Variables Aleatorias Discretas

Variables Aleatorias

- El resultado de un experimento aleatorio puede ser descrito en ocasiones como una cantidad numérica.

- En estos casos aparece la noción de variable aleatoria
  - Función que asigna a cada suceso un número.

- Las variables aleatorias pueden ser *discretas* o *continuas*.

Variables Aleatorias cont.

Una ***variable aleatoria*** *X* es una ***función*** que asocia a cada suceso del espacio muestral *Ω* de un experimento aleatorio un valor numérico real:

$$X : \Omega \to \Re$$

$$w \to X(w)$$

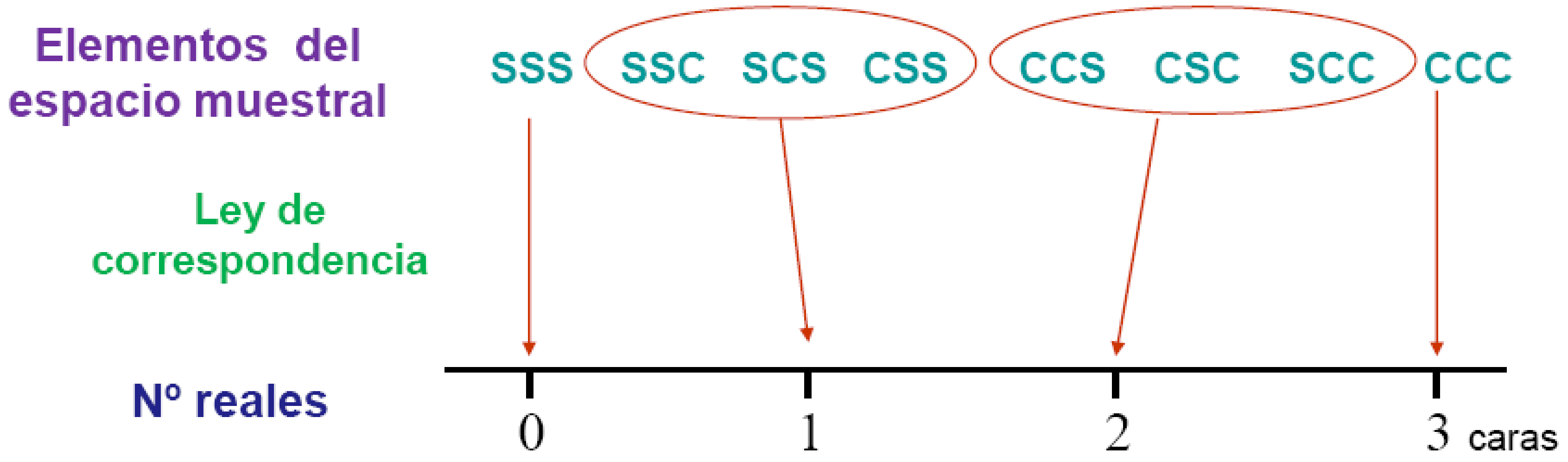
Llamar variable a una función resulta algo confuso, por ello hay que insistir en que *es una función*.

Veremos en esta clase el caso discreto.

Variables Aleatorias cont.

Ejemplo de variable aleatoria discreta:

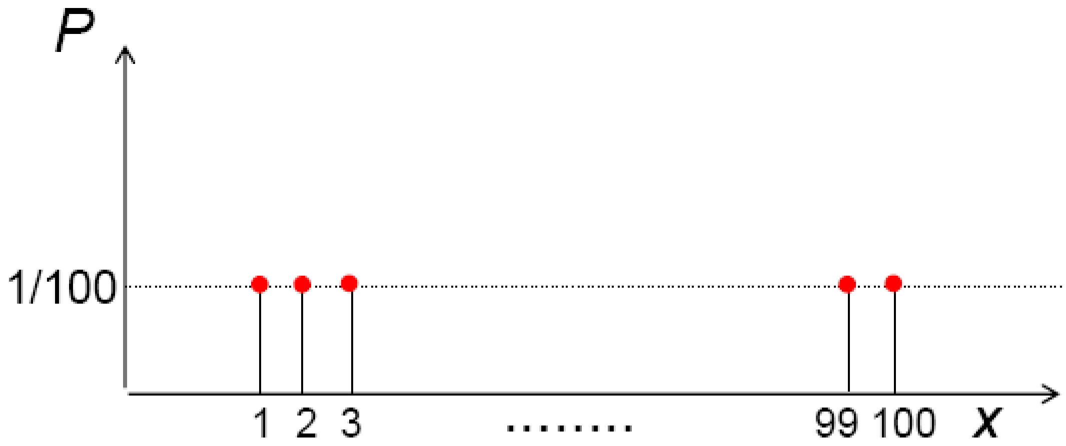
Número de caras al lanzar 3 monedas.



Variables Aleatorias cont.

Pensemos en un número entero del 1 al 100.
“¿Qué numero será?”

Intentaremos representar el estado de incertidumbre mediante una función matemática: la **función de probabilidad**



Función de Probabilidad

Una vez definida una variable aleatoria *X*, podemos definir una función de probabilidad asociada a *X*, de la siguiente forma:

$$p : \Re \rightarrow [0,1]$$

$$x \rightarrow p(x) = P(X = x)$$

La función de probabilidad debe cumplir:

$$(i)\ p(x) \geq 0 \quad \forall x \in \Re$$

$$(ii) \sum_x p(x) = 1$$

(Suma sobre todos los posibles valores que puede tomar la variable aleatoria).

Función de Probabilidad Discreta

Lanzamiento de dos monedas. Sea *X*: número de caras.

Valores	Probabilidad
2	1/4 = 0.25
1	2/4 = 0.50
0	1/4 = 0.25

Ejemplo

Sea el experimento "lanzar dos dados". Definamos el espacio muestral Ω como:
$\Omega = \{(1,1),(1,2),\ldots(1,6),\ldots,(5,6),(6,6)\}$

Definamos la variable aleatoria discreta X como:

$$X : \Omega \rightarrow S \in \Re$$

con $S = \{2,3,\ldots,12\}$ *la suma de puntos.*
Una posible función de probabilidad es:

$$f : \Re \rightarrow [0,1]$$

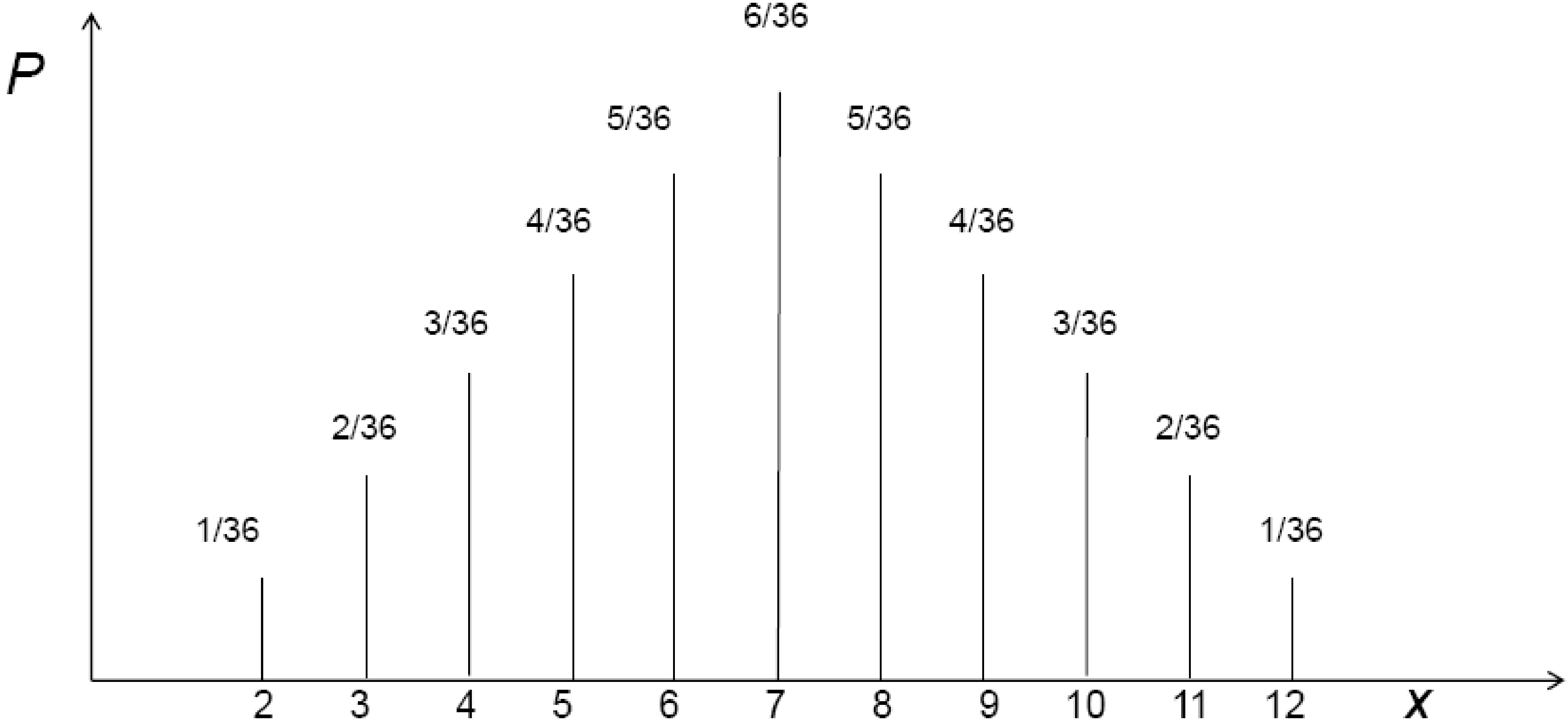
$$f(2) = P(X = 2) = P((1,1)) = 1/36$$

$$f(3) = P(X = 3) = P((1,2) \cup (2,1)) = 2/36$$

$$f(4) = P(X = 4) = P((1,3) \cup (3,1) \cup (2,2)) = 3/36$$

$\ldots$

Función de Probabilidad de una variable X



Observar que cumple las dos condiciones: es siempre positiva y está normalizada (suma 1).

Función de Distribución

Dada una variable aleatoria discreta *X* se llama **función de distribución** a la función *F* definida como:

$$F : \Re \rightarrow [0,1]$$

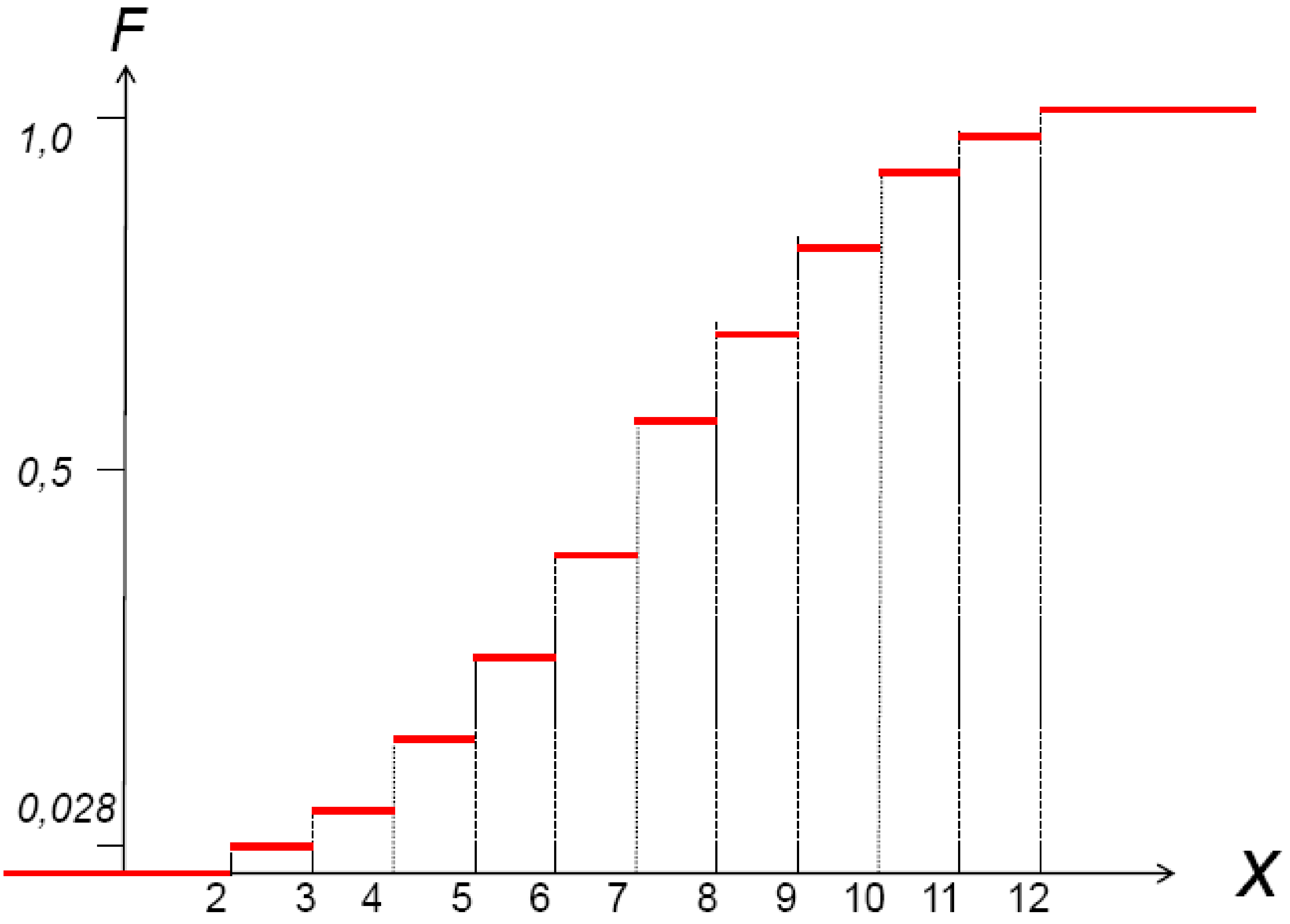
$$x \rightarrow F(x) = P(X \leq x)$$

En nuestro ejemplo de los dos dados:

F(5) = P(X ≤ 5) = P(x = 2 o x = 3 o x = 4 o x = 5)

F(5) = 1/36 + 2/36 +3/36 + 4/36 = 10/36

Función de Distribución de una variable X

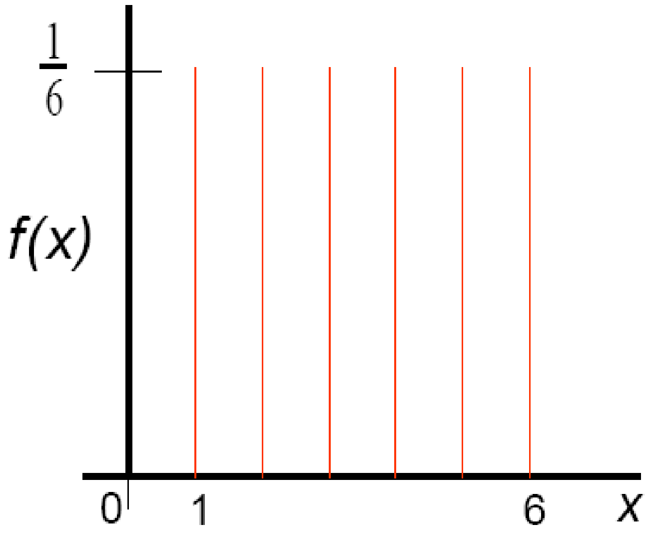


Ejemplo

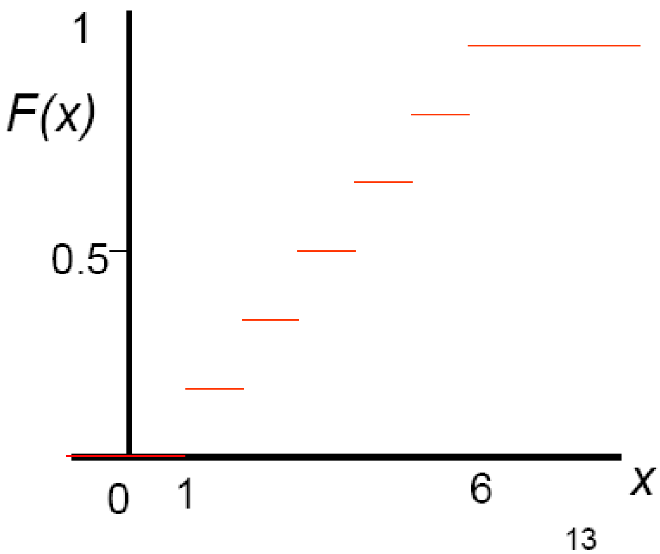
Dibujar la función de probabilidad f(x) y la función de distribución F(x) de una variable discreta definida como:

X = Número en la cara de un dado.

X tiene como posibles valores *x* = 1, 2, 3, 4, 5, 6 cada uno con probabilidad 1/6



Función de probabilidad f(x)



Función de distribución F(x)

Función de Distribución cont.

Algunos problemas de probabilidad están relacionados con la probabilidad $P(a < X \leq b)$ de que X asuma algún valor en un intervalo. Observar que:

$$P(a < X \leq b) = F(b) - F(a)$$

Los sucesos $X \leq a$ y $a < X \leq b$ son mutuamente excluyentes (ya que $\{X \leq a\} \cap \{a < X \leq b\} = \varnothing$). Entonces:

$$F(b) = P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$
$$= F(a) + P(a < X \leq b)$$

En el ejemplo de los dos dados, calcula la probabilidad de que los dos dados sumen al menos 4 pero no más de 8.

$$P(3 < X \leq 8) = F(8) - F(3) = 26/36 - 3/36 = 23/36$$

Propiedades de la Función de Distribución

$$F(+\infty) = \lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \le x) = P(\Omega) = 1$$

$$F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X \le x) = P(\varnothing) = 0$$

$$P(x_1 < X \le x_2) = F(x_2) - F(x_1)$$

F es monótona creciente.

F es continua por la derecha: la probabilidad de que la variable aleatoria discreta X tome un valor concreto es igual al salto de la función de distribución en ese punto.

Valor Esperado y Varianza de una variable aleatoria (v.a.)

- Valor esperado (o Esperanza Matemática)
  - Se representa mediante E[X] ó μ
  - Es el equivalente a la media

- Varianza
  - Se representa mediante VAR[X] o σ^2
  - Es el equivalente a la varianza
  - Se llama desviación estándar o típica a σ

Valor Esperado o Media de una distribución Discreta

$$\boxed{\mu = E[X] = \sum_i x_i P(X = x_i) = \sum_i x_i p(x_i)}$$

Calcular la esperanza de la variable aleatoria *X* en el ejemplo de los dos dados:

$$\mu = E(X) = \sum_{i=2}^{12} P(i) \cdot i =$$

$$\frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \ldots + \frac{6}{36} \cdot 7 + \ldots + \frac{1}{36} \cdot 12 = 7$$

Propiedades del Valor Esperado

- $E(C) = C$ si C es constante
- $E(CX) = C\,E(X)$
- $E(CX+B) = C\,E(X)+B$ C,B constantes
- $E(X+Y) = E(X)+E(Y)$ X,Y var. aleat.
- $E(X-Y) = E(X)-E(Y)$

Ganancia Esperada

A un juego de azar podemos asignarle una **variable aleatoria *X***, cuyos valores son **las ganancias correspondientes a los posibles resultados**. La esperanza matemática de la variable aleatoria *X* representa el **beneficio medio o ganancia media** que se obtiene en cada jugada cuando se juega un número elevado de veces.

Si la esperanza matemática es 0 se dice que el **juego** es **justo.**
Si es mayor que 0 se dice que el **juego** es **favorable** al jugador.
Si es menor que 0 se dice que perjudica al jugador y **no** es **favorable**.

Sea el juego que consiste en sacar una bola de una urna que contiene 7 bolas rojas y 3 bolas negras. Ganamos 500 pesos si la bola extraída es roja y pagamos 1500 pesos en el caso de que sea negra. ¿Qué podemos esperar si jugamos muchas veces?

Ejemplo

Espacio muestral $\Omega = \{R, N\}$. Consideramos las ganancias como positivas y las pérdidas negativas:

Variable aleatoria X		Función de probabilidad
$R \longrightarrow$	$500 \longrightarrow$	0.7
$N \longrightarrow$	$-1500 \longrightarrow$	0.3

$$\mu = 500 \times 0.7 - 1500 \times 0.3 = -100$$

Ganancia media

Momento de orden *k* de una v.a.

De forma más general podemos definir el valor esperado o media no solo para una variable aleatoria *X*, sino para cualquier función *T(X)* como:

$$\mu = E[T(X)] = \sum_i T(x_i) P(X_i = x_i) = \sum_i T(x_i) p(x_i)$$

Tomando como funciones a: $T_1(X) = X^k$

obtenemos los momentos de orden k centrados en el origen:

$$\boxed{m_k = E(X^k) = \sum_i x_i^k p(x_i)}$$

Momentos de orden *k* cont.

Y tomando como funciones a: $T_2(X) = (X-\mu)^k$

obtenemos los momentos de orden k centrados en la media de *X*:

$$\boxed{M_k = E((X-\mu)^k) = \sum_i (x-\mu)^k p(x_i)}$$

Observa que:

$$m_1 = E(X) = \mu$$

$$M_1 = E(X-\mu) = E(X) - \mu = 0$$

Varianza y Desviación Estándar

Varianza

$$\boxed{\sigma^2 = Var(X) = M_2 = E((X-\mu)^2) = \sum_i (x_i - \mu)^2 P(X = x_i)}$$

Desviación estándar o típica

$$\sigma = \sqrt{Var(X)}$$

Ambas miden la "dispersión de los datos". Observar que la desviación estándar lo hace con las mismas unidades que los propios datos.

Ejemplo

X	$P(X)$	$X-\mu$	$(X-\mu)^2$	$(X-\mu)^2 \cdot P(X)$
-1	0.1	-2	4	0.4
0	0.2	-1	1	0.2
1	0.4	0	0	0.0
2	0.2	1	1	0.2
3	0.1	2	4	0.4
				1.2

$$\sigma^2 = \sum_i (x_i - \mu)^2 P(x_i) = 1.2$$

$$\sigma = \sqrt{Var(X)} = 1.10$$

Ejemplo

Calcular la varianza y desviación estándar de la variable aleatoria X en el ejemplo de los dos dados:

$$\mathrm{Var}\ (X) = \sum_{i=2}^{12} P(i)\cdot(i-7)^2 = \frac{1}{36}\cdot(2-7)^2 + \frac{2}{36}\cdot(3-7)^2$$

$$+\ldots+\frac{1}{36}\cdot(12-7)^2 = 5.83$$

$$\sigma = \sqrt{Var(X)} = \sqrt{5.83} = 2.41$$

Propiedades de la Varianza

$$Var(X) = \sigma^2 = \sum_i (x_i - \mu)^2 p(x_i) =$$

$$\sum_i (x_i{}^2 + \mu^2 - 2\mu x_i) p(x_i) =$$

$$\sum_i x_i{}^2 p(x_i) + \mu^2 - 2\mu \sum_i x_i p(x_i) =$$

$$E(X^2) + \mu^2 - 2\mu^2 = E(X^2) - (E(X))^2$$

$$\boxed{\sigma^2 = E(X^2) - (E(X))^2}$$

La varianza es una función de los dos primeros momentos

Propiedades de la Varianza cont.

$\mathrm{Var}(C) = 0$ si C es constante

$$\mathrm{Var}(CX) = C^2\,\mathrm{Var}(X)$$

$$\mathrm{Var}(CX+B) = C^2\,\mathrm{Var}(X)$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

X,Y var. aleat.

Distribución Bernoulli

Experimento de **Bernoulli**: sólo son posibles dos resultados: éxito o fracaso. Podemos definir una variable aleatoria discreta X tal que:

$$\text{éxito} \to 1$$
$$\text{fracaso} \to 0$$

Si la probabilidad de éxito es p y la de fracaso $q = 1 - p$, podemos construir una función de probabilidad:

$$P(X = x) = p^x q^{1-x} \qquad x = 0,1$$

Evidentemente:

$$\sum_{x=0}^{1} P(X = x) = P(X = 0) + P(X = 1) = p + q = 1$$

Esperanza y Varianza de la distribución Bernoulli

$$E[X] = \mu = \sum_{x=0}^{1} x\,P(X = x) =$$

$$0 \cdot P(X = 0) + 1 \cdot P(X = 1) = p$$

$$Var(X) = E[X^2] - (E[X])^2 = \sum_{x=0}^{1} x^2\,P(X = x) - p^2$$

$$= 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) - p^2 =$$

$$p - p^2 = p(1 - p) = pq$$

Distribución Binomial

La distribución binomial aparece cuando estamos interesados en el número de veces que un suceso A ocurre (éxitos) en n intentos independientes de un experimento.

Por ejemplo: número de caras en n lanzamientos de una moneda.

Si A tiene probabilidad p (probabilidad de éxito) en un intento, entonces $q = 1-p$ es la probabilidad de que A no ocurra (probabilidad de fracaso).

En nuestro ejemplo de la moneda, $p = 0.5$ es la probabilidad de que salga cara y $q = 1-p = 1 - 0.5 = 0.5$ es la probabilidad de que no salga cara.

Distribución Binomial cont.

Supongamos que el experimento consta de n intentos y definamos la variable aleatoria:

$$X = \textit{Número de veces que ocurre } A.$$

En nuestro ejemplo: X = Número de veces que sale cara.

Entonces X puede tomar los valores 0, 1, 2, ... n.

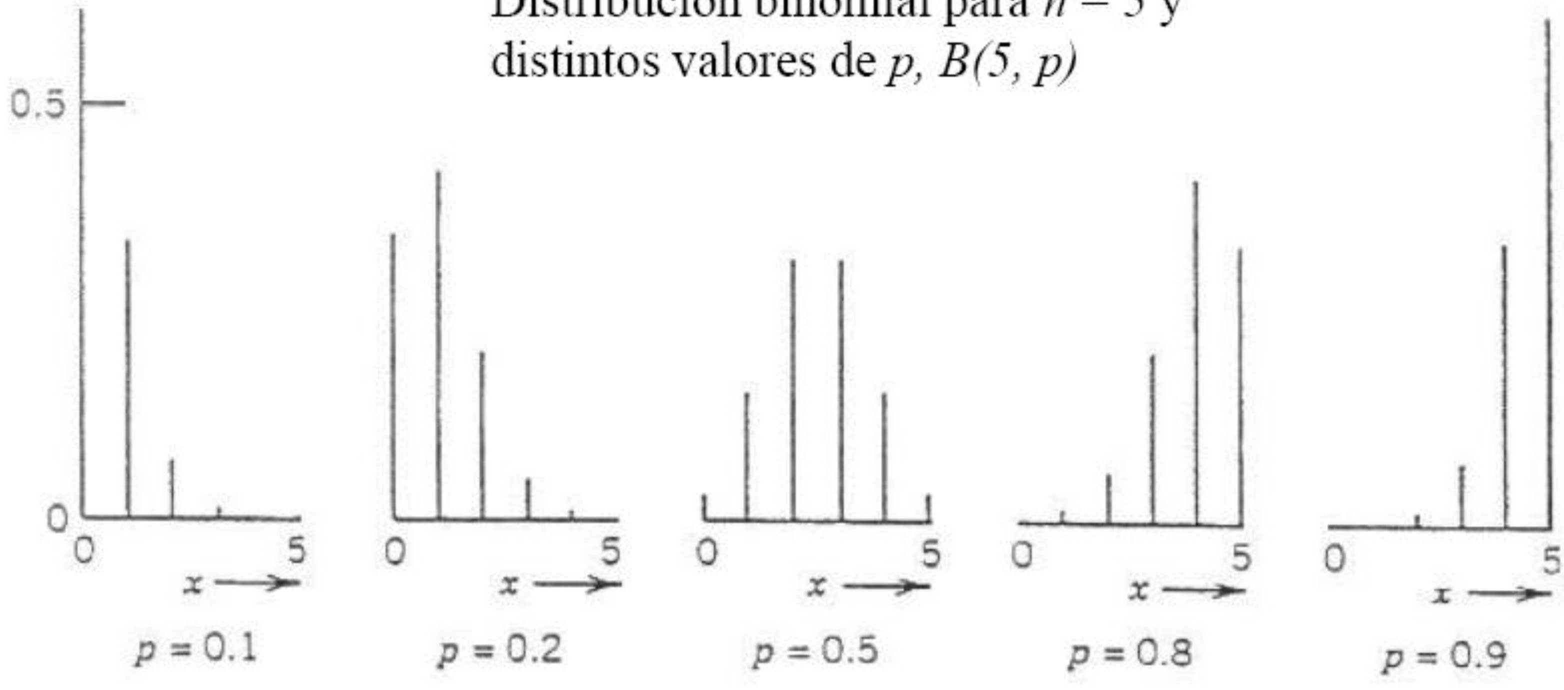
Si consideramos uno de estos valores, digamos el valor x , i.e. en x de los n intentos ocurre A y en n-x no. Entonces la probabilidad de cada posible ordenación es $p^x q^{n-x}$ y existen $\binom{n}{x}$ idénticas ordenaciones.

Distribución Binomial cont.

La distribución de probabilidad $P(X = k)$ será:

$$\boxed{B(n,p):\ p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}}$$

Distribución binomial para $n = 5$ y distintos valores de p, $B(5, p)$



Ejemplo

¿Cuál es la probabilidad de que en una familia de 4 hijos exactamente 2 sean niñas?

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$p = 0.5; \qquad n = 4; \qquad x = 2$$

$$p(2) = \binom{4}{2} (0.5)^2 (1\text{-}0.5)^{4\text{-}2}$$

$$= \frac{4!}{(4-2)!2!} 0.5^4 = \frac{4 \times 3 \times 2 \times 1}{(2 \times 1) \cdot (2 \times 1)} 0.5^4$$

$$= 6 \times 0.5^4 = 0.375$$

Ejemplo

Si una décima parte de personas tiene cierto grupo sanguíneo, ¿cuál es la probabilidad de que entre 100 personas escogidas al azar exactamente 8 de ellas pertenezcan a este grupo sanguíneo?

$$p(x) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

$$p = 0.1; \qquad n = 100; \qquad x = 8$$

$$p(8) = \binom{100}{8} (0.1)^{8}(1\text{-}0.1)^{92} = 0.115$$

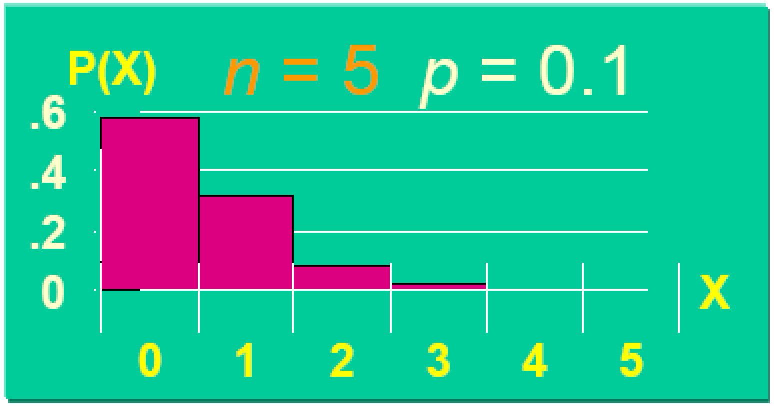
Características de la Distribución Binomial

Media

$\mu = E(X) = n\,p$

$\mu = 5 \cdot 0.1 = 0.5$

$\mu = 5 \cdot 0.5 = 0.25$

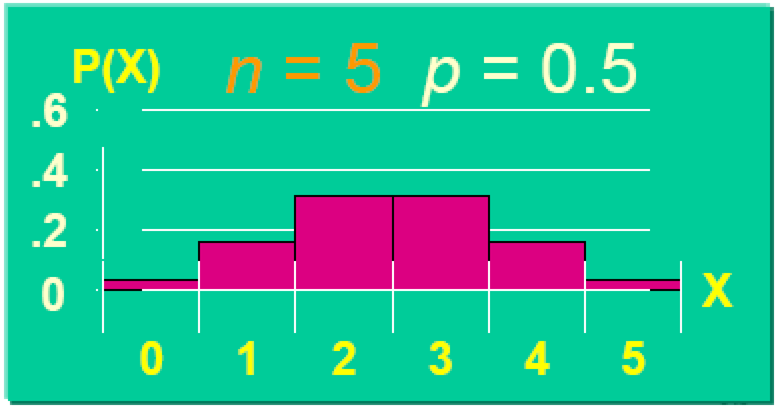


Desviación estándar

$$\sigma = \sqrt{np(1-p)}$$

$$\sigma = \sqrt{5 \cdot 0.1 \cdot (1-0.1)} = 0.67$$

$$\sigma = \sqrt{5 \cdot 0.5 \cdot (1-0.5)} = 1.1$$



Distribución Geométrica

Consideremos el siguiente experimento: Comenzamos de un experimento de Bernoulli donde la probabilidad de que ocurra un suceso es *p* (**éxito**) y la probabilidad de que no ocurra es *q=1-p* (**fracaso**). Repetimos el experimento hasta conseguir el **primer éxito**. Definimos la variable aleatoria *X*, como el número de fracasos hasta que se obtiene el primer éxito. Entonces:

$$\boxed{\begin{array}{l} G(p):\ P(X=x)=pq^{x}, \\ x=0,1,2,\ \ldots \end{array}}$$

Distribución Binomial Negativa (Pascal)

Consideremos el siguiente experimento:

Partimos de un experimento de Bernoulli donde la probabilidad de que ocurra un suceso es *p* (éxito) y la probabilidad de que no ocurra *q = 1- p (fracaso)*. Repetimos nuestro experimento hasta conseguir el r-ésimo éxito. Definimos la variable aleatoria *X*, como el número de fracasos *x* hasta que se obtiene el r-ésimo éxito. Entonces:

$$\boxed{BN(r,p):\ P(X=x)=\binom{x+r-1}{x}p^r q^x, \quad x=0,1,2,\ \ldots}$$

Se denomina binomial negativa porque los coeficiente provienen de la serie binomial negativa:

$$p^{-x}=(1\text{-}q)^{-x}$$

Seleccionar al azar *con reemplazo*

Elegir al azar con reemplazo significa que escogemos al azar un elemento de un conjunto y lo regresamos para elegir de nuevo al azar. Esto garantiza la independencia de las elecciones y nos lleva a una distribución binomial.

Si una caja contiene N bolas de las cuales A son rojas, entonces la probabilidad de escoger al azar una bola roja es: $p = A/N$

Si repetimos el experimento sacando n bolas con reemplazo la probabilidad de que x sean rojas es:

$$P(x) = \binom{n}{x}\left(\frac{A}{N}\right)^{x}\left(1-\frac{A}{N}\right)^{n-x} \qquad (x = 0,1,\ldots n)$$

Seleccionar al azar *sin reemplazo*

Elegir al azar sin reemplazo significa que no devolvemos el elemento elegido al azar al conjunto. De modo que las probabilidades de la siguiente elección dependen de las anteriores.

Si repetimos el experimento anterior sacando n bolas sin reemplazo, ¿cuál será ahora la probabilidad de que x sean rojas?

$$\text{Casos posibles} = \binom{N}{n}$$

Para calcular los casos favorables observa que:
$N = A + (N - A)$. De las A bolas rojas tomaremos x y de las $N - A$ bolas no rojas tomaremos $n - x$.

Distribución Hipergeométrica

$\binom{A}{x}$ = diferentes formas de tomar x bolas rojas de A

$\binom{N-A}{n-x}$ = diferentes formas de tomar $n-x$ bolas no rojas de $N-A$

$$\text{Casos favorables} = \binom{A}{x}\binom{N-A}{n-x}$$

$$H(n, N, A) : P(x) = \frac{\binom{A}{x}\binom{N-A}{n-x}}{\binom{N}{n}} \qquad (x = 0, 1, \ldots n)$$

Ejemplo

Queremos seleccionar al azar dos bolas de una caja que contiene 10 bolas, tres de las cuales son rojas. Encuentrar la función de distribución de la variable aleatoria : X = *Número de bolas rojas en cada elección (con y sin reemplazo).*

Tenemos $N = 10,\ A = 3,\ N - A = 7,\ n = 2$

Escogemos con reemplazo:

$$p(x) = \binom{2}{x}\left(\frac{3}{10}\right)^{x}\left(\frac{7}{10}\right)^{2-x}, \quad p(0) = 0.49,\ p(1) = 0.42,\ p(2) = 0.09$$

Escogemos sin reemplazo:

$$p(x) = \frac{\binom{3}{x}\binom{7}{2-x}}{\binom{10}{2}} \qquad p(0) = p(1) = \frac{21}{45} \approx 0.47,\ p(2) = \frac{3}{45} \approx 0.07$$

Hipergeométrica y Binomial

Hipergeométrica		Binomial	
	$N = 24$	$n = 5$	
	$X = 8$	$p = 8/24 = 1/3$	
	$n = 5$		
x	P(x)	P(x)	Error
0	0.1028	0.1317	-0.0289
1	0.3426	0.3292	0.0133
2	0.3689	0.3292	0.0397
3	0.1581	0.1646	-0.0065
4	0.0264	0.0412	-0.0148
5	0.0013	0.0041	-0.0028

	$N = 240$	$n = 5$	
	$X = 80$	$p = 80/240 = 1/3$	
	$n = 5$		
x	P(x)	P(x)	Error
0	0.1289	0.1317	-0.0028
1	0.3306	0.3292	0.0014
2	0.3327	0.3292	0.0035
3	0.1642	0.1646	-0.0004
4	0.0398	0.0412	-0.0014
5	0.0038	0.0041	-0.0003

Observar que si N, A, N-A son grandes comparados con n no hay gran diferencia en qué distribución empleemos.
La distribución binomial es una aproximación aceptable a la hipergeométrica si $n < 5\%$ de N.

Distribución Poisson

Cuando en una distribución binomial el número de intentos (*n*) es grande y la probabilidad de éxito (*p*) es pequeña y *np* (la media de la distribución binomial) es finito y tiende a λ, entonces la distribución binomial converge a la distribución de Poisson:

$$\boxed{p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!},\ \ x = 0,1,2,\ldots \quad \lambda > 0}$$

Observar que si p es pequeña, el éxito es un "suceso raro".

Características de la Poisson

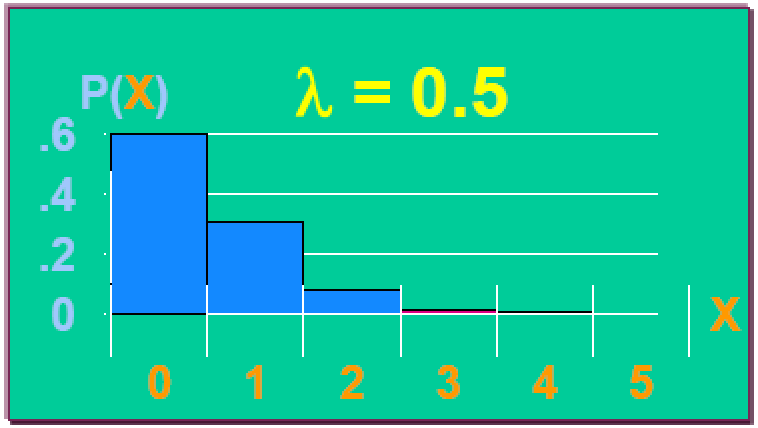
Media

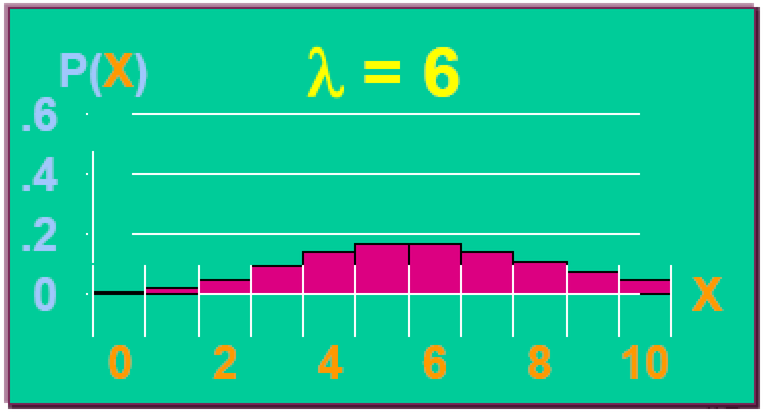
$$\mu = E(X) = \lambda$$

Varianza $Var(X)=\sigma^2$

$$\sigma^2 = \lambda$$

Nota: el máximo de la distribución se encuentra en $x \approx \lambda$

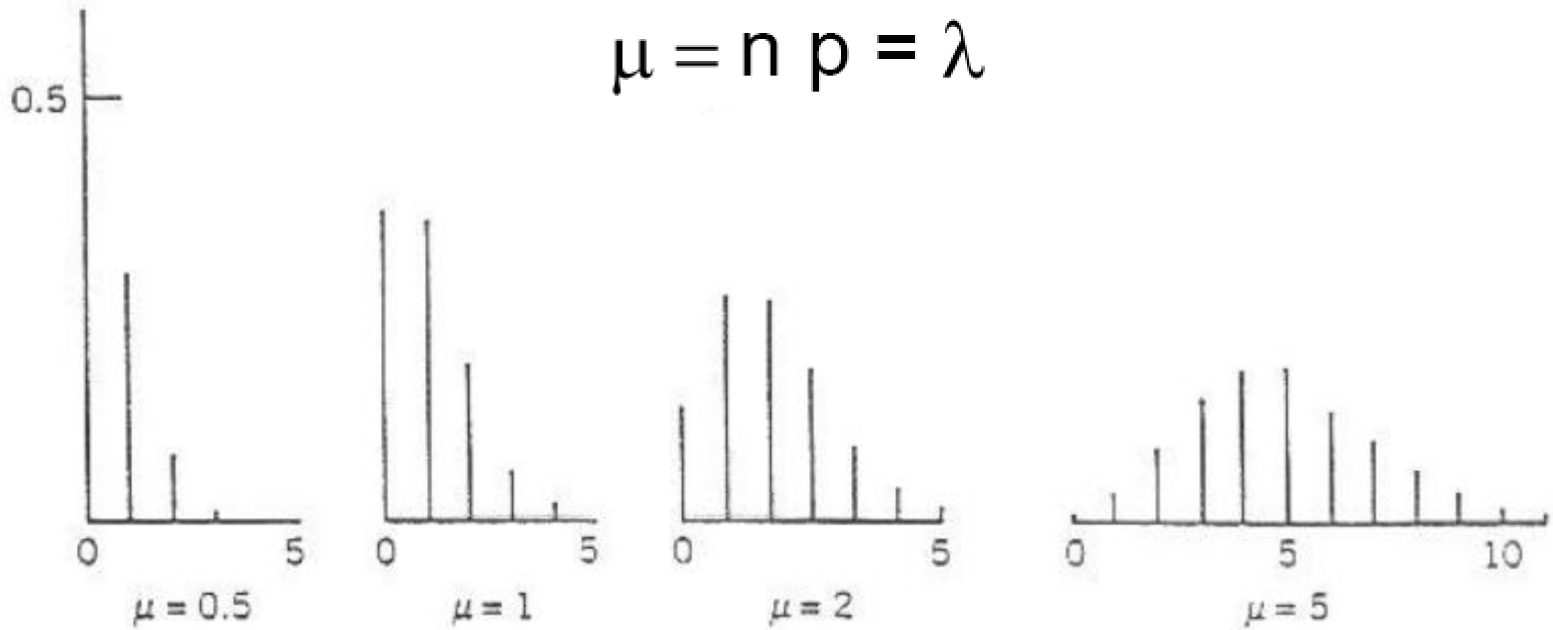




Características de la Poisson

La distribución de Poisson se obtiene como aproximación de una distribución binomial con la misma media, para 'n grande' (n > 30) y 'p pequeño' (p < 0,1). Queda caracterizada por un único parámetro μ (que es a su vez su media y varianza).

$$\mu = n\,p = \lambda$$



Distribución de Poisson para varios valores de μ.

Ejemplo

Si la probabilidad de fabricar un artículo defectuoso es $p = 0.01$, ¿cuál es la probabilidad de que en un lote de 100 artículos contenga más de 2 artículos defectuosos?

La distribución binomial nos daría el resultado exacto:

$$P(A^c) = \binom{100}{0}\left(\frac{99}{100}\right)^{100} + \binom{100}{1}\left(\frac{99}{100}\right)^{99}\left(\frac{1}{100}\right) + \binom{100}{2}\left(\frac{99}{100}\right)^{98}\left(\frac{1}{100}\right)^{2}$$

$$= 0.9206$$

El suceso complementario A^c: *No más de 2 artículos defectuosos* puede aproximarse con una distribución de Poisson con $\mu = np = 1$, sumando p(0) + p(1) + p(2).

$$P(A^c) \approx e^{-1}(1 + 1 + \tfrac{1}{2}) = 0.9197$$

La respuesta es 1-0.9197=0.0803

Proceso de Poisson

Un proceso Poisson es aquél compuesto de eventos discretos que son independientes en el espacio y/o en el tiempo. *Por ejemplo la llegada de fotones a un detector.*

Si el número de eventos esperados, el número medio de eventos en un intervalo de extensión *h* es *m*. *Por ejemplo el detector nos informa de la llegada en promedio de 20 fotones cada 5 segundos.*

Entonces $\lambda = h/m$, será la tasa de eventos por unidad de *h*. *En nuestro caso 4 fotones por segundo.*

La probabilidad de que ocurran *x* eventos en el intervalo *h* vendrá dada por la distribución de Poisson. *En nuestro ejemplo la probabilidad de que lleguen x fotones en 5 segundos.*

Ejemplo

La señal promedio recibida en un telescopio de una fuente celeste es de 10 fotones por segundo. Calcular la probabilidad de recibir 7 fotones en un segundo dado.

Una distribución de Poisson con $\mu = 10$.

$P(7) = 10^7\ e^{-10} / 7! = 0.09$, es decir 9%

Parece muy baja. Comparemos con el valor de máxima probabilidad que ocurrirá para $x = 10$:

$\mu = 10$ $\qquad P(10)=10^{10} \times e^{-10} / 10! = 0.125$, es decir 12.5%

Las probabilidades poissonianas para un número de eventos dado, son siempre pequeñas, incluso en el máximo de la distribución de probabilidad.

Ejemplo

Si en promedio, entran 2 ambulancias por minuto en un hospital, ¿cuál es la probabilidad de que durante un minuto entren 4 o más ambulancias?

Si asumimos que un minuto puede dividirse en muchos intervalos cortos de tiempo independientes y que la probabilidad de que una ambulancia entre en uno de esos intervalos es p (que para un intervalo pequeño será también pequeño) podemos aproximar la distribución a una Poisson con $\mu = np = 2$.

El suceso complementario *"entran 3 ambulancias o menos"* tiene probabilidad:

$$P(A^c) \approx p(0) + p(1) + p(2) + p(3) = e^{-2}\left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}\right) = 0.857$$

y la respuesta es 1 – 0.857 = 0.143