

# Introducción a los Procesos Estocásticos

La teoría de los procesos estocásticos se centra en el estudio y modelización de sistemas que evolucionan a lo largo del tiempo, o del espacio, de acuerdo a unas leyes no determinísticas, esto es, de carácter aleatorio.

La forma habitual de describir la evolución del sistema es mediante sucesiones o colecciones de variables aleatorias. De esta manera, se puede estudiar cómo evoluciona una *v.a.* a lo largo del tiempo. Por ejemplo, el número de personas que espera ante una ventanilla de un banco en un instante  $t$  de tiempo; el precio de las acciones de una empresa a lo largo de un año; el número de parados en el sector de Hostelería a lo largo de un año.

La primera idea básica es identificar un proceso estocástico con una sucesión de *v.a.*  $\{X_n, n \in \mathbb{N}\}$  donde el subíndice indica el instante de tiempo (o espacio) correspondiente.

Esta idea inicial se puede generalizar fácilmente, permitiendo que los instantes de tiempo en los que se definen las *v.a.* sean continuos. Así, se podrá hablar de una colección o familia de *v.a.*  $\{X_t, t \in \mathbb{R}\}$ , que da una idea más exacta de lo que es un proceso estocástico.

Se tenía que una *v.a.*  $X(s)$  es una función que va desde un espacio muestral  $S$  a la recta real, de manera que a cada punto  $s \in S$  del espacio muestral se le puede asociar un número de la recta real.

De este modo, la probabilidad de cada suceso de  $S$  se puede trasladar a la probabilidad de que un valor de  $X$  (*v.a.*) caiga en un cierto intervalo o conjunto de números reales. Si a todo esto se le añade una dimensión temporal, se obtiene un proceso estocástico.

La definición formal es la siguiente:

### Definición (Proceso Estocástico)

Dado el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$  de modo que para todo  $t \in T \subset \mathbb{R}$  fijo

$$\begin{aligned} X_t : (\Omega, \mathcal{A}, P) &\longrightarrow (\mathbb{R}, \mathcal{B}) \\ w &\longrightarrow X_t(w) \in \mathbb{R} \end{aligned}$$

esto es,  $X_t$  es una variable aleatoria y  $\forall w \in \Omega$  fijo,  $X_{\bullet}(w)$  es una función del tiempo.

### Ejemplos:

$X_t$ : número de personas que esperan un autobús en un instante  $t$  donde  $t \in [9, 10]$

$X_t$ : precio de una acción de una empresa en un día  $t$  del mes ( $t = 1, 2, \dots, 30$ ).

$X_t$ : número de parados en el mes  $t$  ( $t = 1, 2, \dots, 12$ ).

Para que un proceso estocástico esté completamente definido hay que determinar completamente las *v.a.*, es decir, determinar e identificar la distribución de probabilidad asociada a cada una de ellas y, es más, la distribución conjunta de todas ellas.

### Definición

Al conjunto  $T \subset \mathbb{R}$  de subíndices se le denomina **conjunto paramétrico** y puede ser continuo o numerable.

De este modo aparece una primera clasificación en procesos estocásticos de parámetro continuo o de parámetro discreto.

### Definición

Se denomina **conjunto de estados**  $E$ , al conjunto de los posibles valores que pueden tomar las *v.a.*  $\{X_t\}_{t \in \mathbb{R}}$

*Nota:* En general, se piensa en el subíndice  $t$  como el indicativo del tiempo y en  $X_t$  como el estado o posición del proceso estocástico en el instante  $t$ .

**Ejemplo:** Se lanza una moneda varias veces. Supóngase que cada vez que sale cara, un jugador gana 1 unidad y si sale cruz pierde 1 unidad. Se puede definir un proceso

estocástico que modeliza la evolución del juego. Así, si se denomina  $X_n$  al número de unidades monetarias que quedan después de  $n$  lanzamientos, el espacio muestral de  $X_n$  es

$$\Omega = \{\text{n-uplas de } c \text{ y } +\}$$

de modo que el cardinal (el número de elementos) del conjunto es

$$\#\Omega = 2^n$$

y el álgebra que se define es  $\mathcal{a} = \mathcal{P}(\Omega)$ , esto es, las *partes* de  $\Omega$  (todos los posibles subconjuntos que se pueden formar en  $\Omega$ ).

Consideramos a todas las posibles  $n$ -uplas equiprobables:  $P(w) = 1/2^n$ .

Es un proceso discreto, porque el conjunto paramétrico es  $T = \{1, 2, 3, \dots\}$  y el posible conjunto de estados es

$$E = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

Podemos estudiar la trayectoria de  $X(w)$ :

Sea  $n = 6$  y fijo, por ejemplo,  $w = (c, c, +, +, +, +)$ , esto es,

$$X_1(w) = 1$$

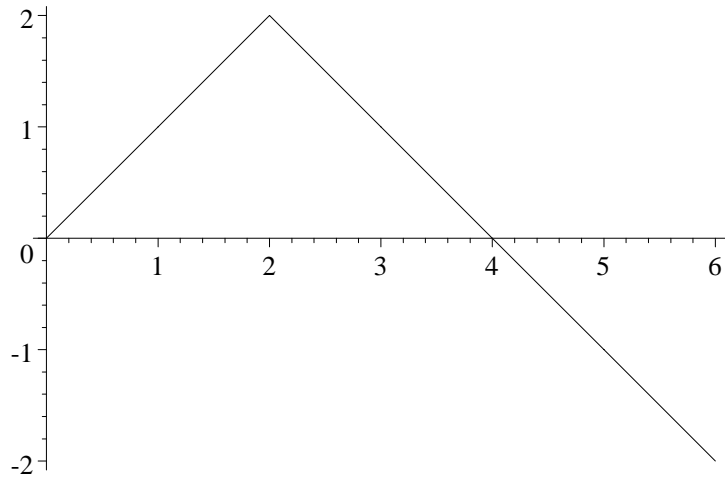
$$X_2(w) = 2$$

$$X_3(w) = 1$$

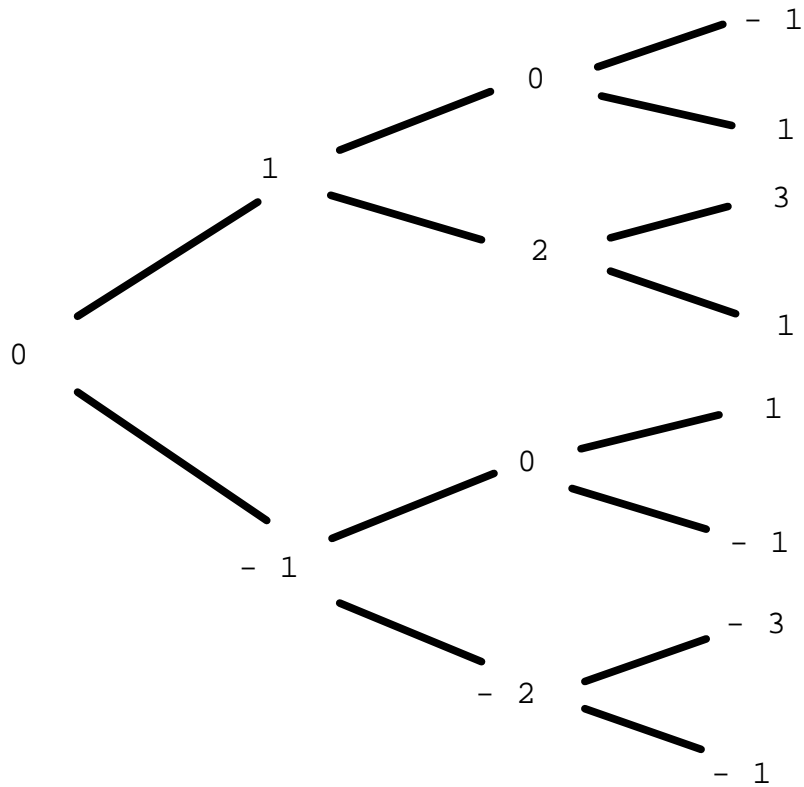
$$X_4(w) = 0$$

$$X_5(w) = -1$$

$$X_6(w) = -2$$



Ahora, si fijamos  $t$ , por ejemplo en  $t = 3$ , se puede calcular la distribución de  $X_3$ . El conjunto de posibles estados de  $X_3$  es:



de modo que  $E = \{-3, -1, 1, 3\}$  y

$$\begin{aligned}P\{X_3 = -3\} &= \frac{1}{2^3} = \frac{1}{8} \\P\{X_3 = -1\} &= 3\frac{1}{2^3} = \frac{3}{8} \\P\{X_3 = 1\} &= 3\frac{1}{2^3} = \frac{3}{8} \\P\{X_3 = 3\} &= \frac{1}{2^3} = \frac{1}{8}\end{aligned}$$

Se puede definir una nueva variable aleatoria:

$Y \equiv$  número de caras obtenidas (éxitos).

Se observa, entonces, que

$$Y \sim Bin\left(3, p = \frac{1}{2}\right)$$

y así

$$\begin{aligned}P\{Y = 0\} &= P\{X_3 = -3\} = \frac{1}{2^3} = \frac{1}{8} \\P\{Y = 1\} &= P\{X_3 = -1\} = 3\frac{1}{2^3} = \frac{3}{8} \\P\{Y = 2\} &= P\{X_3 = 1\} = 3\frac{1}{2^3} = \frac{3}{8} \\P\{Y = 3\} &= P\{X_3 = 3\} = \frac{1}{2^3} = \frac{1}{8}\end{aligned}$$

luego  $X_3$  se distribuye como una  $Bin\left(3, p = \frac{1}{2}\right)$  aunque tomando otros valores que los estrictamente igual a 0, 1, 2, 3.

Se identifica así el proceso estocástico, y se puede preguntar uno cuál es la probabilidad de que a las 10 tiradas se tengan unidades monetarias negativas, esto es, se arruine el jugador.

# Clasificación de Procesos Estocásticos

Se pueden clasificar según

- La estructura del conjunto paramétrico  $T$  y del conjunto de estados  $E$ .
- Las características probabilísticas de las  $v.a.$

## Clasificación según la estructura de $T$ y $E$ .

Los procesos estocásticos se pueden clasificar en cuatro tipos, dependiendo de si  $T$  es un conjunto numerable o continuo, y de si  $E$  es otro conjunto numerable o continuo. Así:

$E \setminus T$	<b>Discreto</b>	<b>Continuo</b>
<b>Discreto</b>	Cadena	Proc. Puntual
<b>Continuo</b>	Suc. de $v.a.$	Proc. Continuo

### Ejemplos:

(i) **Cadena.** Supongamos un sistema hidráulico con una válvula de seguridad que se revisa diariamente. Esta válvula presenta tres posibles estados: *correcto*, *incorrecto* y *deteriorado*. De este modo se puede definir el proceso como  $X_n \equiv$  Estado en el que se encuentra la válvula en el día  $n$ .

(ii) **Proceso Puntual.** Un autobús escolar tiene capacidad para  $k$  personas. Si en una parada hay  $k$  o menos personas esperando, entonces todas subirán al autobús. Si no es así, subirán las  $k$  primeras quedando el resto de personas en espera. Se define, entonces,  $X_t \equiv$  número de personas esperando en la parada en un instante de tiempo  $t$ .

Se pueden plantear asuntos como minimizar el tiempo medio de espera, o reducir el número de personas esperando.

(iii) **Sucesión de v.a.** Una empresa petrolífera debe decidir cuánto petróleo extrae al mes para maximizar los beneficios. Se define, entonces,  $X_n \equiv$  Cantidad de petróleo a extraer en el mes  $n$ .

Esto es un ejemplo particular de la llamada Teoría de Inventarios, donde se estudia la distribución de la demanda, por ejemplo.

(iv) **Proceso Continuo.** En la ventanilla de un banco se contabiliza el tiempo de espera de un cliente que llega en un instante  $t$ . Se puede definir  $X_t \equiv$  Tiempo de espera de un cliente que llega en el instante  $t$ .

Esto está relacionado con la Teoría de Colas. Interesa, por ejemplo, minimizar el tiempo de espera o estudiar la distribución de  $X_t$ .

## Clasificación según las características probabilísticas de las v.a.

En la vida real se producen distintas relaciones entre las variables aleatorias que constituyen un proceso estocástico. Por ejemplo, la ganancia monetaria obtenida en la tirada  $n$ -ésima dependerá obviamente de la ganancia obtenida en la tirada  $(n - 1)$ -ésima.

Las propiedades probabilísticas las *v.a.* son importantes a la hora de identificar y clasificar un proceso estocástico. Se pueden clasificar los procesos en

- Procesos estacionarios.
- Procesos Markovianos.
- Procesos de incrementos independientes.

### 1. Procesos Estacionarios.

Primero es necesario establecer una definición:

**Definición.** (Función de autocovarianzas de un proceso estocástico)

Dado  $\{X_t$  tal que  $t \in T\}$  se llama función de autocovarianzas a la función

$$\gamma(r, s) = Cov(X_r, X_s) = E[(X_r - E(X_r)) \cdot (X_s - E(X_s))]$$

donde  $r, s \in T$ .

Existen dos clases de procesos estacionarios: estacionario débil y estacionario estricto.

**Definición. (Proceso estacionario débil).**

Un proceso  $\{X_t \text{ tal que } t \in T\}$ , tal que  $E(X_t^2) < \infty \forall t \in T$ , es un proceso estacionario débil si

1.  $E(X_t) = m$ , para todo  $t \in T$ .
2.  $\gamma(r, s) = \gamma(r + t, s + t)$ , para todo  $r, s, t \in T$ .

Esto implica, también, que  $Var(X_t)$  es constante para todo  $t \in T$ .

**Observaciones:**

- Debe existir el momento de orden dos de las *v.a.*
- Se observa que todas las *v.a.* tienen la misma media.
- El hecho de que

$$\gamma(r, s) = Cov(X_r, X_s) = Cov(X_{r+t}, X_{s+t})$$

significa que la función de autocovarianzas toma el mismo valor para dos *v.a.* que estén separadas por un retardo  $t$  en el proceso, independientemente de dónde se encuentren situadas estas *v.a.* en el tiempo.

- Si se considera  $t = -s$ , entonces

$$\gamma(r, s) = \gamma(r - s, 0),$$

es decir, es una función de  $(r - s)$ . Esta cantidad es la distancia de separación entre las dos variables  $X_r$  y  $X_s$ . Así, la función de autocovarianzas de un proceso estacionario débil sólo es función de una variable que es la separación  $(r - s)$  entre las variables en consideración.

- Si se toma  $r = s$ , entonces

$$\gamma(r, r) = Cov(X_r, X_r) = Var(X_r) \stackrel{(*)}{=} Var(X_{r+h}), \quad \forall r, h \in T$$

(\*) ya que  $\gamma(r + h, r + h) = Cov(X_{r+h}, X_{r+h}) = \gamma(r, r)$  por (2).



**Definición (Proceso estacionario estricto).**

Un proceso  $\{X_t \text{ tal que } t \in T\}$ , tal que  $E(X_t^2) < \infty \forall t \in T$ , es un proceso estacionario estricto si  $\forall n \in \mathbb{N}, \forall h \in T$  y para todo  $\{t_1, t_2, \dots, t_n\} \in T$  las *v.a.*  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  tienen la misma distribución conjunta que  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ .

P. E. estricto  $\Rightarrow$  P. E. débil, pero no al contrario, en general: P. E. débil  $\not\Rightarrow$  P. E. estricto.

**Observación:** Si  $\{X_t \text{ tal que } t \in T\}$  es un proceso estacionario estricto, tomando  $n = 1$ , se obtiene que todas las *v.a.* tienen la misma distribución y, si se supone la existencia de momentos de orden dos, tienen la misma media, con lo que se cumple la condición (1).

Por otro lado, tomando  $n = 2$ , entonces  $(X_r, X_s)$  se distribuye igual que  $(X_{r+h}, X_{s+h})$ ,  $\forall h \in T$ . De este modo,

$$\gamma(r, s) = \gamma(X_{r+h}, X_{s+h})$$

y sólo depende de cuál sea la diferencia  $(r - s)$ , quedando, así, demostrado (2).

**Observación:** En el caso de la distribución normal, la estacionaridad débil implica estacionaridad estricta dado que, en este caso, la distribución conjunta  $n$ -dimensional queda determinada por las marginales y condicionadas.

Se define un proceso gaussiano como aquel que cumple la propiedad de que  $\forall t_1, \dots, t_n \in T$  la distribución de  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  es normal  $n$ -dimensional.

**2. Procesos markovianos.**

La característica principal de los procesos estocásticos markovianos es que la distribución de  $X_{n+1}$  sólo depende de la distribución de  $X_n$  y no de las anteriores  $(X_{n-1}, X_{n-2}, \dots)$ .

Se puede resumir diciendo que el estado *futuro* del proceso, sólo depende del estado *presente*, y no del resto de estados *pasados*.

Formalmente se expresa como:

$\forall n \in \mathbb{N}$  y  $\forall t_1 < \dots < t_n$

$$P(X_{t_n} \leq x_n | X_{t_1} \leq x_1, \dots, X_{t_{n-1}} \leq x_{n-1}) = P(X_{t_n} \leq x_n | X_{t_{n-1}} \leq x_{n-1}).$$

Cuando el espacio de estados  $E$  es discreto, entonces se puede escribir

$$P(X_{t_n} = x_n | X_{t_1} = x_1, \dots, X_{t_{n-1}} = x_{n-1}) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}).$$

### 3. Procesos de incrementos independientes (ortogonales).

Se dice que un proceso  $\{X_t$  tal que  $t \in T\}$  es de incrementos independientes si  $\forall n \in \mathbb{N}$ ,  $\forall t_1, \dots, t_n \in T$ , con  $t_1 < \dots < t_n$  las *v.a.*

$$\begin{aligned} y_1 &= X_{t_2} - X_{t_1} \\ y_2 &= X_{t_3} - X_{t_2} \\ &\dots \\ y_n &= X_{t_n} - X_{t_{n-1}} \end{aligned}$$

son independientes.

**Proposición** Todo proceso de incrementos ortogonales es un proceso markoviano

*Justificación.*

Suponemos un proceso  $\{X_1, X_2, X_3\}$  y hay que demostrar que

$$P(X_3 = x_3 | X_2 = x_2, X_1 = x_1) = P(X_3 = x_3 | X_2 = x_2).$$

para ver que es markoviano.

Así (dado que se conocen  $X_2 = x_2$  y  $X_1 = x_1$ )

$$\begin{aligned} &P(X_3 = x_3 | X_2 = x_2, X_1 = x_1) = \\ &= P(X_3 - X_2 = x_3 - x_2 | X_2 = x_2, X_2 - X_1 = x_2 - x_1) \stackrel{(*)}{=} \\ &= P(X_3 - X_2 = x_3 - x_2 | X_2 = x_2) = P(X_3 = x_3 | X_2 = x_2). \end{aligned}$$

(\*) por hipótesis, los incrementos son independientes.

Quedando así probado.