Estadística
Modelos probabilísticos discretos

Universidad Europea de Madrid

Laureate International Universities

MODELOS ALEATORIOS

- Al considerar variables aleatorias distintas caemos en la cuenta de que sus comportamientos respecto a la distribución de probabilidad, a sus momentos, etc. **son iguales**.
- Existen algunas distribuciones de probabilidad que sirven de **modelo** para numerosas aplicaciones prácticas.
- Modelos discretos y modelos continuos

Universidad Europea de Madrid

Laureate International Universities

Experimento de Bernoulli

Supongamos un experimento que sólo tiene dos resultados posibles:

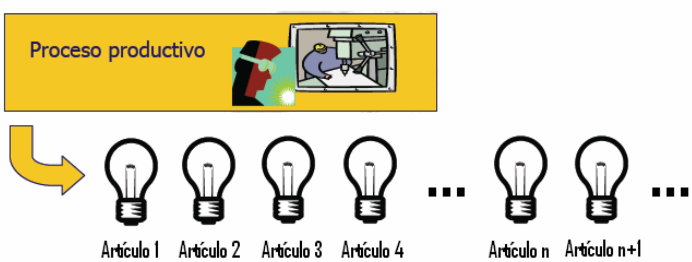
- Éxito/fracaso
- Observamos un atributo/no observamos dicho atributo

Ejemplos:

- Artículo aceptable/defectuoso
- Cliente satisfecho/no satisfecho
- Conexión/bloqueo
- Cara/cruz
- Votará/no votará
- Compra/no compra

Variable aleatoria de Bernoulli

Supongamos un proceso productivo que genera artículos que pueden ser defectuosos o aceptables ...



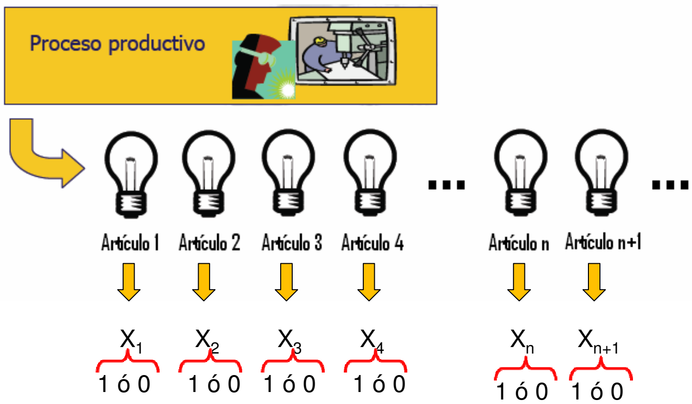
Asignamos una variable X_i a cada artículo

$$X_i = \begin{cases} 1 & si\ el\ artículo\ i\ es\ defectuoso \\ 0 & si\ el\ artículo\ i\ no\ es\ defectuoso \end{cases}$$

Variable aleatoria de Bernoulli

Variable aleatoria de Bernoulli

Supongamos un proceso productivo que genera artículos que pueden ser defectuosos o aceptables ...

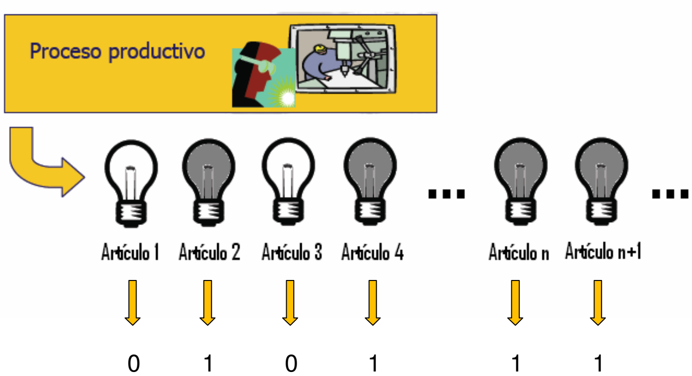


A la secuencia de estas v. aleatorias: X_1 ,X_2 ,X_3 ,…

PROCESO DE BERNOULLI

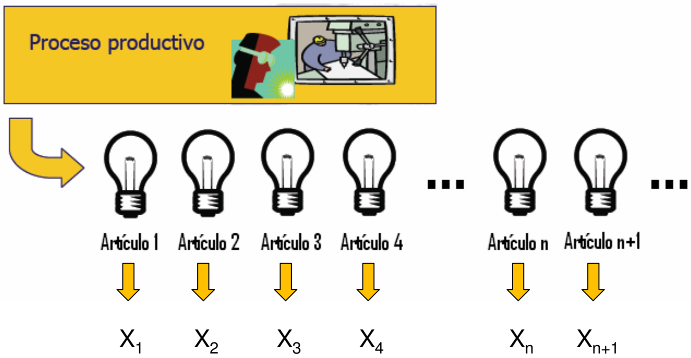
Universidad Europea de Madrid

Variable aleatoria de Bernoulli



Universidad Europea de Madrid

Variable aleatoria de Bernoulli

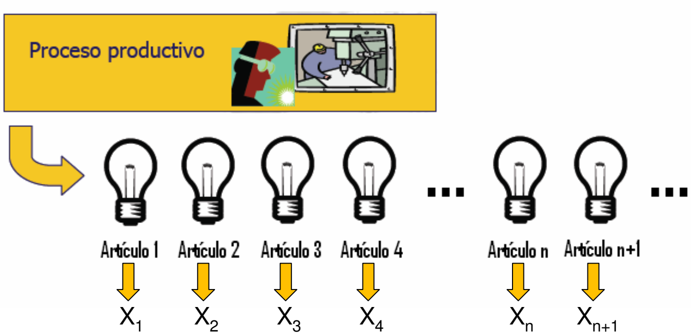


Todas las X_i tienen las mismas propiedades

X_1 , X_2 ,X_3 ,.... es una sucesión de Bernoullis idénticas

Universidad Europea de Madrid

Variable aleatoria de Bernoulli



$$X_i = \begin{cases} 1 & si\ el\ artículo\ i\ es\ defectuoso\ ...P(X_i = 1) = p \\ 0 & si\ el\ artículo\ i\ no\ es\ defectuoso\ ...P(X_i = 0) = 1 - p \end{cases}$$

$E[X_i] = p$ $\quad$ $Var[X] = p \cdot (1-p)$

Universidad Europea de Madrid

Ejemplo Bernoulli

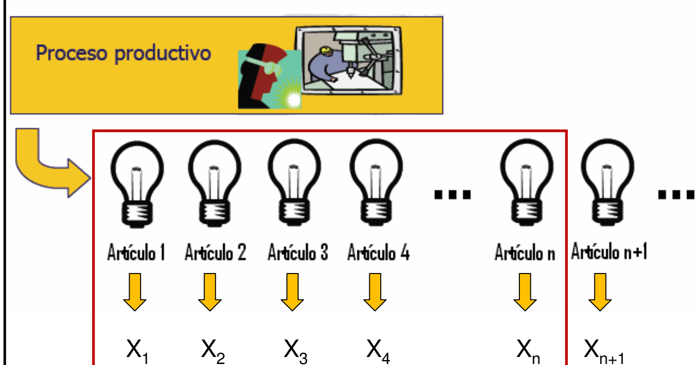
Se sabe que la probabilidad de que un individuo se contagie de gripe es 0.3. Determínese la media y la varianza de esta variable, y la probabilidad de que no padezca esta enfermedad.

Nos encontramos ante un fenómeno dicotómico, es decir, con dos posibles resultados: o se contagia de gripe o no se contagia. Designemos por éxito si se contagia y por fracaso si no lo hace. Así, según el enunciado la probabilidad de éxito es p=0.3 y la de fracaso 1-p=0.7. Este experimento es un experimento de Bernoulli, y por tanto

P(contagio)=0.3 P(no gripe)=0.7

Y E[X]=0.3 V[X]=0.3·0.7=0.21

Distribución binomial



Tomamos una muestra de **n** artículos

X = número de artículos defectuosos en n $\longrightarrow$ X ={ 0,1,2,3,4,...,n}

X = <u>**variable aleatoria BINOMIAL**</u>

Variable aleatoria binomial

X = variable aleatoria BINOMIAL se denota por   X ~ B(n , p)

El modelo BINOMIAL cuenta el número de éxitos que se producen en n experimentos de Bernoulli. Por lo tanto X puede tomar los valores:

{0,1,2,3,…,n}

Su función de masa viene dada por la fórmula:

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad para\ k = 0,1,\ldots,n$$

Función de masa

- Supongamos que repetimos el experimento 4 veces y contamos el número de éxitos. ¿Cuál es la probabilidad de lograr 2 éxitos?

EEFF, EFEF, EFFE, FEEF, FEFE, FFEE

P(EEFF)=pp(1-p)(1-p)=$p^2(1-p)^2$

P(EFEF)=p(1-p)p(1-p)= $p^2(1-p)^2$ $\Rightarrow$ P(2 éxitos)=$6p^2(1-p)^2$

P(EFFE)=p(1-p)(1-p)p= $p^2(1-p)^2$

……………………………

Variable aleatoria binomial

k éxitos (probabilidad p)

n-k fracasos (probabilidad 1-p)

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ \ para\ k = 0, 1, \ldots, n$$

Número de opciones: k éxitos en n repeticiones

Probabilidad de cada una de las opciones

Cálculo de los números combinatorios

Factorial de un número: $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$

Ejemplos:

$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$

$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

$7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$

Número combinatorio: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

n sobre k

Cálculo de números combinatorios

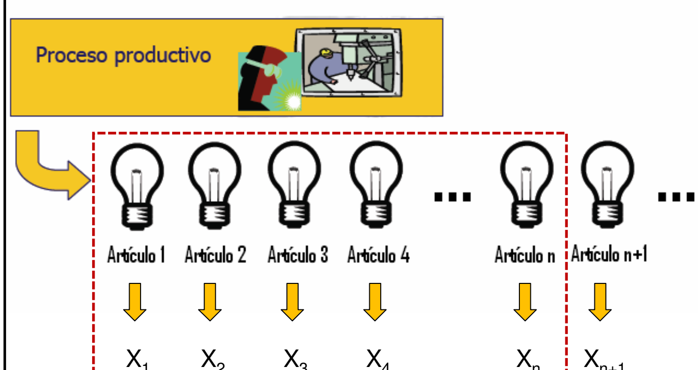
- Ejemplos:

$$\binom{5}{3} = \frac{5!}{3!\cdot 2!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 2\cdot 1} = 10$$

$$\binom{7}{2} = \frac{7!}{2!\cdot 5!} = \frac{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{2\cdot 1\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} = 21$$

$$\binom{9}{3} = \frac{9!}{3!\cdot 6!} = \frac{9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} = 84$$

Variable aleatoria binomial



X = número de artículos defectuosos en n

$$X = X_1 + X_2 + X_3 + \cdots + X_n$$

Suma de n Bernoullis

Variable aleatoria binomial

Propiedades:

1. $X = X_1 + X_2 + X_3 + \cdots + X_n$, con X_1 , X_2 , X_3 , X_n variables Bernoulli independientes

2. $E[X] = E[X_1 + X_2 + X_3 + \cdots + X_n] = \sum E[X_i] = n \cdot p$

3. $Var[X] = Var[X_1 + X_2 + X_3 + \cdots + X_n] = \sum Var[X_i] = n \cdot p \cdot (1-p) = n \cdot p \cdot q$

4. Si $X_1 \sim B(n_1 , p)$ y $X_2 \sim B(n_2 , p)$ son independientes, entonces $X_1 + X_2 \sim B(n_1 + n_2, p)$

Universidad Europea de Madrid

Variable aleatoria binomial

Ejemplo:

Se sabe por información histórica que cuando la máquina M opera en condiciones normales produce un 0.5% de artículos defectuosos.

Después de reparar una avería, se toman n=200 artículos producidos por dicha máquina y 5 de ellos resultan ser defectuosos

¿Deberíamos preocuparnos?

X = número de artículos defectuosos (número de exitos) en n=200

Si la máquina funciona correctamente, supondremos que

$$X \sim B(200,0.005)$$

¿Es fácil encontrar 5 ó más artículos defectuosos en 200?

Universidad Europea de Madrid

Variable aleatoria binomial

Ejemplo:

$$p(0) = P(X = 0) = \binom{200}{0} 0{,}005^{0}(1-0{,}005)^{200} = 0{,}367$$

$$p(1) = P(X = 1) = \binom{200}{1} 0{,}005^{1}(1-0{,}005)^{199} = 0{,}369$$

$$p(2) = P(X = 2) = \binom{200}{2} 0{,}005^{2}(1-0{,}005)^{198} = 0{,}184$$

$$p(3) = P(X = 3) = \binom{200}{3} 0{,}005^{3}(1-0{,}005)^{197} = 0{,}061$$

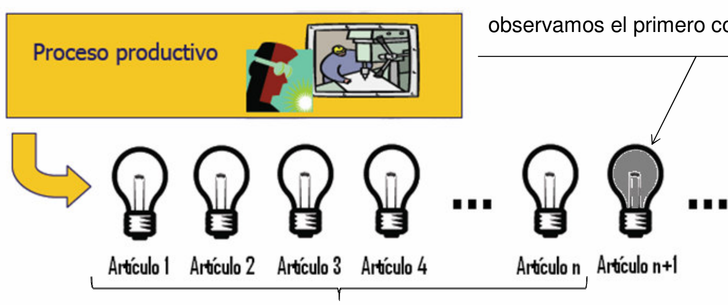
$$p(4) = P(X = 4) = \binom{200}{4} 0{,}005^{4}(1-0{,}005)^{196} = 0{,}015$$

$P(X\geq5)=1-P(X\leq4)=1-(0.367+0.369+0.184+0.061+0.015)=0.004$

Si la máquina funcionase bien, encontrar 5 artículos defectuosos es muy improbable. Debemos preocuparnos por el estado de la máquina tras la reparación.

Variable aleatoria geométrica

Contamos los artículos hasta que observamos el primero con el atributo



Y = número de artículos hasta para observar el primero con el atributo

y=1,2,3,...

Y: Variable aleatoria GEOMÉTRICA

Variable aleatoria geométrica

Y = variable aleatoria GEOMÉTRICA se denota por Y ~ G(p)

El modelo GEOMÉTRICO cuenta el número de fracasos hasta obtener el primer éxito . Su función de masa viene dada por:

$$P(Y = k) = (1 - p)^k p$$

Medidas características:

1. E[Y] = 1-p/p
2. Var[Y] = 1-p/p^2

Función de masa

- Veamos, por ejemplo, la probabilidad de obtener el primer éxito en el 4 intento (o en la 4 repetición): obtener 3 fracasos antes del éxito

FFFE

P(X=3)= (1-p)(1-p)(1-p)p=$(1-p)^3$p

Variable aleatoria geométrica

- Una característica del modelo geométrico es que no tiene memoria, es decir, habiendo realizado a repeticiones del suceso, la probabilidad de que la característica A (éxito) se presente a partir de la repetición a+b es igual a calcular la probabilidad de obtener el éxito después de b repeticiones, olvidándose de que se han producido a repeticiones sin éxito anteriormente.

Universidad Europea de Madrid

Variable aleatoria geométrica

Un triplista de un equipo de baloncesto hace canasta de tres puntos en el 80% de los intentos que realiza. ¿Cuál es la probabilidad de que en un partido logre su primer triple en el quinto intento?.

Lo que queremos es que falle cuatro veces antes de encestar en el quinto tiro. Sabemos que la probabilidad de encestar es de 0.8.

$$P(X=4)= (0.2)^4(0.8)=0.00128$$

La variable aleatoria que cuenta el número de tiros antes de alcanzar el éxito (hacer canasta) se distribuye como una geométrica de parámetro p=0.8

Universidad Europea de Madrid

Variable aleatoria binomial negativa

- Una variable aleatoria que cuenta el número de fracasos hasta obtener el k-ésimo éxito en la k+x repetición sigue una distribución binomial negativa. X~BN(k,p), donde p es la probabilidad de éxito.

- Un caso particular de esta distribución (para k=1) es la distribución geómetrica.

Variable aleatoria binomial negativa

- Función de masa:

x fracasos con probabilidad 1-p

$$P(X=x)=\binom{k+x-1}{x}(1-p)^x p^k$$

Maneras de colocar x fracasos en k+x-1 repeticiones

K éxitos con probabilidad p

Función de masa de BN

- X es la variable aleatoria que cuenta el número de fracasos hasta obtener el k-ésimo éxito:

Supongamos que queremos conseguir 3 éxitos en 5 repeticiones:

Podemos

Obtener los 3 éxitos en las tres primeras repeticiones → EEE → X=0

Obtener los 3 éxitos en 4 repeticiones → FEEE EFEE EEFE → X=1

Obtener los 3 éxitos en 5 repeticiones → FFEEE EFFEE EFEFE FEFEE FEEFE EEFFE

Universidad Europea de Madrid

Variable aleatoria binomial negativa

- Propiedades:

  Esperanza

  $E[X]=k(1-p)/p$

  Varianza

  $V[X]=k(1-p)/p^2$

Universidad Europea de Madrid

Variable aleatoria binomial negativa

- **Ejemplo:** En los play-off de la NBA americana, el vencedor de cada eliminatoria final es el que logra primero la cuarta victoria en 7 encuentros. ¿Cuál es la probabilidad de que un equipo dispute como mucho 6 partidos, si su porcentaje de partidos ganados es 60%?

En cada partido pueden: ganar (éxito) con probabilidad p=0.6 o perder (fracaso) con probabilidad 1-p=0.4. Además suponemos que los partidos son independientes. Lo que queremos es obtener k=4 victorias (4 éxitos).

X={número de partidos que disputamos para lograr la 4 victoria}

X~BN(4,0.6)

Variable aleatoria binomial negativa

Para no llegar al séptimo partido, tienen:

- Que ganar los cuatro primeros partidos.
- Que perder uno de cuatro partidos y ganar el quinto.
- O perder dos de cinco partidos y ganar el sexto.

P(no llegar al séptimo partido)= P(X=0)+P(X=1)+P(X=2)=

$$=\binom{3}{0}0.6^4+\binom{4}{1}0.4\ 0.6^4+\binom{5}{2}0.4^2 0.6^4=0.54$$

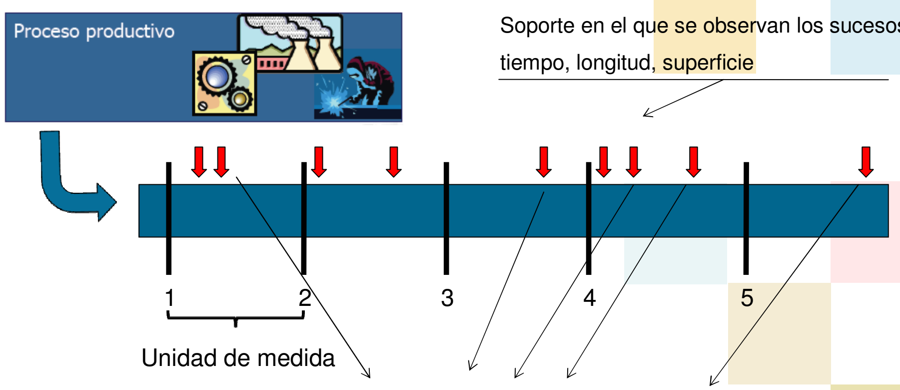
Variable aleatoria de Poisson

La distribución de **Poisson** se utiliza para modelizar algunas situaciones como la distribución de usuarios que llegan a una ventanilla, el número de llamadas que recibe una centralita, el número de erratas de una página, en general, sucesos que ocurren en un intervalo de tiempo o en una región del espacio. A esta distribución también se la conoce como de los *sucesos raros*: sucesos con poca probabilidad de repetirse un número de veces relativamente alto.

Las primeras aplicaciones son ciertamente anecdóticas ya que, por un lado el matemático polaco Bortkievicz la utilizó al estudiar en 1898, el número de soldados del ejercito prusiano muertos por una coz de mula. También se utilizó para modelizar los partos múltiples en Londres.

Variable aleatoria de Poisson

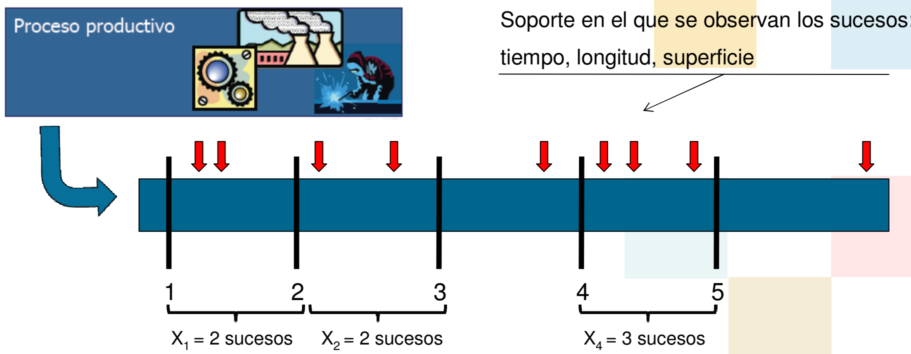
Soporte en el que se observan los sucesos: tiempo, longitud, superficie



X = número de sucesos observados (p.e. defectos) en cada unidad de tiempo

Proceso de Poisson: los sucesos aparecen con una media estable y de forma independiente sobre un soporte continuo: **no hay memoria**

Variable aleatoria de Poisson



X = número de atributos observados (p.e. defectos) en cada unidad de medida

VX = {0,1,2,….}

X es una **variable aleatoria de Poisson**

$X \sim P(\lambda)$

Variable aleatoria de Poisson

X es una variable aleatoria de Poisson $\longrightarrow$ $X \sim P(\lambda)$

• Función de probabilidad:

$$p(k) = P(Y = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \;\; para\; k = 0,1,2, \ldots$$

• Medidas características:

$E[X] = \lambda$

$Var[X] = \lambda$

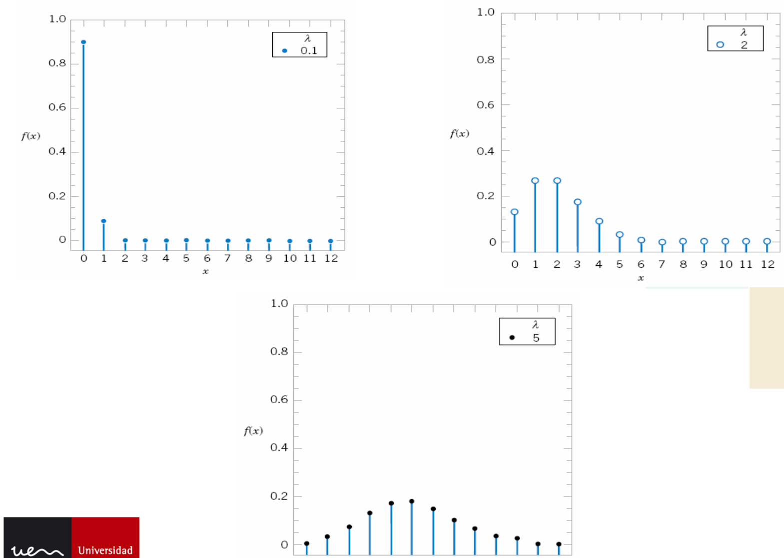
• Aditividad:

Sea X_1 y X_2 dos v. aleatorias de Poisson independientes

$X1 \sim P(\lambda_1)$; $X2 \sim P(\lambda_2)$

Entonces $Y = X_1 + X_2$ es también Poisson $\longrightarrow$ $Y \sim P(\lambda_1 + \lambda_2)$

Variable aleatoria de Poisson



Variable aleatoria de Poisson

Ejemplo 1:

Un servidor de una pequeña red recibe una media de 7 accesos al minuto. Suponiendo que los accesos a dicho servidor suceden de forma independiente y con un ritmo medio constante.

Se quiere calcular la probabilidad de que reciba más de 10 accesos en un minuto, que es el número de accesos a partir del cual el servidor tendría un rendimiento deficiente.

X = número de accesos en un minuto $\longrightarrow$ X ~ P(7)

$$P(X > 10) = 1 - P(X \leq 10) = 1 - \sum_{k=1}^{10} P(X = k)$$

$$= 1 - \sum_{k=1}^{10} \frac{7^k}{k!} e^{-7} = 0{,}099$$

Variable aleatoria de Poisson

Ejemplo 2:

Un sistema para regular la cola de impresión de una impresora colectiva sólo admite un máximo de 4 ficheros en la cola de impresión. Por tanto, si llegan 5 o más ficheros, la cola o bien se bloquea o se pierden los ficheros que pasen del cuarto. Se supone que la llegada de ficheros a esa impresora sigue una Poisson de media 1 fichero por minuto.

a) ¿Cuál es la probabilidad de que en un minuto lleguen 5 ó más ficheros?

b) ¿Cuál es la probabilidad de que en 3 minutos no llegue ningún fichero?

Universidad Europea de Madrid

Variable aleatoria de Poisson

Solución:

a) ¿Cuál es la probabilidad de que en un minuto lleguen 5 ó más ficheros? X = número de ficheros que llegan a la impresora en un minuto

$$X \sim P(1)$$

$P(X > 5) = 1 – P(X = 0) – P(X = 1) – P(X = 2) – P(X = 3) – P(X = 4)$

$= 1 – 0,9963 = 0,0037$

b) ¿Cuál es la probabilidad de que en 3 minutos no llegue ningún fichero?

Y = número de ficheros que llegan a la impresora en 3 minutos

$$Y \sim P(3)$$

$P(Y = 0) = 3^0/0! \cdot e^{-3} = 0,05$

Universidad Europea de Madrid

Poisson como aproximación de Binomial

En la práctica nos encontraremos con situaciones en que se aplica la distribución binomial con una probabilidad de éxito pequeña y n grande, entonces, la distribución Binomial puede aproximarse mediante la distribución de Poisson.

Desde el punto de vista práctico, la aproximación es buena cuando $p<0.1$ y $np<5$.

En estos casos tomamos

$$\lambda = np$$

Universidad Europea de Madrid