.

Antes de estudiarla vamos a ver dos ejemplos que ya conoces:

Actividad resuelta

- *Se lanza un dado. Llamamos "éxito" a que salga un* 5*. Escribe la distribución de probabilidad.*

Número de "*éxitos*":	0	1
Probabilidad:	5/6	1/6



- *Lanzamos ahora* 2 *dados.*

Número de "*éxitos*":	0	1	2
Probabilidad:	25/36	5/36 + 5/36 = 10/36	1/36

Observa que:

$$p(0)=\left(\frac{5}{6}\right)^2=1\cdot\left(\frac{5}{6}\right)^2=\binom{2}{2}\cdot\left(\frac{5}{6}\right)^2$$

$$p(1)=2\cdot\left(\frac{1}{6}\right)\cdot\left(\frac{5}{6}\right)=\binom{2}{1}\cdot\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)$$

$$p(2)=\left(\frac{1}{6}\right)^2=1\cdot\left(\frac{1}{6}\right)^2=\binom{2}{0}\cdot\left(\frac{1}{6}\right)^2$$ donde $\binom{2}{0}$, $\binom{2}{1}$, $\binom{2}{2}$ son los números combinatorios.

Actividad resuelta

- *Se lanza una moneda 3 veces. Llamamos "éxito" a que salga cara. Escribe la distribución de probabilidad.*

Dibujamos el diagrama en árbol y calculamos las probabilidades:

Número de "*éxitos*":	0	1	2	3
Probabilidad:	1/8	3·(1/8)	3·(1/8)	1/8

Observa que:

$$p(0)=\frac{1}{8}=\left(\frac{1}{2}\right)^3=1\cdot\left(\frac{1}{2}\right)^3=\binom{3}{0}\cdot\left(\frac{1}{2}\right)^3$$

$$p(1)=\frac{3}{8}=3\cdot\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)^2=\binom{3}{1}\cdot\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2$$

$$p(2)=\frac{3}{8}=3\cdot\left(\frac{1}{2}\right)^3=\binom{3}{2}\cdot\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)$$

$$p(3)=\frac{1}{8}=\left(\frac{1}{2}\right)^3=1\cdot\left(\frac{1}{2}\right)^3=\binom{3}{3}\cdot\left(\frac{1}{2}\right)^3$$ donde $\binom{3}{0}$, $\binom{3}{1}$, $\binom{3}{2}$, $\binom{3}{3}$ son los números combinatorios.

Los dos ejemplos anteriores son de una distribución binomial.

La **distribución binomial** se caracteriza porque puede ser interpretada como un experimento en el que se consideran sucesos dicotómicos, es decir, el de tener "éxito" y el de no tenerlo, de probabilidades p y $q = 1 - p$ respectivamente. Se realiza el experimento n veces, todas independientes y con la misma probabilidad p.

Se representa a la distribución binomial de parámetro p, probabilidad de "éxito", para n, número de pruebas como $B(n, p)$.

En los ejemplos anteriores hemos obtenido que la probabilidad de tener x éxitos en n pruebas repetidas en una distribución binomial $B(n, p)$ es:

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$$

Esta distribución es importante pues aparece en muchas aplicaciones.

Actividades propuestas

6. Se ha comprobado que la distribución de probabilidad del sexo de un recién nacido es:

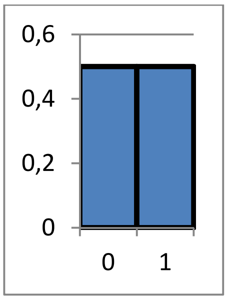
Sexo del recién nacido:	chica	chico
Probabilidad:	0.485	0.515

En un hospital van a nacer hoy 10 bebés. Escribe la expresión de probabilidad de que nazcan 7 chicas.

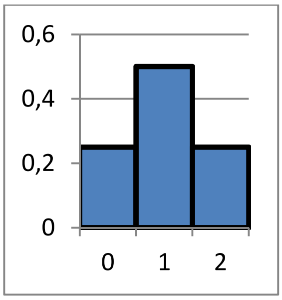
7. Se estima que el porcentaje de hogares que utiliza una determinada marca de tomate frito es del 12 %. En una muestra de 20 hogares, ¿qué probabilidad hay de encontrar entre 6 y 15 que lo utilicen? (No lo calcules, sólo plantea como lo calcularías).

Actividades resueltas

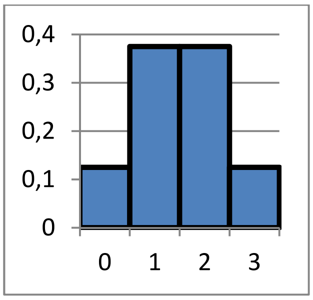
- *Volvemos al problema de lanzar una moneda n* veces. *Dibujamos los histogramas de la distribución de probabilidad binomial.* En este caso $p = q = 1/2$. Y n = 1, 2, 3, 5, 10, 15 y 20.



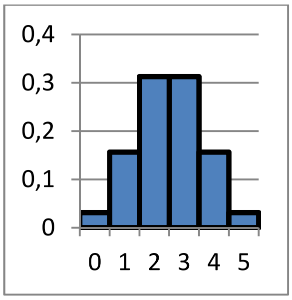
n = 1. B(1, 1/2).



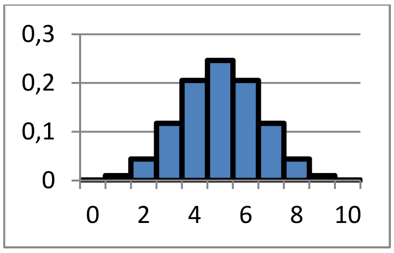
n = 2. B(2, 1/2).



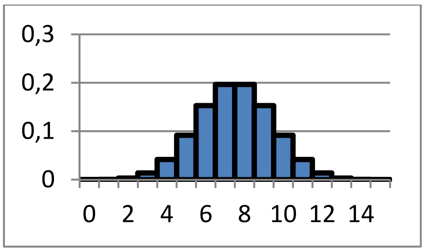
n = 3. B(3, 1/2).



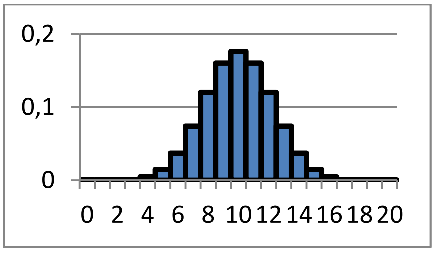
n = 5. B(5, 1/2).



n = 10. B(10, 1/2).



n = 15. B(15, 1/2).



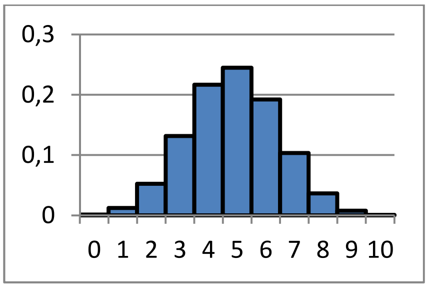
n = 20. B(20, 1/2).

Autora: Raquel Caro
Revisores: Luis Carlos Vidal, Leticia González y Álvaro Valdés
Ilustraciones: Banco de Imágenes de INTEF

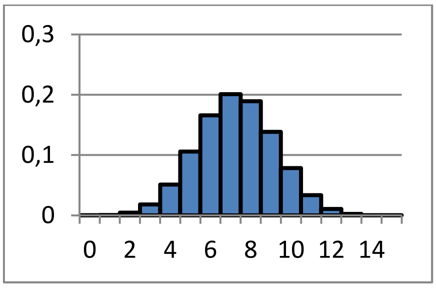
¿Observas alguna diferencia entre los histogramas para *n* par y para *n* impar?

En este caso los histogramas son simétricos pues $p = q = 1/2$.

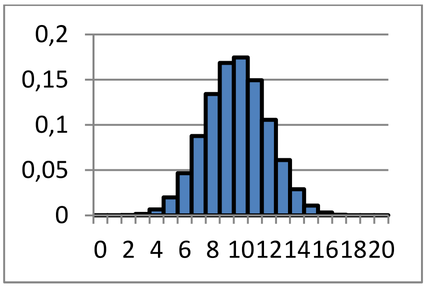
- *Dibujamos los histogramas de la distribución de probabilidad binomial para* $n = 10$, 15 y 20 *del sexo de un recién nacido, donde* $p = 0.485$ *y por tanto* $q = 0.515$.



$n = 10$. $B(10, 0.485)$.



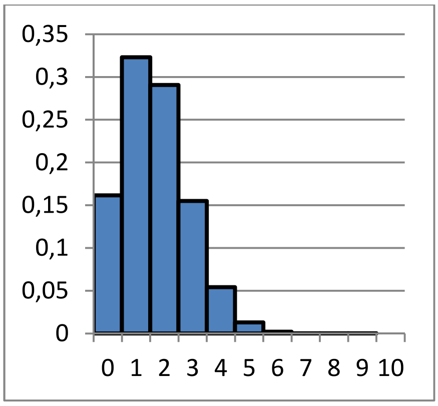
$n = 15$. $B(15, 0.485)$.



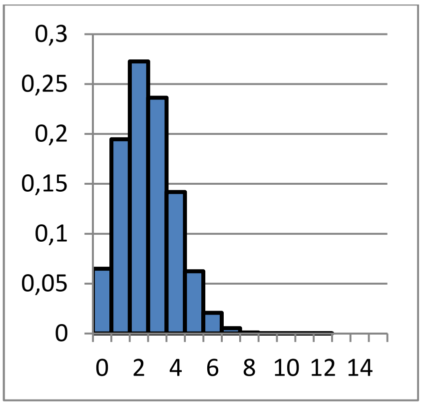
$n = 20$. $B(20, 0.485)$.

Observa como ahora el histograma no es simétrico.

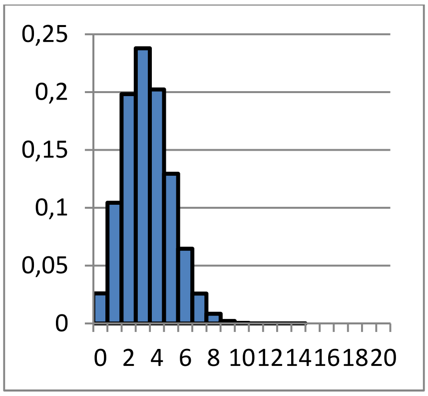
- *Se lanza un dado. Llamamos "éxito" a que salga un 5. Dibujamos los histogramas de la distribución de probabilidad binomial para* $n = 10$, 15 y 20.



$n = 10$. $B(10, 1/6)$.



$n = 15$. $B(15, 1/6)$.



$n = 20$. $B(20, 1/6)$.

La probabilidad viene indicada por el área bajo es histograma.

Actividades propuestas

8. Lanzamos dos monedas y contamos el número de caras. Calcula la media y la desviación típica de dicho experimento.

9. Observa el histograma del experimento de lanzar una moneda 3 veces. Indica las siguientes probabilidades. A) Probabilidad de que el número de caras sea menor que 1. B) Probabilidad de que el número de caras sea menor o igual a 1.

10. Observa el histograma del experimento de lanzar una moneda 5 veces. Indica las siguientes probabilidades. A) Probabilidad de que el número de caras sea menor que 3. B) Probabilidad de que el número de caras sea menor o igual a 3.

11. Escribe la expresión (no lo calcules) de la probabilidad de que al lanzar una moneda 15 veces el número de caras sea menor que 5.

12. Escribe la expresión (no lo calcules) de la probabilidad de que al lanzar un dado 15 veces el número de cincos sea mayor que 10.

Autora: Raquel Caro
Revisores: Luis Carlos Vidal, Leticia González y Álvaro Valdés
Ilustraciones: Banco de Imágenes de INTEF

Parámetros de la distribución binomial

Vamos a dar la expresión de los parámetros de una distribución binomial, su media, varianza y desviación típica. No vamos a demostrar sus expresiones, pero si vamos a calcularlas para algunos casos particulares, que generalizaremos.

Imagina una distribución binomial para $n = 1$, con probabilidad de éxito p: $B(1, p)$. Entonces la distribución de probabilidad es:

X	$p(x)$	$x \cdot p(x)$	$x^2 \cdot p(x)$
0	q	0	0
1	p	p	p
Suma	1	$\mu = p$	$E(x^2) = p$

Luego $\mu = p$ y $\sigma^2 = E(x^2) - E^2(x) = p - p^2 = p(1 - p) = p\,q$.

Hacemos lo mismo para $n = 2$, con probabilidad de éxito p: $B(2, p)$.

X	$p(x)$	$x \cdot p(x)$	$x^2 \cdot p(x)$
0	q^2	0	0
1	$2\,q\,p$	$2\,q\,p$	$2\,q\,p$
2	p^2	$2\,p^2$	$4\,p^2$
Suma	1	μ	$E(x^2)$

Luego $\mu = 2p$ y $\sigma^2 = E(x^2) - E^2(x) = 2\,q\,p + 4\,p^2 - (2p)^2 = 2\,p\,q$.

Y ahora para $n = 3$, con probabilidad de éxito p: $B(3, p)$.

X	$p(x)$	$x \cdot p(x)$	$x^2 \cdot p(x)$
0	q^3	0	0
1	$3\,q^2\,p$	$3\,q^2\,p$	$3\,q^2\,p$
2	$3\,q\,p^2$	$6\,q\,p^2$	$12\,q\,p^2$
3	p^3	$3\,p^3$	$9\,p^3$
Suma	1	μ	$E(x^2)$

Luego $\mu = 3\,q^2\,p + 6\,q\,p^2 + 3\,p^3 = 3\,p(q^2 + 2\,q\,p + p^2) = 3\,p(q + p)^2 = 3\,p$ y

$\sigma^2 = E(x^2) - E^2(x) = 3\,q^2\,p + 12\,q\,p^2 + 9\,p^3 - (3p)^2 =.3\,p(q^2 + 2\,q\,p + p^2) + 6\,q\,p^2 + 6\,p^3 - 9p^2 =.$

$3\,p(q + p)^2 + 6\,p^2\,(q + p) - 9p^2 = 3\,p - 3p^2 = 3\,p\,(1 - p) = 3\,p\,q.$

¡Piensa! Queremos generalizar estos resultados para cualquier valor de n. ¿Cuánto crees que valdrá la media y la varianza?

En efecto:

En una distribución binomial $B(n, p)$ la media vale siempre:

$$E(x) = \mu = n{\cdot}p,$$

la varianza $\sigma^2 = n{\cdot}p{\cdot}q$ y

la desviación típica $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$.

Actividades propuestas

13. En el control de calidad de bombillas de bajo consume de una fábrica se ha comprobado que el 90 % son buenas. Se toma una muestra de 500 bombillas. Por término medio, ¿cuántas serán de buena calidad? Calcula la media, varianza y desviación típica.

14. En el estudio sobre una nueva medicina para la hepatitis C se ha comprobado que produce curaciones completas en el 80 % de los casos tratados. Se administra a mil nuevos enfermos, ¿cuántas curaciones esperaremos que se produzcan?