Probabilidad y Estadística

Unidad 7: Introducción a los procesos estocásticos

Lic. Maite San Martín

Junio 2025

Variables aleatorias y procesos estocásticos
Introducción

En algunos experimentos aleatorios, el resultado que se observa en una función sobre el tiempo o el espacio. Por ejemplo,

- el tráfico en un servidor de streaming varía a lo largo del día,
- en un sistema de procesamiento de imágenes, la intensidad y el color de la imagen varían sobre una región rectangular,
- el precio de una acción en la bolsa, que puede subir o bajar en cada instante con cierta probabilidad,
- la predicción diaria del clima en una ciudad es también una variable aleatoria que fluctúa a lo largo de los días.

Variables aleatorias y procesos estocásticos

Introducción

En algunas situaciones, pueden ser de interés dos o más funciones del tiempo. Por ejemplo, la temperatura en una determinada ciudad y la demanda que se ejerce sobre la compañía eléctrica local varían conjuntamente en el tiempo. O bien, la presión arterial sistólica y diastólica, así como el ritmo cardíaco y el nivel de glucosa en sangre varían aleatoriamente en una persona a lo largo del día (e incluso de su vida).

Las funciones de los ejemplos anteriores pueden considerarse como cantidades numéricas que evolucionan aleatoriamente en el tiempo o en el espacio. Por lo tanto, lo que realmente tenemos es una familia de variables aleatorias indexadas por la variable de tiempo o de espacio. En esta unidad abordaremos (aunque sea de forma introductoria) el estudio de los procesos estocásticos.

Variables aleatorias y procesos estocásticos

Ejemplo

Lanzamos una moneda al aire 6 veces. El jugador gana \$1 cada vez que sale cara (C), y pierde \$1 cada vez que sale cruz (F). Definimos:

$$X_i = \text{estado de cuentas del jugador después de la i−ésima jugada}$$

La familia de variables aleatorias $\{X_1, X_2, \dots, X_6\}$ constituye un proceso estocástico.

$$Y_i = \text{cara de la moneda observada en la i−ésima jugada}$$

La familia de variables aleatorias $\{Y_1, Y_2, \dots, Y_6\}$ constituye un proceso estocástico.

Variables aleatorias y procesos estocásticos
Definición

Un proceso estocástico $\{X(t),\ t \in T\}$ es una colección de variables aleatorias. Esto quiere decir que para cada momento $t \in T$, $X(t)$ es una variable aleatoria. Por ejemplo, $X(t)$ puede ser el número de clientes que entran a un supermercado hasta el momento t, o el número de clientes que están dentro del supermercado en un momento t, o el monto total de ventas que se registraron en el día t, etc.

Un proceso estocástico es una familia de variables aleatorias parametrizadas por el tiempo*. El conjunto T es llamado espacio parametral del proceso; si T es un conjunto numerable $\{X_n, n = 0,1,2, \dots\}$, entonces se trata de un proceso estocástico a tiempo discreto, caso contrario será un proceso a tiempo continuo $\{X(t), t \geq 0\}$.

*Si bien a partir de ahora vamos a referirnos a t como tiempo, todo lo enunciado en esta unidad también aplica a procesos estocásticos definidos sobre el espacio.

Variables aleatorias y procesos estocásticos
Definición

Un proceso estocástico es una familia de variables aleatorias parametrizadas por el tiempo* y definida sobre un espacio muestral Ω:

$$\{X_t(\omega): \Omega \to S, t \in T\}$$

donde S es el conjunto de valores posibles que pueden tomar el proceso, llamado espacio de estados. En términos genéricos, diremos que $X_t(\omega)$ es el estado del proceso al momento t para una realización ω en particular.

Podríamos pensar en X como una función de dos argumentos: $X(t, \omega)$. Para una realización específica $\omega = \omega_i$ del proceso estocástico tenemos una trayectoria posible del proceso -una función determinista, no aleatoria-, y para un instante $t = t_i$ en particular tenemos una variable aleatoria con su distribución de probabilidad a partir de todos los posibles valores que puede tomar el proceso en ese instante t_i.

Variables aleatorias y procesos estocásticos
Ejemplo

Lanzamos una moneda al aire 6 veces. El jugador gana \$1 cada vez que sale cara (C), y pierde \$1 cada vez que sale cruz (F). Entonces:

- $\Omega = \{CCCCCC, CCCCCF, CCCCFC, \dots, FFFFFF\}$
- $\#\Omega = 2^6 = 64$
- $P(\omega) = 1/64, \ \forall \omega \in \Omega$
- $T = \{1, 2, 3, 4, 5, 6\}$
- $S = \{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}$

Variables aleatorias y procesos estocásticos
Ejemplo

Lanzamos una moneda al aire 6 veces. El jugador gana \$1 cada vez que sale cara (C), y pierde \$1 cada vez que sale cruz (F). Para cada momento $t = 1, \dots, 6$ en el que nos paremos podemos definir una variable aleatoria X_t:

X_1	p_{X_1}
-1	0,5
1	0,5

X_2	p_{X_2}
-2	0,25
0	0,5
2	0,25

X_3	p_{X_3}
-3	0,125
-1	0,375
1	0,375
3	0,125

X_4	p_{X_4}
-4	0,0625
-2	0,25
0	0,375
2	0,25
4	0,0625

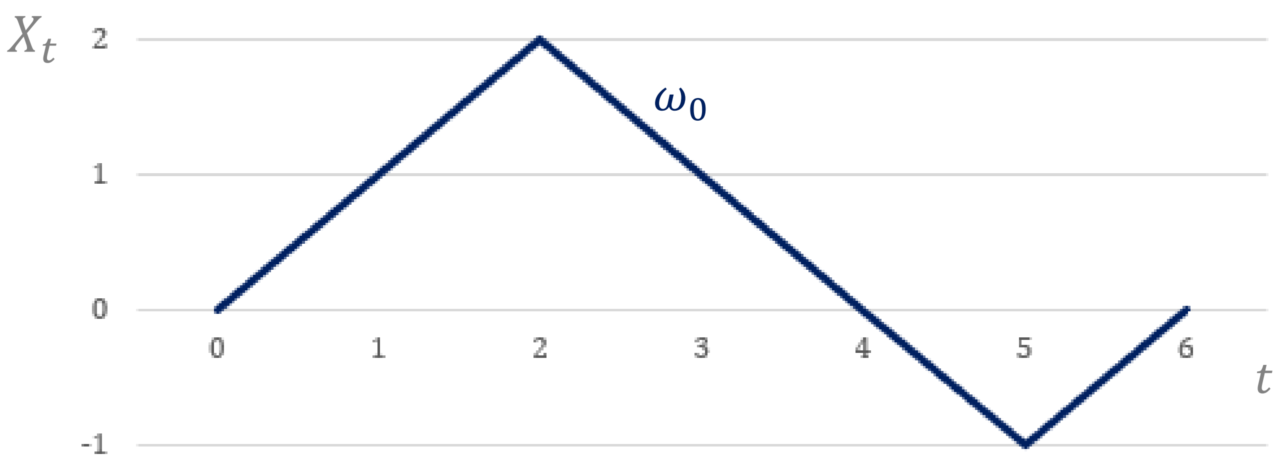
X_5	p_{X_5}
-5	0,03125
-3	0,15625
-1	0,3125
1	0,3125
3	0,15625
5	0,03125

X_6	p_{X_6}
-6	0,015625
-4	0,09375
-2	0,234375
0	0,3125
2	0,234375
4	0,09375
6	0,015625

Variables aleatorias y procesos estocásticos
Ejemplo

Lanzamos una moneda al aire 6 veces. El jugador gana \$1 cada vez que sale cara (C), y pierde \$1 cada vez que sale cruz (F).

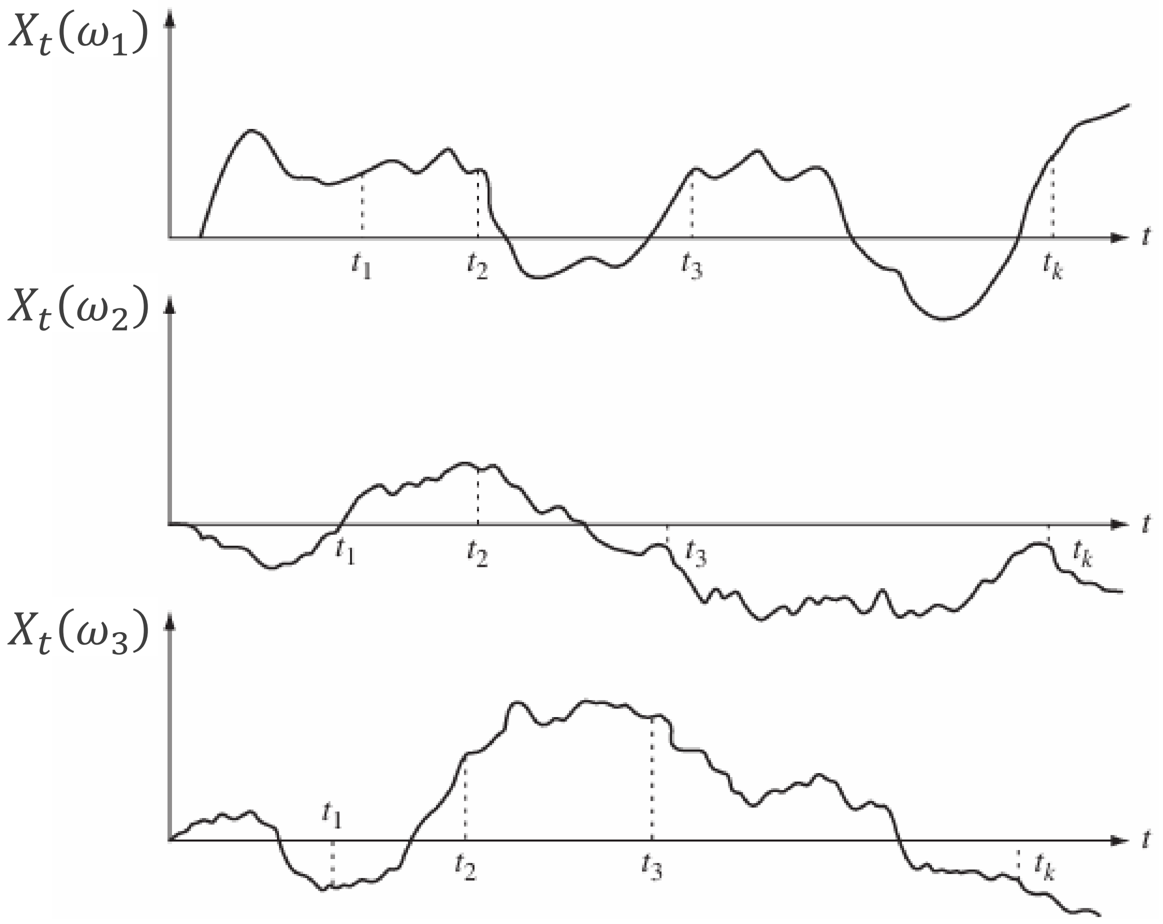
Si fijamos ω, por ejemplo $\omega_0 = CCFFFC$ obtenemos una secuencia de valores completamente determinista: $X_1(\omega_0) = 1, X_2(\omega_0) = 2, X_3(\omega_0) = 1, X_4(\omega_0) = 0, X_5(\omega_0) = -1, X_6(\omega_0) = 0$. Con estos valores es posible graficar la trayectoria del proceso para esta realización.



Variables aleatorias y procesos estocásticos
Ejemplo

A su vez, podríamos ejemplificar distintas trayectorias posibles $(\omega_1, \omega_2, \omega_3)$ de un proceso estocástico a tiempo continuo X_t. Para cada posible valor de t existe un valor de la variable aleatoria X_t en cada trayectoria.

Podría interesarnos estimar probabilidades, por ejemplo, de que el proceso tome un valor dentro de un intervalo al momento t_k.



Variables aleatorias y procesos estocásticos
Introducción

Existen distintos tipos de procesos estocásticos según en función de…

… distintos espacios parametrales,

… distintos espacios de estado,

… distintas relaciones de dependencia estocástica entre las variables aleatorias que conforman el proceso.

Durante esta unidad vamos a ver algunas familias particulares de procesos estocásticos.

Proceso Bernoulli
Introducción

Consideremos un experimento que consiste en una sucesión infinita de ensayos realizados en idénticas condiciones en el cual cada ensayo, todos independientes entre sí, puede tener solo dos resultados posibles: éxito (E) o fracaso (F) con $P(E) = p$ y $P(F) = 1 - p$ constantes a lo largo de todo el proceso. El espacio muestral asociado a un experimento de este tipo es

$$\Omega = \{\omega \ / \ \omega = (\omega_1, \omega_2, \dots, \omega_n, \dots), \omega_i = E \ o \ bien \ \omega_i = F, i \in \mathbb{N}\}$$

Si a la n-ésima repetición del ensayo pudiéramos asociarle una variable aleatoria X_n definida como $X_n(\omega) = \begin{cases} 0, si \omega_n = F \\ 1, si \ \omega_n = E \end{cases}$ con $P(X_n = 1) = p$ y $P(X_n = 0) = 1 - p$, entonces, este experimento puede modelarse como la sucesión de infinitos ensayos Bernoulli.

Proceso Bernoulli
Definición

Se denomina proceso Bernoulli a una colección numerable de variables aleatorias independientes

$$X_n(\omega) = \begin{cases} 0, & si \omega_n = F \\ 1, & si\ \omega_n = E \end{cases}$$

todas con distribución de probabilidad: $P(X_n = 1) = p$ y $P(X_n = 0) = 1 - p$.

Proceso Bernoulli
Ejemplo

En una bifurcación de una ruta, aproximadamente el 62% de los automóviles toma la rama izquierda.

Se define $X_n = \begin{cases} 0, & \text{si el n−ésimo automóvil toma la rama derecha} \\ 1, & \text{si el n−ésimo automóvil toma la rama izquierda} \end{cases}$

Se supone que los conductores eligen su camino independientemente de lo que hacen los otros, eso quiere decir que se puede considerar $X_1, X_2, \dots, X_n, \dots$ independientes con $P(X_n = 1) = 0{,}62 \;\; \forall n$.

Luego, el proceso $\{X_n, n \in \mathbb{N}\}$ es un proceso Bernoulli.

Proceso número de éxitos
Definición

Sea $\{X_n, n \in \mathbb{N}\}$ un proceso Bernoulli con $P(E) = p$. Definamos ahora la variable aleatoria

$$N_n = X_1 + X_2 + \cdots + X_n$$

que describe el número de éxitos en los primeros n ensayos Bernoulli. Si se define como condición inicial del proceso $N_0 = 0$ (es decir, al momento previo de iniciar el proceso Bernoulli no se contabilizaba ningún éxito), entonces el proceso estocástico $\{N_n, n \in \mathbb{N}_0\}$ es conocido como proceso estocástico número de éxitos, describe la cantidad de éxitos en un proceso Bernoulli y tiene a $\mathbb{N}_0$ como espacio de estados.

Proceso número de éxitos

Definición

Cabe destacar que si quisiéramos calcular $P(N_n = k)$, deberíamos tener en cuenta que esto significa que en n repeticiones del experimento Bernoulli se observaron exactamente k éxitos, por lo que

$$P(N_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \forall k = 0, 1, \dots, n$$

Es decir, si $\{N_n, n \in \mathbb{N}_0\}$ es un proceso estocástico número de éxitos, entonces $N_n \sim Bi(n, p)$.

Proceso número de éxitos
Ejemplo

En una bifurcación de una ruta, aproximadamente el 62% de los automóviles toma la rama izquierda.

Se define $X_n = \begin{cases} 0, & \text{si el n−ésimo automóvil toma la rama derecha} \\ 1, & \text{si el n−ésimo automóvil toma la rama izquierda} \end{cases}$

Se supone que los conductores eligen su camino independientemente de lo que hacen los otros, eso quiere decir que se puede considerar $X_1, X_2, \dots, X_n, \dots$ independientes con $P(X_n = 1) = 0{,}62 \ \ \forall n$. Luego, $N_n = X_1 + X_2 + \cdots + X_n$ describe la cantidad de automovilistas que tomaron la rama izquierda entre los primeros n autos, y el proceso $\{N_n, n \in \mathbb{N}_0\}$ es un proceso estocástico número de éxitos.

Proceso número de éxitos
Ejemplo

¿Cuál es la probabilidad de que entre los primeros 20 autos solo 5 de ellos tomen la rama izquierda?

$$P(X_{20} = 5) = \binom{20}{5} 0{,}62^5 0{,}38^{15} = 0{,}0007$$

De cada 50 autos que pasan por esa esquina, ¿cuántos se espera que tomen la rama izquierda?

$$E(X_{50}) = np = 50 * 0{,}62 = 31$$

Proceso número de éxitos
Propiedad Markoviana

La cantidad de éxitos k observados en n ensayos depende exclusivamente de la cantidad de ensayos (n) y no de instante en el que se comienza a observar el proceso. Si contáramos con dos variables aleatorias N_m y N_{m+n}, entonces podríamos modelar la cantidad de éxitos registrados entre los instantes m y $m+n$ como

$$P(N_{m+n} - N_m = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \forall k = 0, 1, \ldots, n$$

Teniendo esto en cuenta, podríamos incluso pensar en las variables aleatorias $N_{n_1}, N_{n_2} - N_{n_1}, \ldots, N_{n_m} - N_{n_{m-1}}$ $(0 < n_1 < n_2 < \cdots < n_m)$ y deducir que se trata de variables aleatorias independientes que gozan de la propiedad Markoviana: el futuro inmediato del proceso sólo depende del presente y no del pasado:

$$P(N_{m+1} = k \,/\, N_0, N_1, \ldots, N_m) = P(N_{m+1} = k \,/ N_m)$$

Proceso número de éxitos
Transiciones

Si se sabe que hasta el instante n-ésimo hubo i éxitos, ¿cuál es la probabilidad de contar j éxitos en el instante siguiente (n+1-ésimo)?

$$P(N_{n+1} = j \ / \ N_n = i) = P(N_n + X_{n+1} = j \ / \ N_n = i) = P(X_{n+1} = j - i)$$

Luego, como $P(X_{n+1} = j - i) = \begin{cases} p, \ si \ j - i = 1 \\ 1 - p, \ si \ j - i = 0 \\ 0, \ en \ otro \ caso \end{cases}$

tenemos entonces que

$$P(N_{n+1} = j \ / \ N_n = i) = \begin{cases} p, \ si \ j - i = 1 \\ 1 - p, \ si \ j - i = 0 \\ 0, \ en \ otro \ caso \end{cases}$$

Notar que estas probabilidades son independientes de n, el momento del proceso en el que nos estamos parando. Recordar además que p se mantiene constante a lo largo de todo el proceso

Proceso número de éxitos

Transiciones

Estas probabilidades pueden interpretarse como la probabilidad de transición en un paso entre dos estados del proceso, y pueden representarse a través de una matriz o un grafo.

$$
T = \begin{array}{cc}
\begin{array}{cccccc} 0 & 1 & 2 & \cdots & n & \cdots \end{array} & \\
\begin{pmatrix}
1-p & p & 0 & \cdots & 0 & \cdots \\
\boxed{0} & 1-p & \boxed{p} & \cdots & 0 & \cdots \\
\vdots & \vdots & \ddots & & \vdots & \\
0 & 0 & \cdots & 1-p & p & \cdots \\
0 & 0 & 0 & \cdots & 1-p & \cdots \\
\vdots & \vdots & \vdots & & \vdots & \ddots
\end{pmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \\ \cdots \\ n \\ \cdots \end{array}
\end{array}
$$

Hablamos de probabilidades de transición de un paso porque estamos en procesos estocásticos a tiempos discretos

$$P\left(N_{j+1} = 0 \;/\; N_j = 1\right) = 0$$

$$P\left(N_{j+1} = 2 \;/\; N_j = 1\right) = p$$

Proceso número de éxitos
Transiciones

Estas probabilidades pueden interpretarse como la probabilidad de transición en un paso entre dos estados del proceso, y pueden representarse a través de una matriz o un grafo.

Los nodos representan los valores que puede tomar el proceso, es decir, los elementos del espacio de estados, mientras que las aristas representan probabilidades de transición positivas (si una probabilidad de transición es nula, entonces no se grafica). Notemos que son aristas dirigidas, ya que indican la probabilidad de pasar de un estado a otro en un paso

Proceso Poisson
Introducción

Ya vimos que el modelo Poisson es una distribución de probabilidades que permite modelar fenómenos de conteo en un tiempo o espacio: llamadas telefónicas a una central, partículas detectadas por un contador radioactivo, personas que llegan a un local, etc.

Un proceso Poisson será una colección de variables aleatorias Poisson indexadas en el tiempo que vamos a usar para modelar la ocurrencia de eventos sobre un intervalo continuo de tiempo. Se trata de un tipo de proceso estocástico con espacio de estados discreto (ya que son conteos, cantidad de veces que se presenta cierto evento) y espacio parametral continuo.

Proceso Poisson
Definición

Dada una serie de ocurrencias del evento en cuestión que ocurren de manera aleatoria a partir de un momento inicial $t = 0$, entonces N_t denota la cantidad de ocurrencias de dicho suceso que ocurrieron al momento t, es decir la cantidad de ocurrencias que tuvieron lugar durante el intervalo $[0, t]$. Dado que N_t contabiliza ocurrencias, para $t_1 < t_2$ entonces tendremos que $N_{t_1} \leq N_{t_2}$ (función no decreciente).

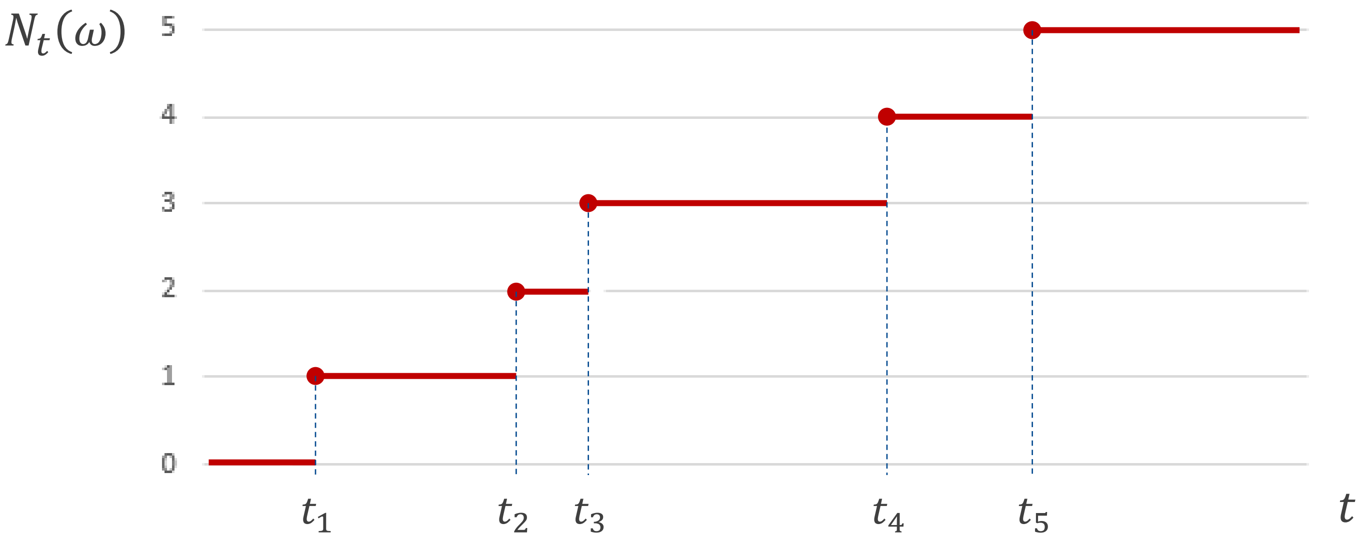
Para un momento t fijo, N_t es una variable aleatoria con distribución Poisson que contabiliza la cantidad de ocurrencias durante el intervalo $[0, t]$:

$$P(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \qquad \forall k \in \mathbb{N}_0, \forall t \geq 0$$

donde λ es la cantidad promedio de ocurrencias del suceso en una unidad de tiempo.

Proceso Poisson
Trayectoria

Para un ω cualquiera podemos graficar la trayectoria del proceso: una función escalonada con saltos de tamaño 1. Esos saltos son los momentos en los que ocurren los eventos:



t_i representa el momento en el que se presenta la i-ésima ocurrencia del evento

Proceso Poisson
Hipótesis de trabajo

- En el instante inicial no se registra ninguna ocurrencia del evento: $N_0 = 0$.
- $t \to N_t$ tiene solo saltos de magnitud 1. No se considera la posibilidad de que se registren dos ocurrencias exactamente en el mismo instante t, siempre puede determinarse cuál de ellas ocurrió primero.
- La cantidad de ocurrencias en un intervalo de longitud t de la forma $[s, s+t]\ \forall s, t \geq 0$ es independiente del inicio del intervalo (s) y solo depende de su longitud (t).
- Para intervalos de tiempo no superpuestos, la cantidad de ocurrencias en un intervalo es independiente de la cantidad de ocurrencias en cualquier otro: sean $0 \leq q < r \leq s < t$, entonces $N_t - N_s$ y $N_r - N_q$ son variables aleatorias independientes.

Proceso Poisson
Ejemplo

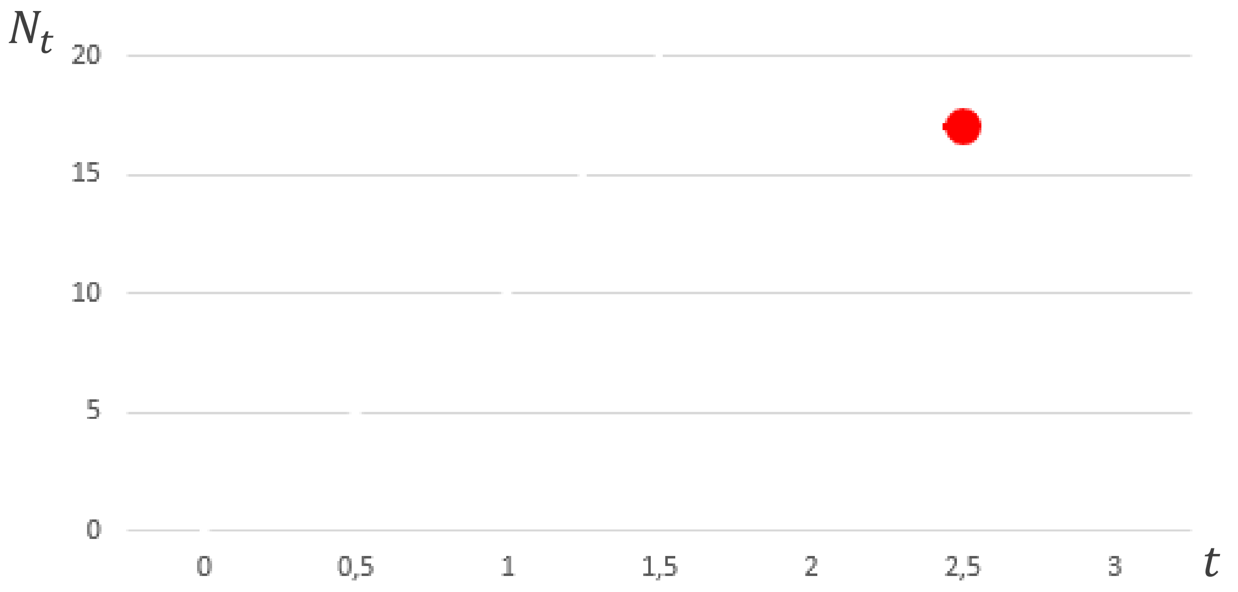
La panadería de Marta abre a las 6am, y se sabe que la cantidad de clientes que entran a su negocio puede modelarse a través de un proceso Poisson con promedio $\lambda = 8\ clientes\ por\ hora$.

¿Cuál es la probabilidad de que hayan entrado 17 clientes hasta las 8:30 am?

$$¿P(N_{2,5} = 17)?$$

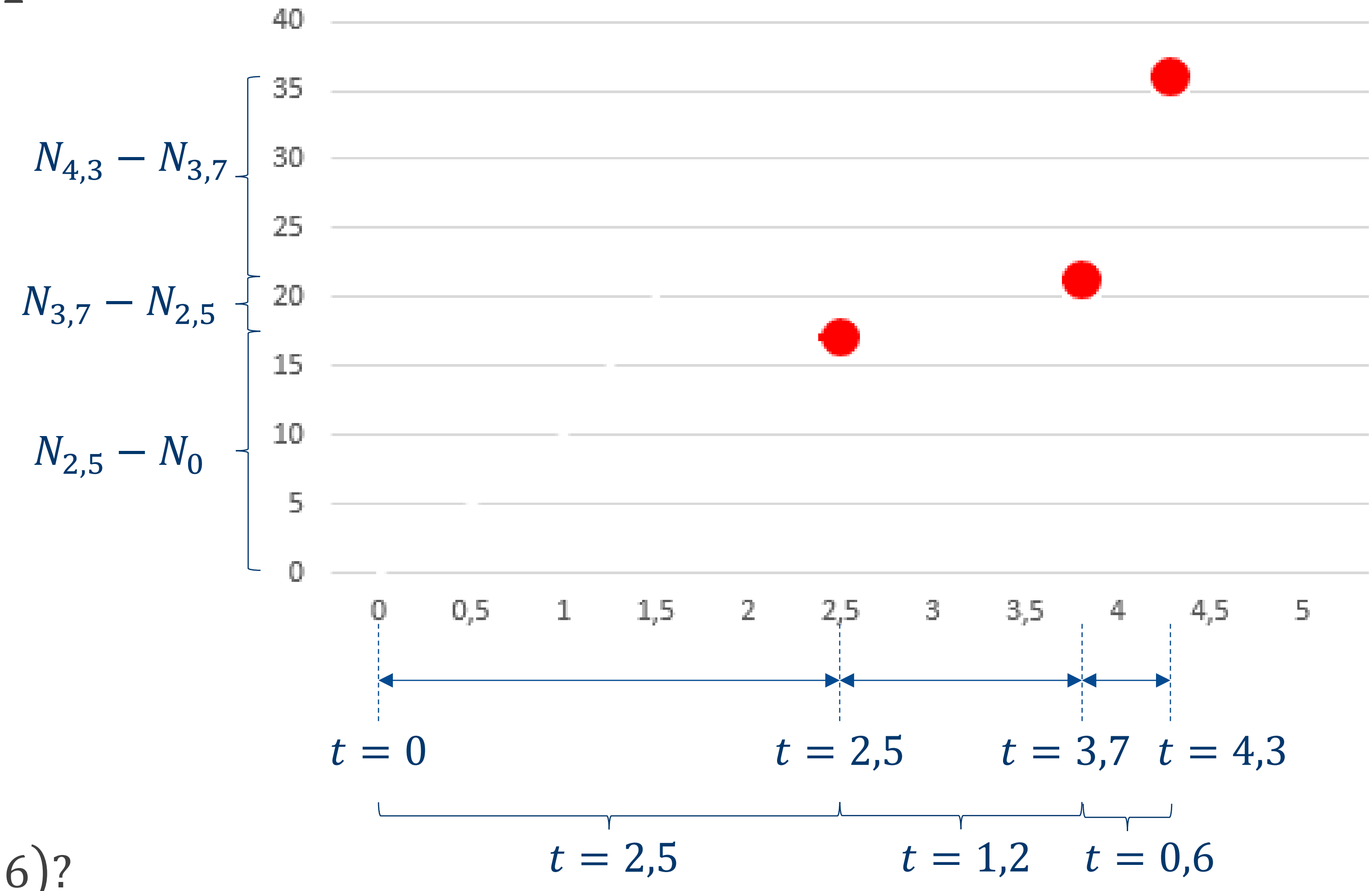
$$P(N_{2,5} = k) = \frac{e^{-8*2,5}(8*2,5)^k}{k!}, \qquad \forall k \in \mathbb{N}_0$$

$$P(N_{2,5} = 17) = \frac{e^{-8*2,5}(8*2,5)^{17}}{17!} = 0,0759$$



Proceso Poisson
Ejemplo

¿Cuál es la probabilidad de que hayan entrado 17 clientes hasta las 8:30 am, 22 hasta las 9:42 am y 36 clientes hasta el horario de cierre (12:18 am)?



$$¿P\left(N_{2,5} = 17, N_{3,7} = 22, N_{4,3} = 36\right)?$$

Proceso Poisson
Ejemplo

¿Cuál es la probabilidad de que hayan entrado 17 clientes hasta las 8:30 am, 22 hasta las 9:42 am y 36 clientes hasta el horario de cierre (12:18 am)?

Dado que en los procesos Poisson las variables aleatorias N_t correspondientes a intervalos no superpuestos son independientes, podemos definir:

$$N_{2,5} \sim Po(8 * 2{,}5), \qquad X_{1,2} \sim Po(8 * 1{,}2), \qquad X_{0,6} \sim Po(8 * 0{,}6)$$

$$P\left(N_{2,5} = 17, N_{3,7} = 22, N_{4,3} = 36\right) = P\left(N_{2,5} = 17\right) * P\left(N_{1,2} = 5\right) * P\left(N_{0,6} = 14\right)$$

Proceso Poisson
Tiempo entre ocurrencias

Cabe destacar que a partir del proceso Poisson emerge otro proceso estocástico: si nombramos con t_k al momento en el que tiene lugar la k-ésima ocurrencia, entonces $\{t_k, k \in \mathbb{N}\}$ representa el proceso estocástico de los momentos de ocurrencia. Tiene espacio de estados $\mathbb{R}^+$ y espacio paramétrico discreto.

A su vez, si pensamos en los tiempos transcurridos entre dos ocurrencias sucesivas del proceso:

$$Y_1 = t_1 - t_0, \qquad Y_2 = t_2 - t_1, \qquad \dots , \qquad Y_k = t_k - t_{k-1}, \qquad \dots \;\; \forall k \in \mathbb{N}$$

entonces resulta que las variables aleatorias Y_k son variables aleatorias independientes, todas idénticamente distribuidas con $Y_k \sim exp(\lambda)$

$$f(y_k) = \begin{cases} \lambda e^{-\lambda * y_k}, & y_k \geq 0 \\ 0, & en\ otro\ caso \end{cases} \qquad\qquad E(Y_k) = \frac{1}{\lambda}$$

Proceso Poisson
Tiempo entre ocurrencias

De este modo, t_k puede escribirse como la suma de k variables aleatorias independientes e idénticamente distribuidas:

$$t_k = (t_1 - t_0) + (t_2 - t_1,) + \cdots + (t_k - t_{k-1}), \qquad \ldots \;\; \forall k \in \mathbb{N}$$

de lo que se desprende que $\mathrm{E}(t_k) = \frac{k}{\lambda}$

Esta es una forma alternativa de definir un proceso Poisson: Si $\{t_k, k \in \mathbb{N}\}$ es un proceso de tiempos de ocurrencia en un proceso $N = \{N_t, t \geq 0\}$ y t_1, $(t_2 - t_1)$, .., $(t_k - t_{k-1})$ son variables aleatorias independientes, idénticamente distribuidas con distribución exponencial de parámetro λ, entonces N es un proceso Poisson de tasa λ.

Proceso Poisson
Ejemplo

Supongamos un equipo que tiene un componente que puede fallar, pero el equipo no debe dejar de funcionar. Por este motivo, se conectan varias de esas componentes de forma que al fallar una, la siguiente la reemplaza automáticamente, y así con todas.

N_t podría definirse como la cantidad de veces que la componente tuvo que ser reemplazada al momento t: $\{N_t, t \geq 0\}$ es un proceso Poisson.

Si $\lambda = 0{,}0002$ y el tiempo está medido en horas, ¿cuál es la vida útil esperada de una componente?

$$Y = t_k - t_{k-1}, \qquad \forall k \in \mathbb{N}$$

$$Y \sim \exp(\lambda = 0{,}0002)$$

$$E(Y) = \frac{1}{\lambda} = \frac{1}{0{,}0002} = 5000\ horas$$

Proceso Poisson
Ejemplo

Si el equipo contiene 3 componentes conectadas y reemplazables ante la falla, ¿cuál es la vida útil del equipo?

$$T = Y_1 + Y_2 + Y_3, \qquad con\ Y_1, Y_2, Y_3\ iid$$
$$Y_i \sim \exp(\lambda = 0{,}0002)$$

$$E(T) = E(Y_1 + Y_2 + Y_3) = E(Y_1) + E(Y_2) + E(Y_3) = \frac{3}{\lambda} = \frac{3}{0{,}0002} = 15.000\ horas$$