

## 5 DISTRIBUCIONES BINOMIAL Y DE POISSON

La repetición sucesiva de  $n$  pruebas (ensayos) de BERNOULLI, de modo independiente y manteniendo constante la probabilidad de éxito,  $p$ , da lugar a la variable aleatoria discreta  $X$  cuyo espacio imagen está formado por el número de éxitos posibles:  $\Omega^*=\{0,1,2,3, \dots,n\}$ . Se dice que  $X$  sigue una **LEY BINOMIAL**  $\mathcal{B}(n;p)$ , cuya función o ley de probabilidad viene dada por

$$b(k;n;p) = b_{n;p}(k) = P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad \text{para cada } k \in \{0,1,2,3,\dots,n\}.$$

Su función de distribución acumulativa es

$$x \in \mathbb{R} \mapsto B(x;n;p) = B_{n;p}(x) = P(X \leq x) = \sum_{k \leq x} P(X = k) = \sum_{k \leq x} \binom{n}{k} p^k q^{n-k}.$$

La hoja de cálculo EXCEL, de Microsoft Office, en su librería de funciones estadísticas, tiene las dos funciones en una cuya sintaxis es

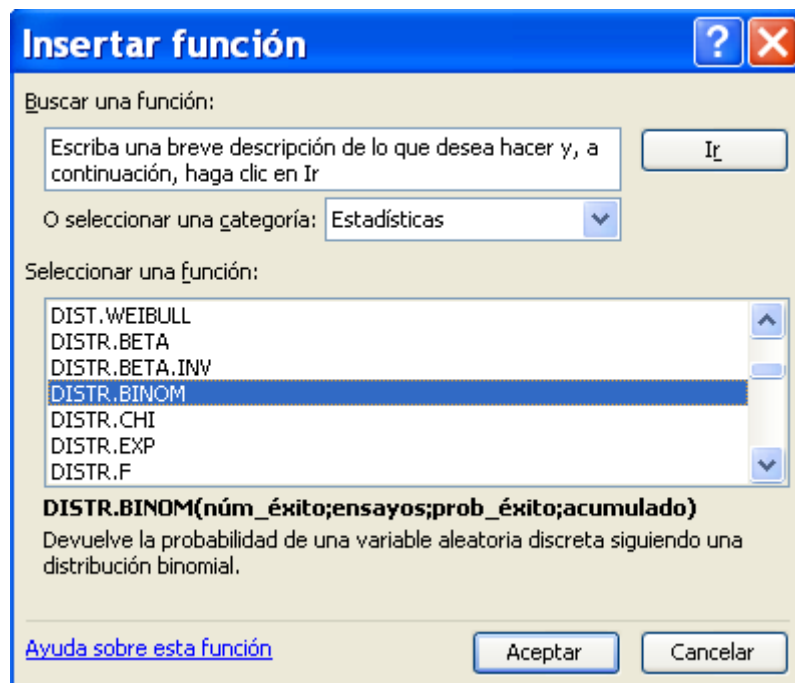
**DISTR.BINOM(núm\_éxitos;ensayos;prob\_éxito;acumulado),**

donde acumulado es un valor lógico, de modo que

**DISTR.BINOM(k;n;p;FALSO) =  $b(k;n;p) = P(X = k)$ ,**

mientras que

**DISTR.BINOM(x;n;p;VERDADERO) =  $B(x;n;p) = P(X \leq x)$**



Una variable aleatoria  $X$  que toma valores naturales sigue una **LEY DE POISSON** si su función o ley de probabilidad viene dada por

$$p(k; \lambda) = p_{\lambda}(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{para cada } k \in \{0, 1, 2, 3, \dots\},$$

Su función de distribución acumulativa es

$$x \in \mathbb{R} \mapsto P(x; \lambda) = P_{\lambda}(x) = P(X \leq x) = \sum_{k \leq x} P(X = k) = \sum_{k \leq x} \frac{\lambda^k}{k!} e^{-\lambda}.$$

La hoja de cálculo EXCEL, también cuenta en su librería de funciones estadísticas con dos funciones en una cuya sintaxis es

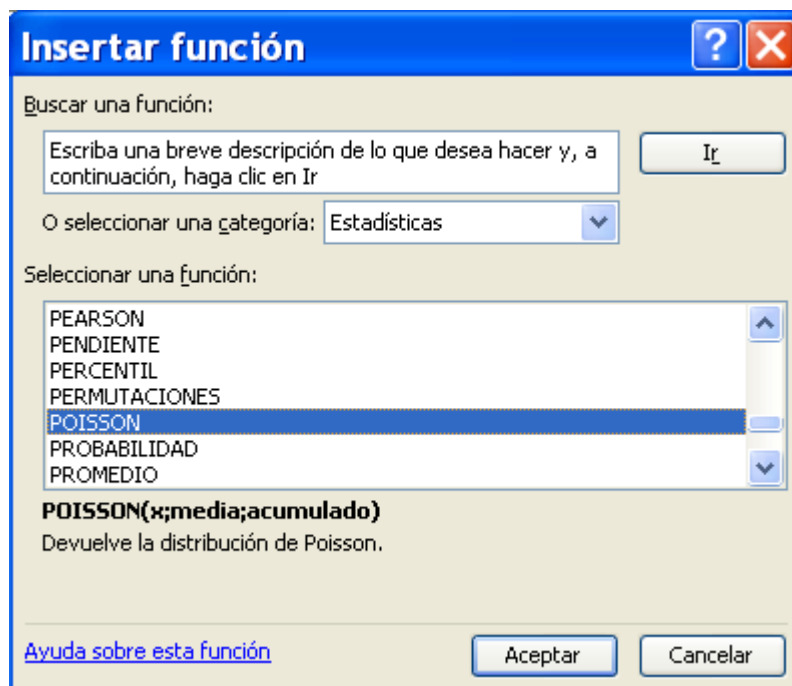
### **POISSON(x;media;acumulado)**

donde acumulado es un valor lógico tal que

$$\text{POISSON}(k; \lambda; \text{FALSO}) = p(k; \lambda) = P(X = k),$$

mientras que

$$\text{POISSON}(x; \lambda; \text{VERDADERO}) = P(x; \lambda) = P(X \leq x)$$



Puede ser útil leer la ayuda que se proporciona para cada función.

En los ejercicios que siguen utilizaremos estas funciones siempre que sea posible.

- 1 En una población, el 25% de las personas adultas poseen automóvil. Se eligen doce personas de dicha población (con reemplazamiento e independientemente e unas de otras). Calcular las probabilidades de los sucesos siguientes:
- A lo sumo cuatro de ellas poseen coche.
  - Seis de ellas no tienen coche.
  - Ocho al menos no tienen auto, sabiendo que ocho a lo sumo sí tienen.

La variable aleatoria a tratar,  $X =$  "número que tienen coche entre las doce personas elegidas", sigue una binomial  $\mathcal{B}(n=12; p=0,25)$ .

$$b(k;12;0,25) = P(X = k) = \binom{12}{k} (0,25)^k \cdot (0,75)^{12-k}, \quad \text{para cada } k \in \{0,1,2,3,\dots,12\}$$

Hemos de calcular

$$P(X \leq 4) = F(4), \quad P(12-X=6) = P(X=6) \quad \text{y} \quad P(12-X \geq 8 \mid X \leq 8) = P(X \leq 4 \mid X \leq 8),$$

siendo

$$P(X \leq 4 \mid X \leq 8) = \frac{P("X \leq 4" \cap "X \leq 8")}{P(X \leq 8)} = \frac{P(X \leq 4)}{P(X \leq 8)} = \frac{B(4;12;0,25)}{B(8;12;0,25)}$$

	A	B	C	D	E
1	<b>BINOMIAL</b>				
2	n = 12				
3	p = 0,25				
4					
5	k = 4	P(X<=4)=B(4;12;0,25)=	DISTR.BINOM(4;12;0,25;VERDADERO)=	0,842686373	
6					
7	k = 6	P(X=6)=b(6;12;0,25)=	DISTR.BINOM(6;12;0,25;FALSO)=	0,04014945	
8					
9	k = 8	P(X<=8)=B(8;12;0,25)=	DISTR.BINOM(8;12;0,25;VERDADERO)=	0,999608338	
10					
11					
12		P(X<=4   X<=8)=	B(4;12;0,25) / B(8;12;0,25)=	0,84301655	

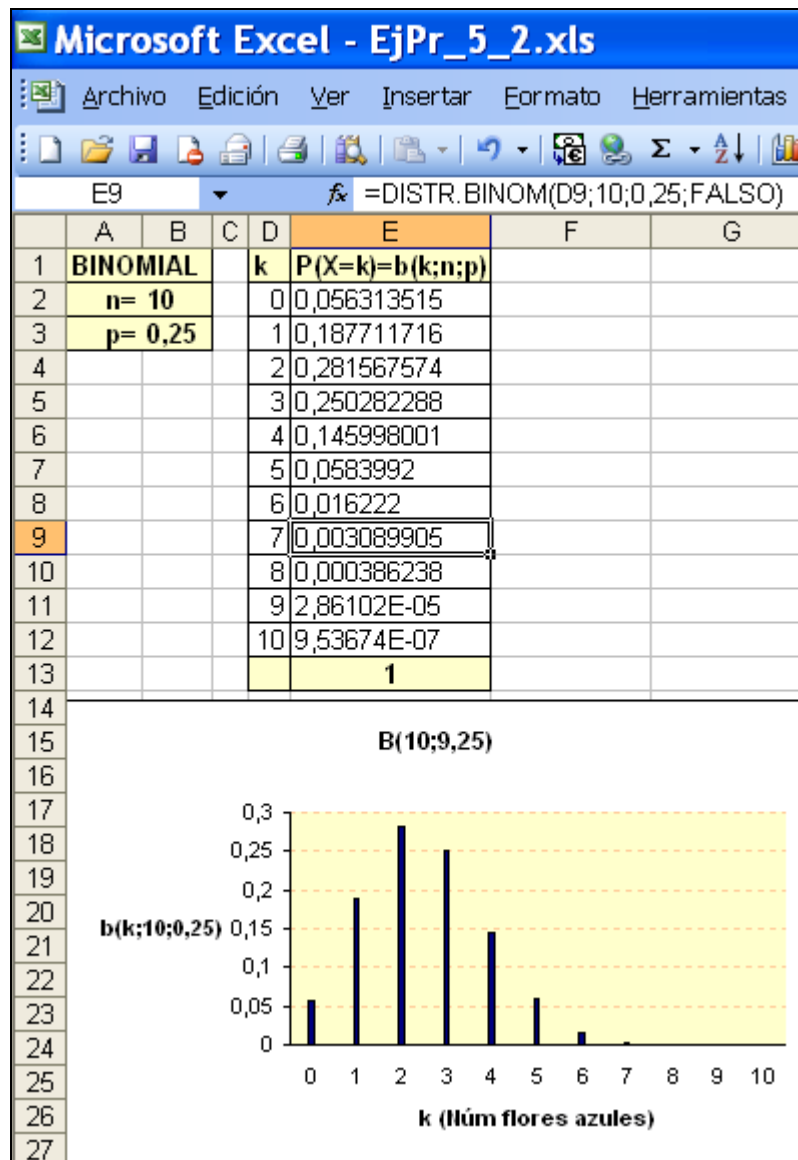
2 En una cierta familia de flores, las flores rojas son tres veces más numerosas que las azules. Se eligen al azar diez flores de esta familia y se supone que las elecciones son independientes las unas de las otras.

- a) ¿Cuál es la probabilidad de no obtener ninguna azul?
- b) ¿Cuáles son las probabilidades de obtener  $k$  flores azules con  $k \in \{1, 2, \dots, 10\}$ ?
- c) Construir el diagrama de barras de esta distribución.

Se tiene que  $P(\text{Roja}) = 3 \cdot P(\text{Azul})$ . Si  $P(\text{Azul}) = p$ ,  $P(\text{Roja}) = 3p$ , entonces  $p + 3p = 1$  y  $p = P(\text{Azul}) = 1/4$ ; luego  $P(\text{Roja}) = 3/4$

La variable aleatoria discreta  $X =$  "número de flores azules entre las 10 elegidas al azar" sigue una distribución binomial  $\mathcal{B}(n=10; p=0,25)$ :

$$b(k; 10; 0,25) = P(X = k) = \binom{10}{k} (0,25)^k \cdot (0,75)^{10-k}, \quad \text{para cada } k \in \{0, 1, 2, 3, \dots, 10\}$$



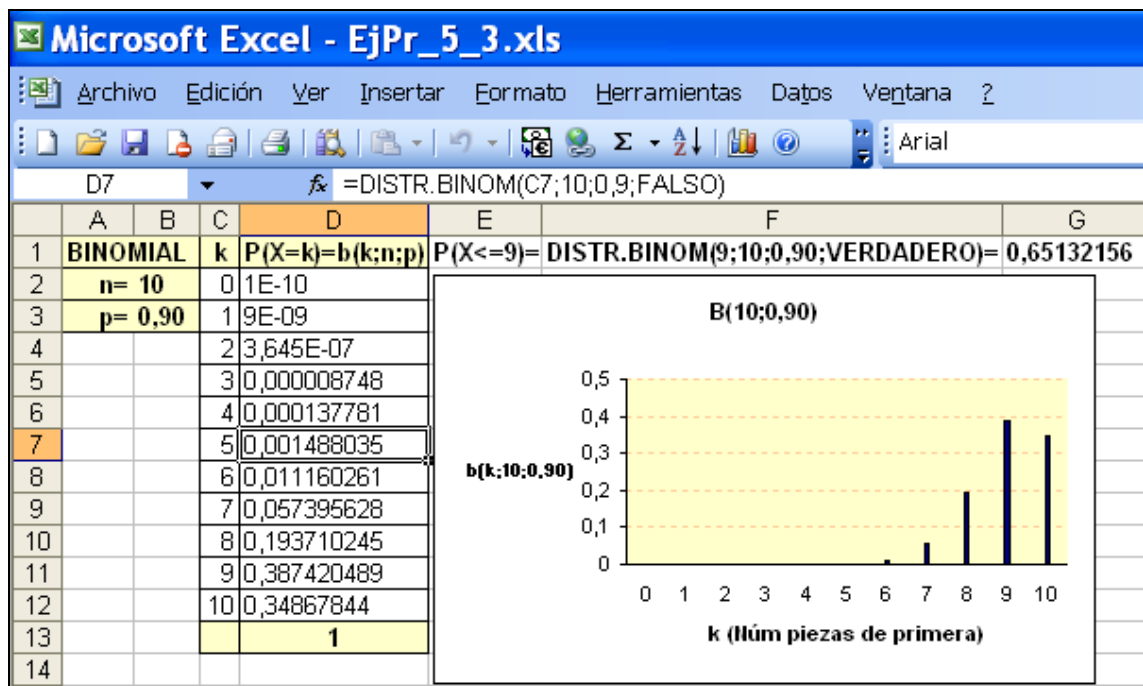
3 Una fábrica produce piezas para automóviles. Sobre un gran número de piezas fabricadas, se constata que se pueden considerar de primera calidad el 90%.

Una pieza tomada al azar tiene, por tanto, la probabilidad 0,90 de ser de primera calidad.

1. Se toman 10 piezas al azar.
  - a) El número de piezas de primera calidad contenidas en este conjunto es una variable aleatoria,  $X$ . ¿Cuál es su ley de probabilidad?
  - b) Si  $k$  es un número natural ( $\leq 10$ ), calcular  $p_k = P(X = k)$
  - c) Calcular la probabilidad de encontrar a lo sumo 9 piezas de primera calidad.
2. Por otra parte, se ha observado que en el conjunto de la producción hay que rechazar un 2% de piezas por defectuosas.
  - a) Expresar la probabilidad de encontrar exactamente dos piezas defectuosas en un conjunto de 100 elegidas al azar y dar un valor aproximado de la misma.
  - b) Calcular la probabilidad de encontrar en el conjunto de 100 al menos dos piezas a rechazar.

1. La variable aleatoria discreta  $X =$  “número de piezas de primera calidad entre las 10 elegidas al azar” sigue una distribución binomial  $\mathcal{B}(n=10; p=0,90)$ :

$$b(k; 10; 0,90) = P(X = k) = \binom{10}{k} (0,90)^k \cdot (0,10)^{10-k}, \quad \text{para cada } k \in \{0, 1, 2, 3, \dots, 10\}$$



2. La variable aleatoria discreta  $Y =$  “número de piezas defectuosas entre las 100 elegidas al azar” sigue una distribución binomial  $\mathcal{B}(n=100; p=0,02)$ :

$$b(k;100;0,02) = P(Y = k) = \binom{100}{k} (0,02)^k \cdot (0,98)^{100-k}, \quad \text{para cada } k \in \{0,1,2,3,\dots,100\}$$

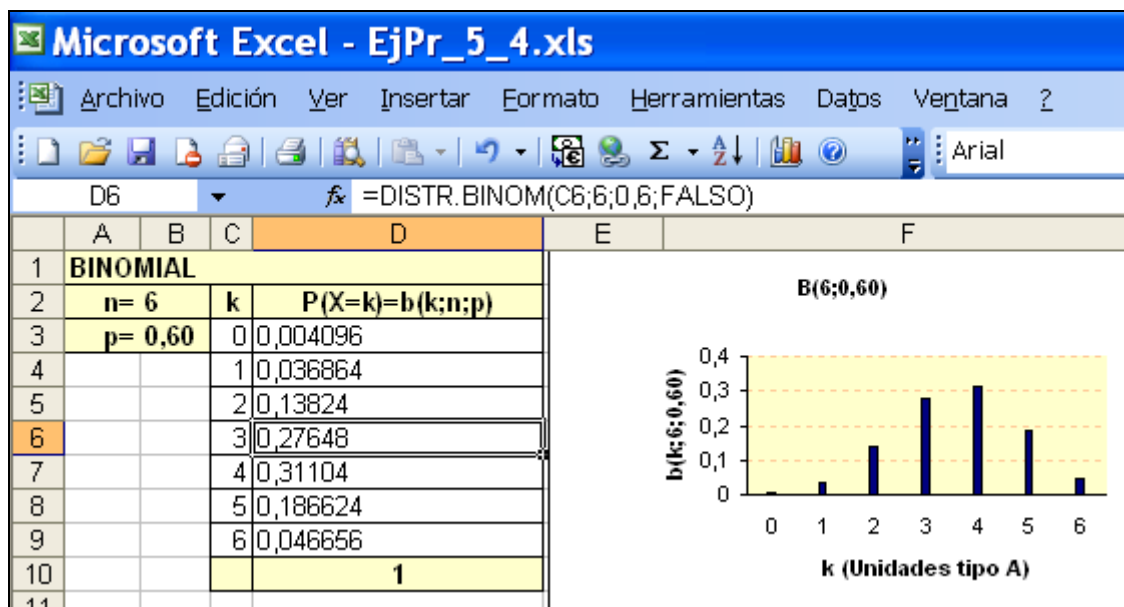
Como  $n = 100$  es grande y  $p = 0,02$  pequeño, la variable  $Y$  sigue aproximadamente una distribución de POISSON con  $\lambda = n \cdot p = 2$ .

	A	B	C	D
1	<b>DISTRIBUCIÓN BINOMIAL</b>			
2	n= 100			
3	p= 0,02			
4				
5	P(Y=2)=	b(2;100;0,02)=	0,273413912	
6				
7	P(Y>=2)=	1-P(Y<2)=	1-P(Y<=1)=	0,596728289
8				
9	<b>APROX. POR POISSON</b>			
10	λ = 2			
11				
12	P(Y=2)≈	p(2;2)=	0,270670566	
13				
14	P(Y>=2)=	1-P(Y<2)=	1-P(Y<=1)=	0,59399415

- 4 Una población de 20 unidades consta de 12 de tipo A y 8 de tipo B.
- a) Se extraen, **con reemplazamiento**, muestras de 6 unidades. Sea X la variable aleatoria que hace corresponder a cada muestra el número de unidades de tipo A. Calcular las probabilidades  $p_k = P(X = k)$ , con  $k \in \{0,1,2,3,4,5,6\}$ .
- b) Se extraen, **sin reemplazamiento**, muestras de 6 unidades. Sea Y la variable aleatoria que hace corresponder a cada muestra el número de unidades de tipo A. Calcular las probabilidades  $p_k = P(Y = k)$ , con  $k \in \{0,1,2,3,4,5,6\}$ .

a) La variable aleatoria discreta X = “número de tipo A (éxito) entre las 6 elegidas al azar”, al realizar la extracciones **con reemplazamiento**, sigue una distribución binomial  $\mathcal{B}(n=6; p=12/20=0,60)$ :

$$b(k;6;0,60) = P(X = k) = \binom{6}{k} (0,60)^k \cdot (0,40)^{6-k}, \quad \text{para cada } k \in \{0,1,2,3,\dots,6\}$$



b) Cuando de una población con  $N=20$  individuos, de los cuales  $N_1=12$  son del tipo A y  $N_2=N-N_1=8$  son del tipo B, se extraen muestras de tamaño  $n=6$  **sin reemplazamiento** (lo que equivale a escoger un subconjunto de  $n$  elementos de la población), la variable aleatoria Y = “núm. de elementos de tipo A en la muestra” sigue una **distribución hipergeométrica**  $\mathcal{H}(N;N_1;n)$  de función de probabilidad (ver pág. 287 del texto)

$$h_{N;N_1;n}(k) = P(Y = k) = \frac{\binom{N_1}{k} \cdot \binom{N - N_1}{n - k}}{\binom{N}{n}}, \quad \text{para cada } k \in \{0,1,2,3,\dots,n\},$$

donde ha de ser  $\max(0, n+N_1-N) \leq k \leq \min(n, N_1)$ .

La hoja de cálculo EXCEL, cuenta en su librería de funciones estadísticas con la ley de probabilidad de esta distribución. Su sintaxis es

**DISTR.HIPERGEOM(muestra\_éxito;núm\_de\_muestra;población\_éxito;núm\_de\_población)**

donde

**muestra\_éxito** es el número **k** de éxitos (tipo A) en la muestra;

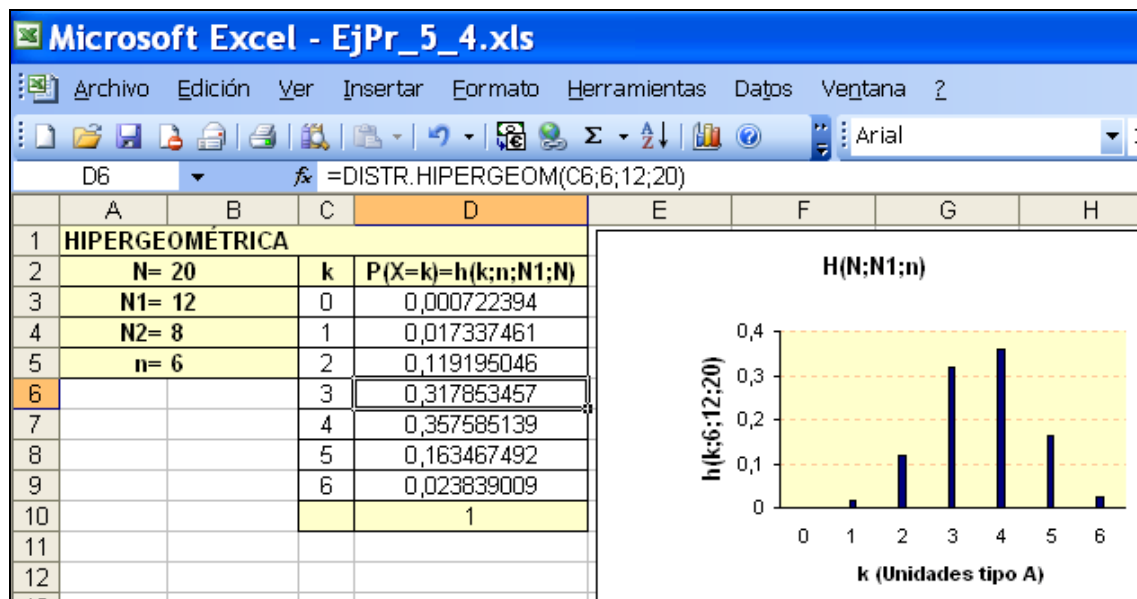
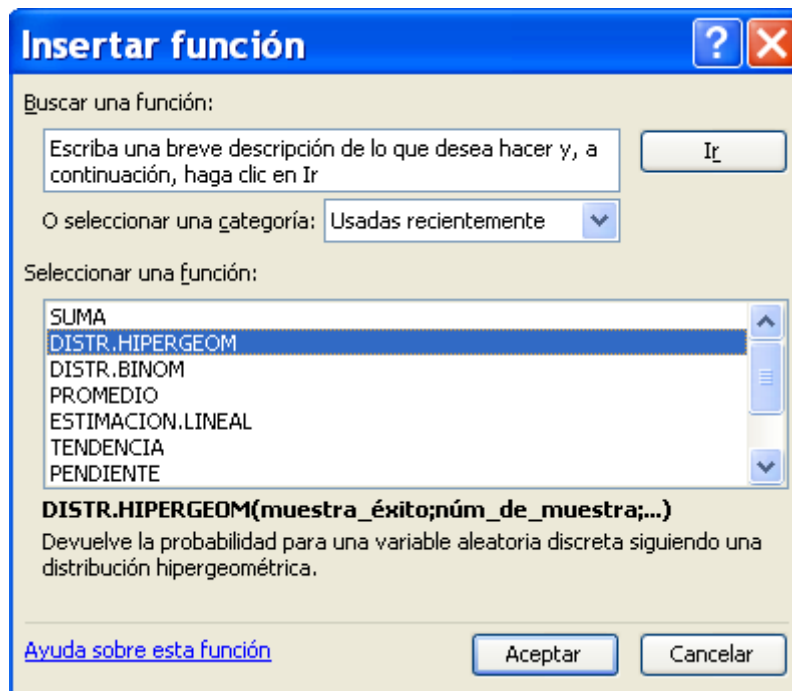
**núm\_de\_muestra** es el tamaño **n** de la muestra;

**población\_éxito** es el número de éxitos **N<sub>1</sub>** en la población; y

**núm\_de\_población** es el tamaño **N** de la población,

de modo que :

$$\text{DISTR.HIPERGEOM}(k;n;N_1;N) = P(X = k) = h_{N;N_1;n}(k)$$



5 Un dado cúbico con cuatro caras blancas y dos rojas se arroja tres veces consecutivas. Sea  $X$  la variable aleatoria que a cada resultado (terna ordenada) le asigna el número de caras blancas obtenidas.

Describir el espacio muestral  $\Omega$  y el espacio imagen de  $X$ ,  $X(\Omega) = \Omega^*$ . Representar la distribución de probabilidad de  $X$  (diagrama de barras) y la función de distribución acumulativa,  $F(x)$ . Calcular  $E(X)$  y  $Var(X)$ .

Si  $D = \{b, r\}$  el espacio muestral de experimento de BERNOUILLI consistente en lanzar el dado una vez, con  $p = P(b) = 4/6 = 2/3$  y  $q = 1 - p = 1/3$ , el espacio muestral al realizar tres lanzamientos sucesivos es  $\Omega = D^3 = D \times D \times D$ .

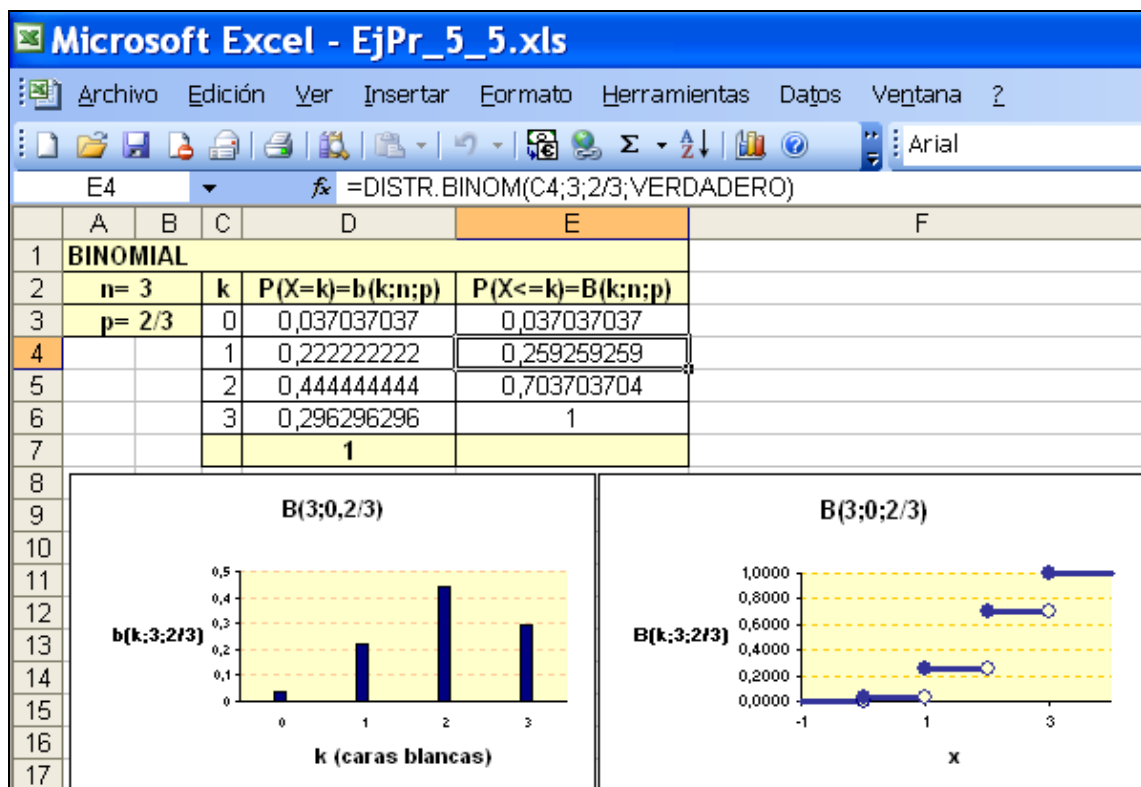
La variable aleatoria  $X: \Omega \rightarrow \mathbb{R}$  que a cada elemento de  $\Omega$  le hace corresponder el número de caras "b" tiene por imagen:  $\Omega^* = X(\Omega) = \{0, 1, 2, 3\}$  y sigue una distribución binomial  $\mathcal{B}(n=3; p=2/3)$ .

$$\Omega = D^3 = \{(r,r,r); (r,r,b),(r,b,r),(b,r,r); (b,b,r),(b,r,b),(r,b,b); (b,b,b)\}$$

$$X$$

$$\mathbb{R} \supseteq \Omega^* = \{ 0, \quad 1, \quad 2, \quad 3 \}$$

Por tanto:  $E[X] = np = 3 \cdot \frac{2}{3} = 2$       y       $VAR[X] = npq = 3 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{3}$



6 El promedio de llegadas de automóviles a una gasolinera entre las 11 AM y las 12 AM es de 2 por minuto.

1. Calcular la probabilidad de que entre las 11h 30m y las 11h 31m lleguen a la gasolinera: a) más de 3 autos; b) algún auto; y c) un número de autos comprendido entre 2 y 5, ambos inclusive.
2. Calcular la probabilidad de que entre las 11h 30m y las 11h 35m lleguen a la gasolinera: a) 2 autos; b) menos de 2 autos; y c) no menos de 2 autos.

1. La variable  $X =$  “número de llamadas por minuto” sigue una distribución de POISSON de parámetro (media)  $\lambda = 2$  con función de probabilidad

$$p(k;2) = p_2(k) = P(X = k) = \frac{2^k}{k!} e^{-2}, \quad \text{para cada } k \in \{0,1,2,3,\dots\}$$

y hay que calcular **a)**  $P(X > 3) = 1 - P(X \leq 3) = 1 - P_{\lambda=2}(3)$ , **b)**  $P(X \geq 1) = 1 - P(X = 0)$  y **c)**  $P(2 \leq X \leq 5) = P(X \leq 5) - P(X \leq 1) = P_{\lambda=2}(5) - P_{\lambda=2}(1)$

The screenshot shows an Excel spreadsheet titled "Microsoft Excel - EjPr\_5\_6.xls". The active cell is B7, containing the formula `=POISSON(5;2;VERDADERO)-POISSON(1;2;VERDADERO)`. The spreadsheet has two columns, A and B. Row 1 is a header for "1. POISSON (lambda=2)". Row 2 contains the calculation for part a: `a) P(X > 3) = 1 - P(X <= 3) = 1 - POISSON(3;2;VERDADERO) =` with the result **0,14287654** in column B. Row 3 is empty. Row 4 contains the calculation for part b: `b) P(X >= 1) = 1 - P(X = 0) = 1 - POISSON(0;2;FALSO) =` with the result **0,86466472** in column B. Row 5 is empty. Row 6 contains the calculation for part c: `c) P(2 <= X <= 5) = P(X <= 5) - P(X <= 1) =`. Row 7 contains the final result for part c: `=POISSON(5;2;VERDADERO) - POISSON(1;2;VERDADERO) =` with the result **0,57743054** in column B.

	A	B
1	<b>1. POISSON (lambda=2)</b>	
2	a) $P(X > 3) = 1 - P(X \leq 3) = 1 - \text{POISSON}(3;2;\text{VERDADERO}) =$	<b>0,14287654</b>
3		
4	b) $P(X \geq 1) = 1 - P(X = 0) = 1 - \text{POISSON}(0;2;\text{FALSO}) =$	<b>0,86466472</b>
5		
6	c) $P(2 \leq X \leq 5) = P(X \leq 5) - P(X \leq 1) =$	
7	$= \text{POISSON}(5;2;\text{VERDADERO}) - \text{POISSON}(1;2;\text{VERDADERO}) =$	<b>0,57743054</b>

2. La variable  $Y =$  “número de llamadas cada 5 minutos” sigue una distribución de POISSON de parámetro (media)  $\Lambda = \lambda t = 2 \cdot 5 = 10$  con función de probabilidad

$$p(k;10) = p_{10}(k) = P(X = k) = \frac{10^k}{k!} e^{-10}, \quad \text{para cada } k \in \{0,1,2,3,\dots\}$$

The screenshot shows an Excel spreadsheet titled "Microsoft Excel - EjPr\_5\_6.xls". The active cell is B4, containing the formula `=POISSON(2;10;VERDADERO)`. The spreadsheet has two columns, A and B. Row 1 is a header for "2. POISSON (LAMBDA=5\*2=10)". Row 2 contains the calculation for part a: `a) P(X=2) = POISSON(2;10;FALSO) =` with the result **0,002269996** in column B. Row 3 is empty. Row 4 contains the calculation for part b: `b) P(X <= 2) = P(X <= 1) = POISSON(1;10;VERDADERO) =` with the result **0,002769396** in column B. Row 5 is empty. Row 6 contains the calculation for part c: `c) P(X >= 2) = 1 - P(X < 2) = 1 - P(X <= 1) = 1 - POISSON(1;10;VERDADERO) =` with the result **0,999500601** in column B.

	A	B
1	<b>2. POISSON (LAMBDA=5*2=10)</b>	
2	a) $P(X=2) = \text{POISSON}(2;10;\text{FALSO}) =$	<b>0,002269996</b>
3		
4	b) $P(X \leq 2) = P(X \leq 1) = \text{POISSON}(1;10;\text{VERDADERO}) =$	<b>0,002769396</b>
5		
6	c) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - \text{POISSON}(1;10;\text{VERDADERO}) =$	<b>0,999500601</b>

7 Una centralita telefónica atiende 100 abonados. La probabilidad de que durante un minuto llame un abonado es 0,02. ¿Cuál de los dos sucesos es más probable: que en un minuto llamen 3 abonados o que llamen 4?

La variable  $X = \text{“número de llamadas por minuto de un abonado”}$  sigue una distribución de POISSON de parámetro (media)  $\lambda = 0,02$  con función de probabilidad

$$p(k;0,02) = p_{0,02}(k) = P(X = k) = \frac{(0,02)^k}{k!} e^{-0,02}, \quad \text{para cada } k \in \{0,1,2,3,\dots\}$$

La variable  $Y = \text{“número de llamadas por minuto de 100 abonados”}$  sigue una distribución de POISSON de parámetro (media)  $\Lambda = \lambda t = 0,02 \cdot 100 = 2$  con función de probabilidad

$$p(k;2) = p_2(k) = P(X = k) = \frac{2^k}{k!} e^{-2}, \quad \text{para cada } k \in \{0,1,2,3,\dots\}$$

	A	B
1	<b>POISSON (LAMBDA=100*0,02=2)</b>	
2	P(X=3) = POISSON(3;2,FALSO)=	
3		<b>0,180447044</b>
4	P(X=4) = POISSON(4;2,FALSO)=	
5		<b>0,090223522</b>

- 8 La probabilidad de que un individuo al que se le inyecta una vacuna sufra una reacción adversa es 0,04. Se someten 100 individuos a dicha vacuna ( $\lambda = np$ ). Calcular la probabilidad de que:
- Tengan reacción 4 individuos.
  - Ninguno tenga reacción.
  - Al menos 3 individuos tienen reacción.
  - Tengan reacción 4, si se sabe que al menos 3 la han tenido.

Microsoft Excel - EjPr\_5\_8.xls

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana

B4      fx =POISSON(0;4;FALSO)

	A	B
1	<b>POISSON (LAMBDA=n·p=100·0,04=4)</b>	
2	a) $P(X=4) = \text{POISSON}(4;4;\text{FALSO}) =$	<b>0,195366815</b>
3		
4	b) $P(X=0) = \text{POISSON}(0;4;\text{FALSO}) =$	<b>0,018315639</b>
5		
6	c) $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2) = 1 - \text{POISSON}(2;4;\text{VERDADERO}) =$	<b>0,761896694</b>
7		
8	d) $P(X=4   X \geq 3) = [P("X=4" \text{ y } "X \geq 3")]/P(X \geq 3) = P(X=4)/P(X \geq 3) =$	<b>0,256421660</b>

- 9 El número de accidentados que acude al servicio de urgencias de un hospital cada 10 minutos es, en promedio, 1,6. Calcular la probabilidad de que entre las 18h y las 18h 20m acudan:
- 0 accidentados.
  - 1 accidentado.
  - 2 accidentados.
  - Al menos 2 accidentados.
  - Más de 2 accidentados.

Microsoft Excel - EjPr\_5\_9.xls

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana

B10      fx =B8-B6

	A	B
1	<b>POISSON (LAMBDA=2·1,6=3,2)</b>	
2	a) $P(X=0) = \text{POISSON}(0;3,2;\text{FALSO}) =$	<b>0,040762204</b>
3		
4	b) $P(X=1) = \text{POISSON}(1;3,2;\text{FALSO}) =$	<b>0,130439053</b>
5		
6	c) $P(X=2) = \text{POISSON}(2;3,2;\text{FALSO}) =$	<b>0,208702484</b>
7		
8	d) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - \text{POISSON}(1;3,2;\text{VERDADERO}) =$	<b>0,828798743</b>
9		
10	e) $P(X > 2) = P(X \geq 2) - P(X=2) =$	<b>0,620096259</b>

- 10 La probabilidad de que se agote el stock de cierto producto en un almacén en una semana es del 3%. Calcular la probabilidad de que en un período de 50 semanas
- No haya ningún fallo de stock.
  - Haya 1 fallo.
  - Haya 2 fallos.
  - Haya al menos 2 fallos.
  - Haya más de 2 fallos.

Microsoft Excel - EjPr\_5\_10.xls

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana

B6  $=\text{POISSON}(2;1,5;\text{FALSO})$

	A	B
1	<b>POISSON (LAMBDA=50*0,03=1,5)</b>	
2	a) $P(X=0) = \text{POISSON}(0;1,5;\text{FALSO}) =$	<b>0,223130160</b>
3		
4	b) $P(X=1) = \text{POISSON}(1;1,5;\text{FALSO}) =$	<b>0,334695240</b>
5		
6	c) $P(X=2) = \text{POISSON}(2;1,5;\text{FALSO}) =$	<b>0,251021430</b>
7		
8	d) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - \text{POISSON}(1;1,5;\text{VERDADERO}) =$	<b>0,442174600</b>
9		
10	e) $P(X > 2) = P(X \geq 2) - P(X = 2) =$	<b>0,191153169</b>