

Procesos Estocásticos y Series Temporales

Francisco Javier Mercader Martínez

Índice

1	Introducción a los procesos estocásticos	2
1.1	Concepto de proceso estocástico, trayectorias y ejemplos	2
1.2	Funciones asociadas a un proceso estocástico	4
1.3	Procesos estacionarios y débilmente estacionarios	6
1.4	Procesos gaussianos	7
1.5	Ejemplos y simulación	8
1.5.1	Ruido blanco gaussiano	8
1.5.2	Movimiento Browniano	9
1.5.3	Procesos gaussianos estacionarios isotrópicos	12
2	Cadenas de Markov	18
2.1	Cadenas de Markov de tiempo discreto	18
2.2	Matriz de transición de n pasos	21
2.3	Clasificación de los estados	22

Tema 1: Introducción a los procesos estocásticos

Los procesos estocásticos modelizan cantidades numéricas que cambian con el tiempo de manera aleatoria. Ejemplos del mundo real de estos procesos incluyen:

1. Resultados sucesivos en un juego de azar.
2. Número de plazas ocupadas en un parking.
3. Porcentaje de cielo cubierto en el cielo de Madrid.
4. Evolución del precio de activos financieros: acciones, tipos de cambio de divisas o criptomonedas, materias primas, etc.
5. Indicadores económicos como la inflación, el precio de la luz, y el IBEX 35.

Los procesos estocásticos proporcionan marcos matemáticos para modelar y comprender estos fenómenos, lo que permite hacer predicciones y tomar decisiones informadas.

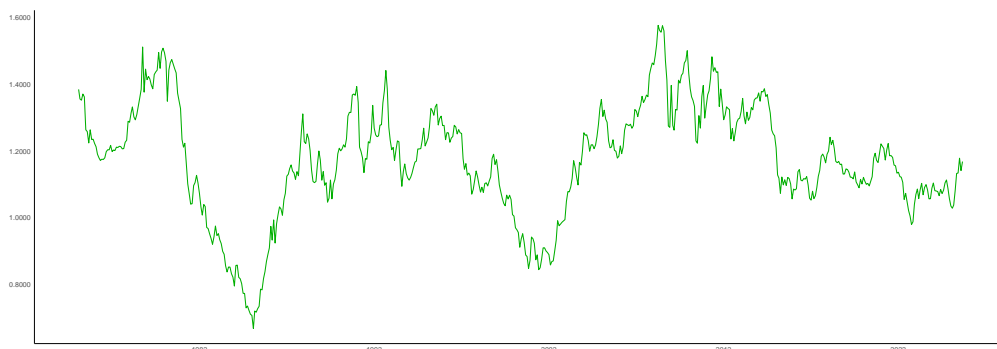


Figure 1: Evolución del tipo de cambio Euro/dólar (Fuente: Google Finance)

1.1) Concepto de proceso estocástico, trayectorias y ejemplos

A lo largo de este tema, vamos a fijar un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y un subconjunto $\mathbb{T} \subset [0, \infty)$.

Definición 1.1.1 Un **proceso estocástico** $(X_t)_{t \in \mathbb{T}}$ es una colección de variables aleatorias reales X_t definidas en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Interpretamos t como el tiempo (medido en cierta unidad).

- Si \mathbb{T} es contable (por ejemplo, $\mathbb{T} = \{0 < 1 < 2 < \dots\}$), diremos que el proceso estocástico $(X_t)_{t \in \mathbb{T}}$ es de tiempo discreto.
- Si \mathbb{T} es un intervalo (por ejemplo, $\mathbb{T} = [0, T]$ o $\mathbb{T} = [0, \infty)$), diremos que el proceso estocástico es de tiempo continuo.

Para cada $t \in \mathbb{T}$ tenemos una variable aleatoria X_t . La variable aleatoria X_t tomará un valor numérico $X_t(\omega)$ para cada $\omega \in \Omega$. A los posibles valores que toma un proceso estocástico se les llama **estados**.

- Si para cada $t \in \mathbb{T}$ la variable aleatoria X_t es de tipo discreto, diremos que el proceso estocástico $(X_t)_{t \in \mathbb{T}}$ es de estado discreto.
- Si para cada $t \in \mathbb{T}$ la variable aleatoria X_t es de tipo continuo, diremos que el proceso estocástico $(X_t)_{t \in \mathbb{T}}$ es de estado continuo.

	t Discreto	t Continuo
X Discreta	Proceso de estado discreto y tiempo discreto (Unidades prod. mensualmente de un producto)	Proceso de estado discreto y tiempo continuo (Unidades producidas hasta t)
X Continua	Proceso de estado continuo y tiempo discreto (Toneladas de producción diaria de un producto)	Proceso de estado continuo y tiempo continuo (Velocidad de un vehículo en el instante t)

Figure 2: Tipos de procesos estocásticos

En lo que sigue, siempre supondremos que $0 \in \mathbb{T}$. Esta condición no es realmente restrictiva, ya que podemos desplazar el tiempo por una constante para garantizar que dicha condición se cumple.

Definición 1.1.2 Dado un proceso estocástico $(X_t)_{t \in \mathbb{T}}$, para cada realización $\omega \in \Omega$, la colección de números reales $(X_t(\omega))_{t \in \mathbb{T}}$ se llama **trayectoria** del proceso.

- Para un proceso estocástico $(X_t)_{t=0,1,2,\dots}$ de tiempo discreto, cada trayectoria define una sucesión de números reales $(x_t)_{t=0,1,2,\dots}$.
- Para un proceso estocástico $(X_t)_{t \in [0,\infty)}$ de tiempo continuo, cada trayectoria define una función real $t \mapsto x(t) : [0, \infty) \rightarrow \mathbb{R}$.

Cada trayectoria corresponde a una observación particular de la evolución del valor de un proceso estocástico en el tiempo.

Ejemplo. Supón que inviertes 100 euros en una cuenta bancaria con tipo de interés anual R , compuesto anualmente. Es decir, si X_t es la cantidad de dinero en el año t , entonces

$$X_t = 100 \cdot (1 + R)^t \text{ para } t = 0, 1, 2, \dots$$

Supongamos también que el valor de $R > 0$ es fijado cuando metes el dinero en la cuenta y sigue una distribución $\text{Exp}(1)$.

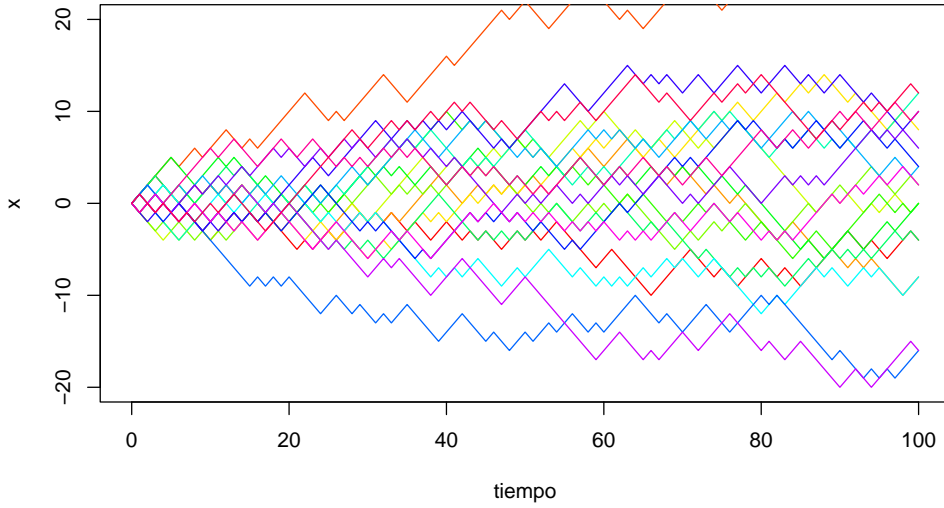
Para cada observación particular $R = r$ de la variable R tendremos una trayectoria del proceso estocástico $(X_t)_{t=0,1,2,\dots}$. Podemos simular y dibujar cinco trayectorias de este proceso estocástico mediante el siguiente código en R:

```
set.seed(145)
nsim <- 20
nlanzamientos <- 100
y <- sample(c(-1,1), size=nlanzamientos, replace = TRUE)
x <- c(0, cumsum(y))
tiempo <- 0:nlanzamientos

colores <- rainbow(nsim)

plot(tiempo, x, col=colores[1], type = "l", lty = 1, ylim=c(-20, 20))

for (i in 2:nsim){
  y <- sample(c(-1,1), size=nlanzamientos, replace = TRUE)
  x <- c(0, cumsum(y))
  lines(tiempo, x, col=colores[i])
}
```



1.2) Funciones asociadas a un proceso estocástico

Dado un proceso estocástico $(X_t)_{t \in \mathbb{T}}$, cada variable aleatoria X_t tendrá su propia distribución de probabilidad la cual puede ser discreta o continua. Observar un proceso estocástico en un solo tiempo de forma aislada no es útil para describir su comportamiento como función del tiempo, ya que los valores observados en un tiempo pueden condicionar su comportamiento en tiempos distintos. Por ejemplo, en el caso del camino aleatorio simétrico en el que apostamos un euro a cara, si hemos acumulado muchas ganancias en el pasado, cual disminuirá a lo sumo en un euro en cada lanzamiento.

Por ese motivo, es necesario estudiar la distribución de probabilidad conjunta a lo largo de varios tiempos.

Definición 1.2.1 Supongamos que $(X_t)_{t \in \mathbb{T}}$ es un proceso estocástico. Dada una sucesión finita de tiempos $\{t_1 < t_2 < \dots < t_n\} \subset \mathbb{T}$, la función $F_{t_1, t_2, \dots, t_n} : \mathbb{R}^n \rightarrow [0, 1]$ definida por

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n),$$

se llama **función de distribución (marginal) finito dimensional** del proceso $(X_t)_{t \in \mathbb{T}}$.

La función de distribución finito dimensional describe el comportamiento probabilístico del proceso estocástico observado en los distintos tiempo t_1, t_2, \dots, t_n . La distribución de probabilidad de un proceso está caracterizada por el conjunto de todas las distribuciones finito dimensionales. En particular, dos procesos estocásticos con las mismas funciones de distribución finito dimensionales tendrán un comportamiento probabilístico similar.

En el caso de que $(X_t)_{t \in \mathbb{T}}$ sea de estado discreto, entonces dicho comportamiento probabilístico puede ser descrito por medio la **función puntual de probabilidad finito dimensional**

$$P_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = \mathbb{P}(X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n).$$

En el caso de que $(X_t)_{t \in \mathbb{T}}$ sea de estado continuo, consideramos la **función de densidad finito dimensional** $f_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n)$, la cual verifica

$$\mathbb{P}(a_1 < X_{t_1} < b_1, \dots, a_n < X_{t_n} < b_n) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) dx_1 \dots dx_n$$

para todo $-\infty \leq a_i < b_i \leq \infty$.

Definición 1.2.2 Dado un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ la **función medio o función de medias** $\mu_X : \mathbb{T} \rightarrow \mathbb{R}$ se define como

$$\mu_X(t) = \mathbb{E}(X_t).$$

- Para un proceso estocástico $(X_t)_{t=0,1,2,\dots}$ de tiempo discreto, la función media define una sucesión de números reales $(\mu_X(t))_{t=0,1,2,\dots}$ de tiempo discreto, la función media define una sucesión de números reales $(\mu_X(t))_{t=0,1,2,\dots}$.
- Para un proceso estocástico $(X_t)_{t \in [0,\infty)}$ de tiempo continuo, la función media define una función real $t \mapsto \mu_X(t) : [0, \infty) \rightarrow \mathbb{R}$.

La función de medias da una idea de cómo el proceso estocástico se comporta en promedio a lo largo del tiempo.

Ejemplo. Volvamos al ejemplo donde una moneda es lanzada reiteras veces y $(X_t)_{t=0,1,2,\dots}$ son las ganancias acumuladas al apostar un euro a cara en cada lanzamiento (paseo aleatorio simétrico). En este caso, $\mu_X(0) = \mathbb{E}(X_0) = \mathbb{E}(0) = 0$ y, si $t > 0$,

$$\mathbb{E}(X_t) = \mathbb{E}(Y_1 + Y_2 + \dots + Y_t) = \mathbb{E}(Y_1) + \mathbb{E}(Y_2) + \dots + \mathbb{E}(Y_t) = 0,$$

donde hemos aplicado que $\mathbb{E}[Y_s] = 0$ para todo s . En definitiva, la ganancia acumulada promedio es 0 para cualquier número de lanzamientos efectuado.

Definición 1.2.3 Dado un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ la **función de covarianza** $C_X : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ se define como

$$C_X(s, t) = \text{Cov}(X_s, X_t) = \mathbb{E}((X_s - \mu_X(s))(X_t - \mu_X(t))) = \mathbb{E}(X_s X_t) - \mu_X(s)\mu_X(t).$$

La **función de varianza** $\sigma_X^2 : \mathbb{T} \rightarrow \mathbb{R}$ se define como

$$\sigma_X^2(t) = \text{Var}(X_t) = C_X(t, t),$$

y la **función de correlación** $\rho_X : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ se define como

$$\rho_X(s, t) = \frac{C_X(s, t)}{\sigma_X(s)\sigma_X(t)}.$$

Ejemplo. Consideremos el proceso estocástico $(X_t)_{t \in [0,\infty)}$ definido como

$$X_t = A + Bt \quad \forall t \in [0, \infty),$$

donde A y B son variables aleatorias $N(1, 1)$ independientes.

Tenemos por un lado que la función de medias viene dada por:

$$\mu_X(t) = \mathbb{E}(X_t) = \mathbb{E}(A + Bt) = 1 + t \quad \forall t \in [0, \infty).$$

Y por otro lado la función de varianzas es:

$$C_X(s, t) = \mathbb{E}(X_s X_t) - \mu_X(s)\mu_X(t)$$

Observar que la esperanza del producto vale:

$$\begin{aligned} \mathbb{E}(X_s X_t) &= \mathbb{E}((A + Bt)(A + Bs)) \\ &= \mathbb{E}(A^2 + ABs + ABt + B^2 st) \\ &= \mathbb{E}(A^2) + \mathbb{E}(A)\mathbb{E}(B)(s + t) + \mathbb{E}(B^2)st \\ &= 2 + s + t + 2st \end{aligned}$$

Y el producto de esperanza vale:

$$\mu_X(s)\mu_X(t) = (1 + s)(1 + t) = 1 + s + t + st.$$

Por tanto, para este proceso estocástico la función de covarianzas es:

$$C_X(s, t) = (2 + s + t + 2st) - (1 + s + t + st) = 1 + st,$$

la función de varianzas:

$$\sigma_X^2(t) = \text{Var}(X_t) = C_X(t, t) = 1 + t^2,$$

y la función de correlaciones:

$$\rho_X(s, t) = \frac{1 + st}{\sqrt{(1 + t^2)(1 + s^2)}}$$

Ejemplo. Volvamos al ejemplo del paseo aleatorio simétrico $(X_t)_{t=0,1,2,\dots}$. Fijados dos números naturales n, m con $n \leq m$, tenemos

$$\begin{aligned} C_X(n, m) &= \mathbb{E}(X_n X_m) \\ &= \mathbb{E}(X_n(X_m - X_n + X_n)) \\ &= \mathbb{E}(X_n(X_m - X_n)) + \mathbb{E}(X_n^2) \\ &= \mathbb{E}(X_n)\mathbb{E}(X_m - X_n) + \mathbb{E}(X_n^2) \\ &= n = \min(n, m), \end{aligned}$$

donde hemos usado que X_n y $X_m - X_n$ son variables aleatorias independientes. Además, obsérvese que, como $X_n = Y_1 + Y_2 + \dots + Y_n$ y las variables Y_1, Y_2, \dots, Y_n son independientes, se tiene que:

$$\text{Var}(X_n) = \text{Var}(Y_1^2) + \text{Var}(Y_2^2) + \dots + \text{Var}(Y_n^2) = 1 + 1 + \dots + 1 = n,$$

y por tanto:

$$\mathbb{E}(X_n^2) = \text{Var}(X_n) + (\mathbb{E}(X_n))^2 = n + 0 = n.$$

1.3) Procesos estacionarios y débilmente estacionarios

Los procesos estocásticos se pueden dividir entre estacionarios (en sentido estricto), débilmente estacionarios y no estacionarios. Intuitivamente hablando, un proceso estocástico es estacionario (en sentido estricto) si su comportamiento probabilístico o distribución no cambia con el tiempo.

Definición 1.3.1 Un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ se dice que es **estacionario** si para toda sucesión finita de tiempos t_1, t_2, \dots, t_n y todo $s > 0$ se verifica que los vectores aleatorios

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \quad \text{y} \quad (X_{t_1+s}, X_{t_2+s}, \dots, X_{t_n+s})$$

tienen la misma función de distribución, es decir,

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = F_{t_1+s, t_2+s, \dots, t_n+s}(x_1, x_2, \dots, x_n)$$

para cualesquiera números reales x_1, x_2, \dots, x_n .

En particular, si un proceso es estacionario, entonces todas las variables aleatorias X_t del proceso tienen la misma distribución, es decir, están idénticamente distribuidas. Esto no implica que los vectores aleatorios que pueden formar con las variables del proceso en diferentes instantes tengan todos la misma distribución.

En la práctica es muy útil saber si un proceso estocástico es estacionario cuando es necesario predecir el comportamiento futuro de dicho proceso. Si sabemos que el proceso es estacionario, entonces observando su comportamiento en el pasado podremos inferir información de su comportamiento en el futuro.

A continuación introducimos una noción de estacionaridad menos exigente, la cual es también muy útil para predecir comportamientos futuros de procesos estocásticos reales.

Definición 1.3.2 Un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ se dice que es **débilmente estacionario** si:

- 1) $\mu_X(t) = \mu_X(0)$ para todo $t \in \mathbb{T}$. Es decir, la función de medias es constante.
- 2) $C_X(s, t) = C_X(0, t - s)$ para todo $s, t \in \mathbb{T}$ con $s \leq t$. Es decir, la función de covarianzas depende sólo del salto entre tiempos.

Obsérvese que, en particular, si el proceso es débilmente estacionario tendrá función de varianzas constante.

Dado que la media $\mu_X(t)$ es determinada por la función de distribución marginal $F_t(x)$, y la covarianza $C_X(s, t)$ es determinada por la función de distribución marginal $F_{s,t}(x)$, todo proceso estacionario será también débilmente estacionario. Sin embargo, es posible que un proceso sea débilmente estacionario pero no estacionario.

Ejemplo. Consideremos el proceso estocástico $(X_t)_{t \in [0, \infty)}$ definido por

$$X_t = Y \cos(U + t),$$

donde Y y U son variables aleatorias independientes con $Y \sim N(0, 1)$ y $U \sim U[0, 2\pi]$. Entonces

$$\mu_X(t) = 0,$$

y además, para $s \leq t$,

$$\begin{aligned} C_X(s, t) &= \mathbb{E}(X_s X_t) - \mu_X(s) \mu_X(t) \\ &= \mathbb{E}((Y \cos(U + s))(Y \cos(U + t))) \\ &= \mathbb{E}(Y^2) \mathbb{E}(\cos(U + s) \cos(U + t)) \\ &= \mathbb{E}(\cos(U + s) \cos(U + t)) \\ &= \mathbb{E}\left(\frac{1}{2} \cos(2U + s + t) + \frac{1}{2} \cos(t - s)\right) \\ &= \mathbb{E}\left(\frac{1}{2} \cos(2U + s + t)\right) + \frac{1}{2} \cos(t - s) \\ &= \frac{1}{2} \int_0^{2\pi} \left(\cos(2u + s + t) \frac{1}{2\pi}\right) du + \frac{1}{2} \cos(t - s) \\ &= 0 + \frac{1}{2} \cos(t - s) \\ &= \frac{1}{2} \cos(t - s) \end{aligned}$$

En particular, tenemos que $C_X(s, t) = \frac{1}{2} \cos(t - s) = C_X(0, t - s)$, de donde deducimos que el proceso estocástico $(X_t)_{t \in [0, \infty)}$ es débilmente estacionario.

1.4) Procesos gaussianos

A continuación introducimos una de las familias de procesos estocásticos más importantes: los procesos gaussianos. Este tipo de procesos tienen importantes aplicaciones en Machine Learning.

Recordemos primero la noción de variable **Normal multivariante**. Una propiedad importante de las variables normales multivariadas es que su función de densidad (y, por tanto, su distribución) está completamente determinada por el vector de medias y la matriz de covarianzas.

Definición 1.4.1 Se dice que el vector aleatorio $\vec{X} = (X_1, X_2, \dots, X_n)$ sigue una distribución Normal multivariante si su función de densidad viene dada por:

$$f_X(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|C_X|}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_X)^\top C_X^{-1} (\vec{x} - \vec{\mu}_X) \right\} \text{ para todo } \vec{x} \in \mathbb{R}^n,$$

donde $\vec{\mu}_X$ es el vector de medias de \vec{X} y C_X es la matriz de covarianzas de \vec{X} .

Propiedad: Se puede probar, aunque no es inmediato, que una condición equivalente para el vector aleatorio (X_1, X_2, \dots, X_n) siga una distribución Normal multivariante es que, para cualesquiera números reales a_1, a_2, \dots, a_n la variable aleatoria real

$$a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

sigue una distribución Normal univariante.

Otra importante propiedad de las variables aleatorias normales multivariadas que conviene recordar es que las transformaciones lineales de estas variables también son normales multivariadas. En otras palabras, si \vec{X} es un vector aleatorio Normal multivariante de tamaño n , A es una matriz constante de tamaño $m \times n$, y \vec{b} es un vector constante de tamaño m , entonces el vector aleatorio $\vec{Y} = A\vec{X} + \vec{b}$ también es un vector aleatorio Normal multivariante.

Definición 1.4.2 Un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ se dice que es **gaussiano** (o **Normal**) si para toda sucesión de tiempos $t_1, t_2, \dots, t_n \in \mathbb{T}$ el vector aleatorio

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n})$$

es una variable Normal multivariante.

Una importante propiedad de los procesos gaussianos es que las nociones de estacionaridad y estacionaridad débil son equivalentes para este tipo de procesos. En la práctica esto significa que para verificar que un proceso gaussiano es estacionario basta analizar su media y covarianza. Concretamente, tenemos lo siguiente.

Teorema 1.4.1 Sea $(X_t)_{t \in \mathbb{T}}$ un proceso estocástico gaussiano. Si $(X_t)_{t \in \mathbb{T}}$ es débilmente estacionario, entonces es también estacionario.

1.5) Ejemplos y simulación

En lo que sigue vamos a estudiar y simular en R algunos ejemplos particulares de procesos estocásticos gaussianos, analizando en cada paso si se trata de procesos estacionarios o no estacionarios.

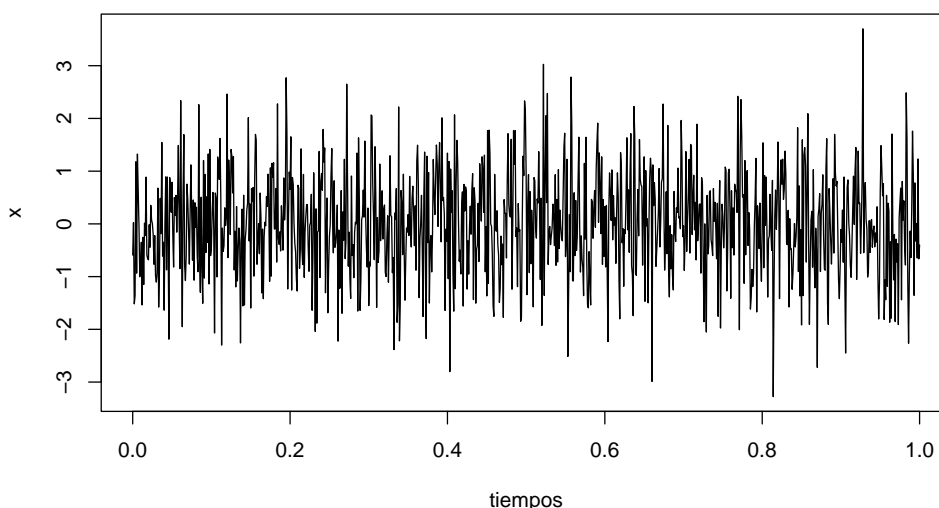
1.5.1) Ruido blanco gaussiano

Definición 1.5.1 Un proceso estocástico $(X_t)_{t \in \mathbb{T}}$ se llama **ruido blanco gaussiano** si las variables aleatorias X_t son independientes y siguen una distribución $N(0, \sigma^2)$.

Claramente el ruido blanco gaussiano es un proceso gaussiano. Además, $\mu_X(t) = 0$, $C_X(t, s) = 0$ si $t \neq s$, y $C_X(t, t) = \sigma^2$. Lo cual en particular implica que el ruido blanco gaussiano es un proceso estacionario.

Podemos simular en R una trayectoria del ruido blanco gaussiano:

```
set.seed(11)
sigma <- 1
tiempos = seq(from = 0, to = 1, by = 0.001)
x=rnorm(length(tiempos), 0, sigma)
plot(tiempos, x, type = "l", lty = 1)
```



Obtenemos una trayectoria extremadamente irregular. Además, observamos a simple vista un comportamiento estacionario del proceso.

1.5.2) Movimiento Browniano

Definición 1.5.2 Un proceso estocástico $(X_t)_{t \in [0, \infty)}$, de tiempo continuo se dice que es un **movimiento Browniano** o proceso de Wiener si cumple:

- 1) $X_0 = 0$.
- 2) Para todo par de tiempos $s \leq t$, la variable aleatoria $X_t - X_s$, sigue una distribución $N(0, t - s)$.
- 3) Para cualquier sucesión de tiempo $t_1 < t_2 < \dots < t_n$ se tiene que las variables aleatorias

$$X_{t_2} - X_{t_1}, \quad X_{t_3} - X_{t_2}, \quad \dots, \quad X_{t_n} - X_{t_{n-1}}$$

son independientes.

- 4) Las trayectorias $t \mapsto X_t(\omega)$ son funciones continuas.

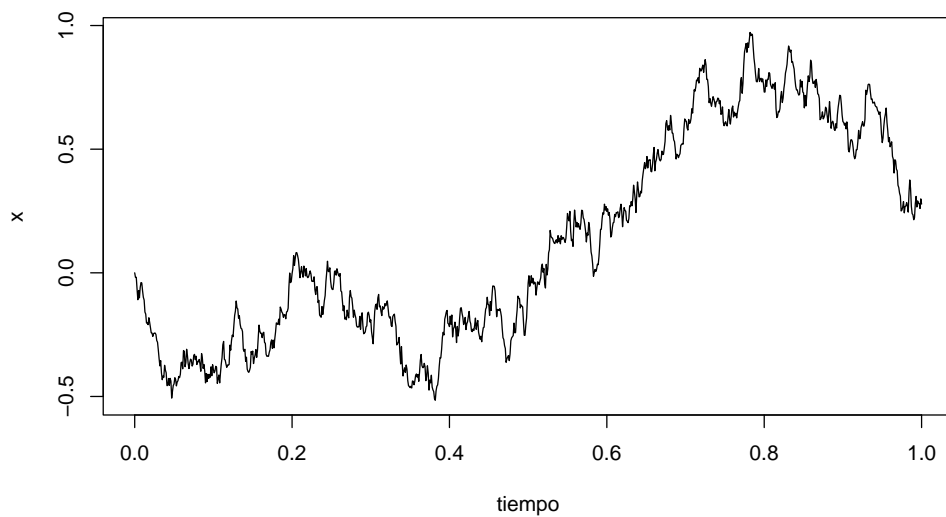
El movimiento Browniano se utiliza en finanzas para modelar la evolución del precio de activos financieros con las acciones o los tipos de cambio. En particular, se aplica a la valoración de opciones financieras mediante el modelo de Black-Scholes, y el análisis de riesgos.

A continuación vamos a simular una trayectorias del movimiento Browniano. Para ello fijamos un incremento de tiempo $\Delta t > 0$ y un número natural n . Deseamos simular una trayectoria en los instantes de tiempo $t_0 = 0, t_1 = \Delta t, \dots, t_n = n\Delta t$. En vista de la definición de movimiento Browniano, tenemos que los incrementos

$$\Delta X_{t_1} = X_{t_1} - X_{t_0}, \quad \Delta X_{t_2} = X_{t_2} - X_{t_1}, \quad \dots, \quad \Delta X_{t_n} = X_{t_n} - X_{t_{n-1}},$$

son variables aleatorias independientes y siguen todas ellas una distribución $N(0, \Delta t)$. De este modo, nuestra estrategia será simular los valores de los incrementos mediante la generación de una muestra aleatoria de n variables $N(0, \Delta t)$ independientes, y reconstruir la trayectoria del proceso a partir de dichos incrementos.

```
set.seed(11)
n <- 1000
dt <- 1/n
tiempo <- seq(from=0, to=1, by=dt)
incrementos <- rnorm(n, 0, sqrt(dt))
x <- c(0, cumsum(incrementos))
plot(tiempo, x, type="l", lty=1)
```



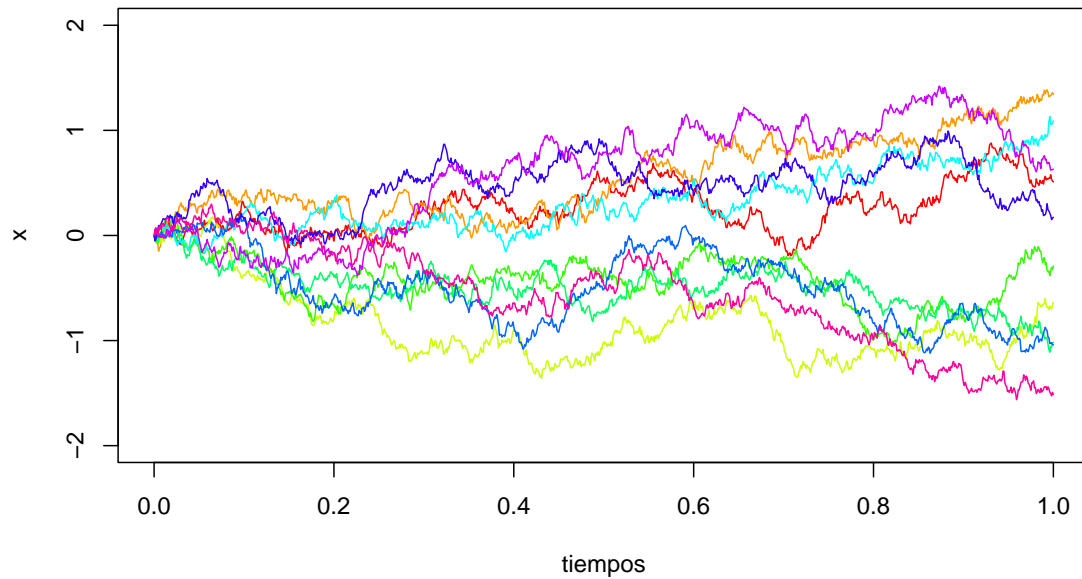
Para analizar el comportamiento probabilístico, simularemos diez trayectorias del proceso:

```
set.seed(123)
nsim <- 10
n <- 1000
dt <- 1/n
tiempos <- seq(from=0, to=1, by=dt)
incrementos <- rnorm(n, 0, sqrt(dt))
x <- c(0, cumsum(incrementos))

colores <- rainbow(nsim)

plot(tiempos, x, ylim=c(-2,2), type = "l", lty = 1, col=colores[1])

for (i in 2:nsim) {
  incrementos=rnorm(n, 0, sqrt(dt))
  x <- c(0, cumsum(incrementos))
  lines(tiempos, x, col=colores[i])
}
```



A simple vista observamos un comportamiento no estacionario del proceso. Notamos que a medida que avanza el tiempo, las trayectorias tienden a volverse más dispuestas, lo que indica que el comportamiento probabilístico de las trayectorias varía con el tiempo.

El movimiento Browniano es un proceso gaussiano. De hecho, la distribución de probabilidad del vector aleatorio $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$, para $t_1 < t_2 < \dots < t_n$, es normal multivariada porque se obtiene como una transformación lineal del vector $(X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}})$, el cual es normal multivariado porque sus componentes son independientes y normales.

Además, $\mu_X(t) = \mathbb{E}(X_t) = 0$ y, para $s \leq t$,

$$\begin{aligned} C_X(s, t) &= \mathbb{E}(X_s X_t) \\ &= \mathbb{E}(X_s(X_t - X_s + X_s)) \\ &= \mathbb{E}(X_s(X_t - X_s)) + \mathbb{E}(X_s^2) \\ &= s = \min(s, t). \end{aligned}$$

En resumen, el movimiento Browniano es un proceso gaussiano con funciones de media y covarianza $\mu_X(t) = 0$, y $C_X(s, t) = \min(s, t)$, respectivamente. En particular, observamos que este proceso no es estacionario (ni siquiera en sentido débil), ya que en general no se verifica que $C_X(s, t) = C_X(0, t - s)$. Por ejemplo, $C_X(3, 4) = \min(3, 4) = 3$; sin embargo, $C_X(0, 4 - 3) = \min(0, 1) = 0$.

A continuación vamos a detallar otra estrategia para simular el movimiento Browniano, la cual se basa en que dicho proceso es gaussiano. Aunque nos vamos a centrar en el caso particular del movimiento Browniano, este procedimiento sirve para simular cualquier proceso gaussiano conocidas sus funciones de media y covarianza. Fijados los tiempo $t_1 < \dots < t_n$ sabemos que el vector aleatorio $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ sigue una distribución multivariada de media $\mu_X = 0$ y matriz de covarianzas $C_X(c_{i,j})_{n \times n}$ donde $c_{i,j} = \min\{t_i, t_j\}$. De esta manera, podremos simular una trayectoria mediante la obtención de una muestra aleatoria del anterior vector aleatorio. Para ello utilizamos la librería `mvtnorm` de R, la cual permite simular variables aleatorias normales multivariadas:

```
set.seed(23)
n <- 1000
dt <- 1/n
tiempo = seq(from=0, to=1, by=dt)

# Definimos la matriz de covarianzas
```

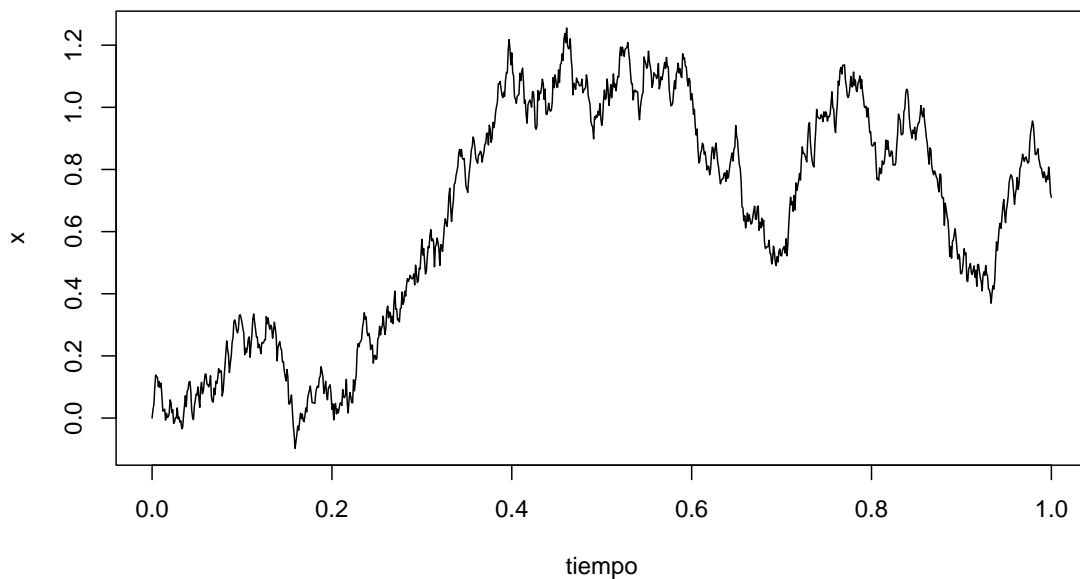
```

C <- matrix(NA, nrow = n, ncol = n)
C <- dt*pmin(row(C), col(C))

# Simulamos una normal multivariada
library(mvtnorm)
x <- rmvnorm(1, sigma=C)

# Añadimos el valor 0 inicial
x <- c(0, x)
plot(tiempo, x, type = "l", lty = 1)

```



1.5.3) Procesos gaussianos estacionarios isotrópicos

Un tipo particular de proceso gaussiano estacionario son los procesos gaussianos estacionarios isotrópicos.

Definición 1.5.3 Decimos que un proceso gaussiano estacionario $(X_t)_{t \in \mathbb{T}}$ es **isotrópico** si su función de medias es constante, $\mu_X(t) = \mu$, y la función de covarianza depende sólo de la distancia $|t - s|$, es decir,

$$C_S(s, t) = K(|t - s|)$$

para cierta función $K : [0, \infty) \rightarrow [0, \infty)$ llamada núcleo.

Tomemos por ejemplo $\mu_X(t) = 0$ y el núcleo exponencial cuadrático $K_l(x) = \exp\left(-\frac{x^2}{2l^2}\right)$, donde l es un parámetro de escala. Vamos a simular en R una trayectoria de dicho proceso:

```

set.seed(112)
n <- 1000
dt <- 1/n
tiempo <- seq(from=0, to=1, by=dt)

# Definimos la matriz de covarianzas
l <- 1
C <- matrix(NA, nrow = n+1, ncol = n+1)

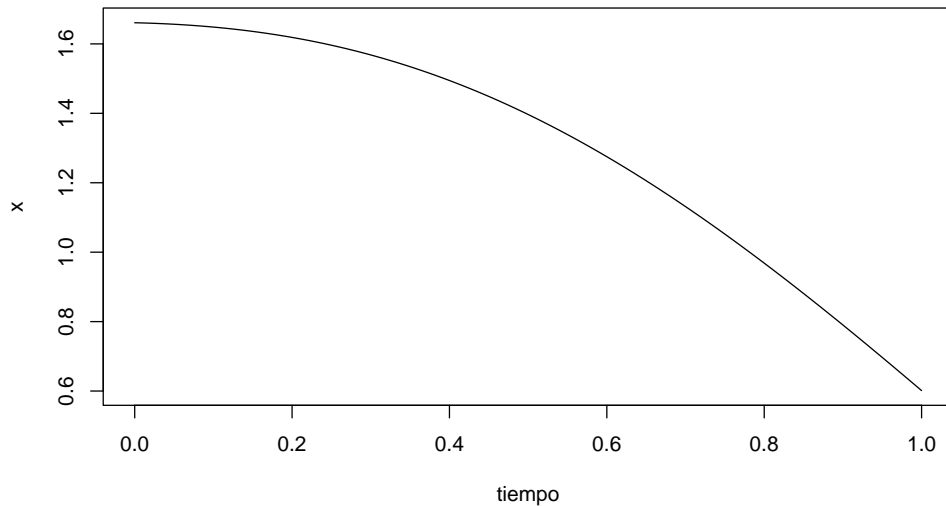
```

```

C <- dt*abs(row(C)-col(C))
C <- exp(-C^2/(2*1^2))

# Simulamos una normal multivariada
library(mvtnorm)
x <- rmvnorm(1, sigma = C)
plot(tiempo, x, type = "l", lty = 1)

```



Para analizar el comportamiento probabilístico, podemos simular diez trayectorias del proceso:

```

set.seed(112)
nsim <- 10
n <- 1000
dt <- 1/n
tiempo <- seq(from=0, to=1, by=dt)

# Definimos la matriz de covarianzas
l <- 1
C <- matrix(NA, nrow = n+1, ncol = n+1)
C <- dt*abs(row(C)-col(C))
C <- exp(-C^2/(2*1^2))

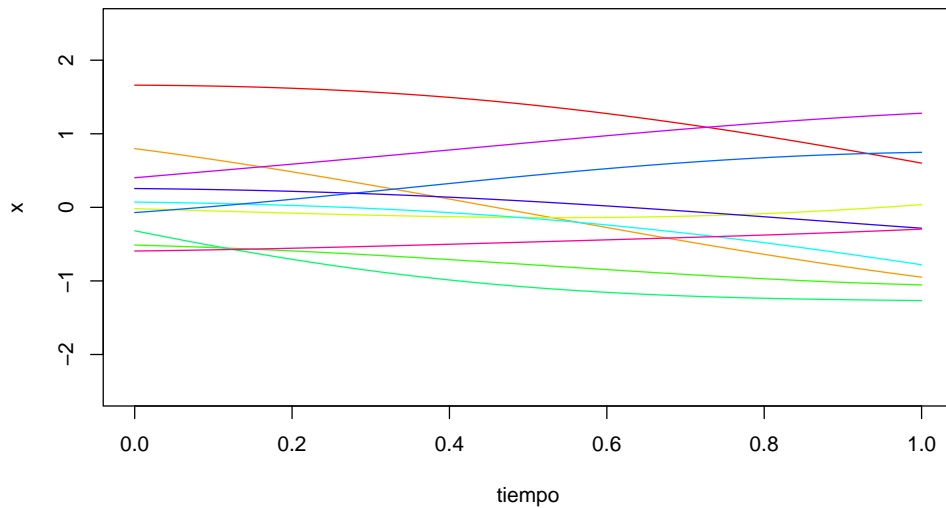
# Simulamos una normal multivariada
library(mvtnorm)
x <- rmvnorm(1, sigma = C)

colores <- rainbow(10)

plot(tiempo, x, type = "l", lty = 1, ylim = c(-2.5,2.5), col=colores[1])

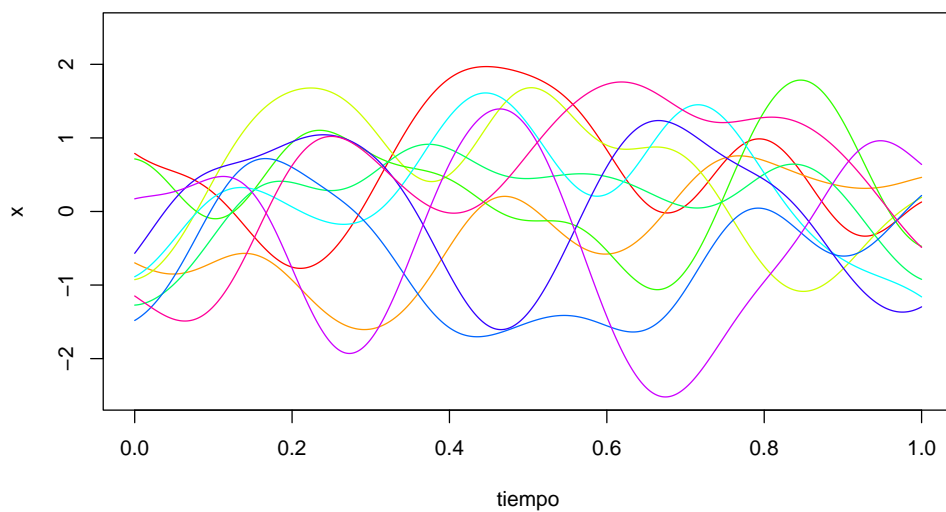
for (i in 2:nsim) {
  x <- rmvnorm(1, sigma = C)
  lines(tiempo, x, col=colores[i])
}

```



A diferencia del movimiento Browniano, vemos que las trayectorias presentan un comportamiento estacionario, ya que su distribución es la misma a lo largo del tiempo. Otra diferencia con el movimiento Browniano es que las trayectorias tienen una apariencia suave.

Podemos cambiar el parámetro de escala, tomando $l = 0.1$. De este modo obtenemos las trayectorias:



Tomemos ahora el núcleo de Ornstein-Uhlenbeck $K_l(x) = \exp\left(-\frac{|x|}{l}\right)$

```
set.seed(114)
nsim <- 10
n <- 1000
dt <- 1/n
tiempo <- seq(from=0, to=1, by=dt)

# Definimos la matriz de covarianzas
l <- 1
C <- matrix(NA, nrow = n+1, ncol = n+1)
C <- dt*abs(row(C)-col(C))
C <- exp(-C/l)
```

```

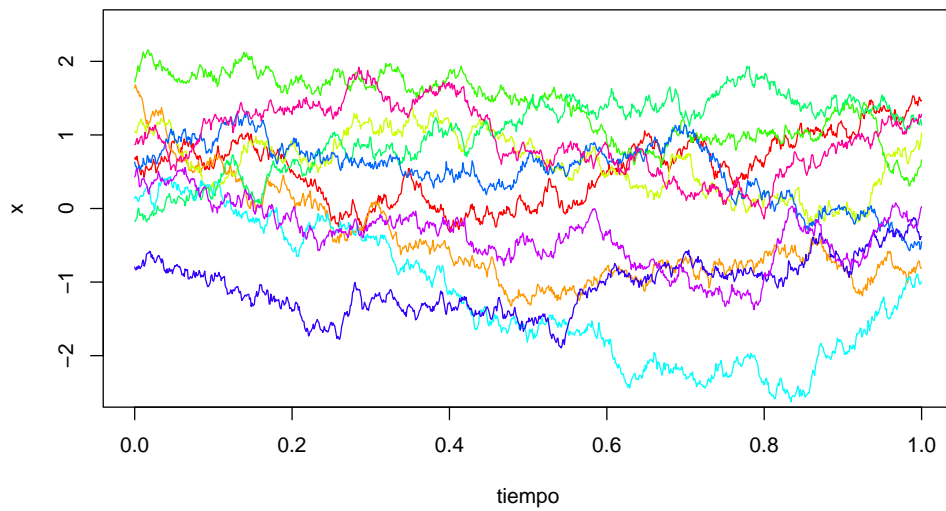
# Simulamos una normal multivariada
library(mvtnorm)
x <- rmvnorm(1, sigma = C)

colores <- rainbow(10)

plot(tiempo, x, type = "l", lty = 1, ylim = c(-2.5,2.5), col=colores[1])

for (i in 2:nsim) {
  x <- rmvnorm(1, sigma = C)
  lines(tiempo, x, col=colores[i])
}

```



Vemos que las trayectorias tienen una apariencia más irregular la cual nos recuerda al movimiento Browniano. Sin embargo, en este caso tenemos un proceso estacionario cuyas trayectorias se comportan de manera similar a lo largo del tiempo.

Hoja 1: Problemas de Introducción a los Procesos Estocásticos

- 1) ¿Qué es un proceso estocástico y cómo se define formalmente?
- 2) ¿Cuáles son los tipos principales de procesos estocásticos en función del tiempo?
- 3) ¿Cómo se clasifica un proceso estocástico según los estados que puede formar?
- 4) ¿Qué es una trayectoria en el contexto de un proceso estocástico?
- 5) ¿Qué es un paseo aleatorio y cómo se define en términos de variables aleatorias? Simular, usando R, 4 trayectorias de un paseo aleatorio con variables aleatorias i.i.d. dadas por la distribución $N(0, 1)$.
- 6) ¿Qué características tiene un paseo aleatorio simple y en qué se diferencia del paseo aleatorio simétrico? Simular, usando R, 6 trayectorias de un paseo aleatorio simple $p = \frac{1}{3}$.
- 7) ¿Qué es una función de distribución finito dimensional en un proceso estocástico?
- 8) ¿Cómo se define la función de medias de un proceso estocástico? ¿Y la de covarianzas y correlaciones?
- 9) ¿Qué relación existe entre la función de varianza y la función de covarianza en un proceso estocástico?
- 10) ¿Qué condiciones deben cumplirse para que un proceso sea estacionario y en qué se diferencia de la estacionariedad débil?
- 11) ¿Qué diferencias hay entre un proceso estocástico gaussiano y un ruido blanco gaussiano? Simular, usando R, 2 trayectorias de un ruido blanco gaussiano con varianza σ^2 .
- 12) ¿Cuál es la definición de un proceso de Wiener o movimiento Browniano? Simular, usando R, 3 trayectorias de un movimiento Browniano.
- 13) ¿Cuál de las siguientes es una condición necesaria para que un proceso estocástico sea débilmente estacionario?
 - (a) La varianza depende del tiempo.
 - (b) La media es constante en el tiempo.
 - (c) La covarianza es cero para todos los tiempos.
- 14) Si $(X_t)_{t \in [0, \infty)}$ es ruido blanco (gaussiano) con varianza σ^2 , ¿cuál de las siguientes afirmaciones es verdadera?
 - (a) $\text{Cov}(X_t, X_{t+h}) = \sigma^2$ para todo t y h .
 - (b) $\text{Cov}(X_t, X_{t+h}) = \sigma^2$ si $h = 0$, $\text{Cov}(X_t, X_{t+h}) = 0$ si $h \neq 0$.
 - (c) $\text{Cov}(X_t, X_{t+h}) = \frac{1}{2}\sigma^2$ para todo t y h .
- 15) ¿Cuál de las siguientes afirmaciones es verdadera para un ruido blanco?
 - (a) Tiene una función de correlación que decrece exponencialmente.
 - (b) Su media es no nula.
 - (c) Sus valores en diferentes tiempos son independientes.
- 16) ¿Qué describe la función de covarianza de un proceso estocástico?
 - (a) La relación lineal en el proceso en dos tiempos diferentes.
 - (b) La suma de todas las realizaciones del proceso.
 - (c) La frecuencia con la que el proceso cruza la media.
- 17) Una condición necesaria para que un proceso estocástico sea débilmente estacionario es que:
 - (a) La media debe ser cero.
 - (b) La covarianza depende del tiempo.

(c) La función de covarianza depende solo de la diferencia entre los tiempos.

18) Consideremos el proceso estocástico de tiempo discreto $(X_t)_{t=0,1,2,\dots}$, con $X_0 = a \in \mathbb{R}$ una constante cualquiera y $X_t = 3 + \varepsilon_t + 2\varepsilon_{t-1}$ para todo $t = 1, 2, \dots$, siendo $(\varepsilon_t)_{t=0,1,2,\dots}$ ruido blanco (gaussiano) de varianza σ^2 .

(a) Calcular la función de medias, covarianzas y correlaciones del proceso $(X_t)_{t=0,1,2,\dots}$.

(b) Justificar si el proceso es estacionario en sentido débil.

(c) Justificar si el proceso es estacionario (en sentido fuerte).

Nota: El proceso del enunciado es un proceso denominado MA(1) que estudiará en el tema posterior.

19) Consideremos el proceso estocástico de tiempo discreto $(X_t)_{t=0,1,2,\dots}$, con $X_0 = a \in \mathbb{R}$ una constante cualquiera y $X_t = \alpha X_{t-1} + \varepsilon_t$, para todo $t = 1, 2, \dots$, siendo $(\varepsilon_t)_{t=1,2,\dots}$ ruido blanco (gaussiano) de varianza σ^2 . Suponiendo el proceso $(X_t)_{t=0,1,2,\dots}$ es débilmente estacionario, obtener la función de medias y correlaciones.

Nota: El proceso del enunciado es un proceso denominado AR(1) que estudiará en el tema posterior.

20) Sea $(X_t)_{t=0,1,2,\dots}$ un paseo aleatorio simple. Es decir, $X_0 = 0$ y $X_t = X_{t-1} + Y_t$, $(Y_t)_{t=0,1,2,\dots}$ variables aleatorias i.i.d verificando $P(Y_t = 1) = p$, $P(Y_t = -1) = 1 - p$, $p \in (0, 1)$.

(a) Calcular la función de medias, covarianzas y correlaciones del proceso $(X_t)_{t=0,1,2,\dots}$.

(b) Justificar si el proceso es estacionario en sentido débil.

(c) Justificar si el proceso es estacionario (en sentido fuerte).

Tema 2: Cadenas de Markov

En este tema vamos a estudiar un tipo particular de procesos estocásticos: las cadenas de Markov. Este tipo de procesos describen cambios de estado en un sistema, con la peculiaridad de que dichos cambios dependen únicamente del estado actual y no están influenciados por ningún estado que haya tomado previamente.

Por ejemplo, imaginemos un aparcamiento y consideremos la variable X_n representando el número de plazas de aparcamiento ocupadas en cada instante de tiempo $n \in \{0, 1, 2, \dots\}$. Claramente el valor de X_{n+1} dependerá de X_n ya que X_{n+1} se obtiene de X_n sumándole los coches que han aparcado y restándole los que se han ido entre los instantes n y $n + 1$. Por tanto, conocido X_n , parece razonable que la cantidad X_{n+1} no dependa de los valores previos $X_0, X_1, X_2, \dots, X_{n-1}$.

Este tipo de procesos, en los que el valor X_{n+1} depende exclusivamente de X_n y no se ve influenciado por los estados previos $X_0, X_1, X_2, \dots, x_{n-1}$ es formalizado matemáticamente mediante el concepto de cadena de Markov.

2.1) Cadenas de Markov de tiempo discreto

Definición 2.1.1 Un proceso estocástico $(X_n)_{n=0,1,2,\dots}$ de tiempo discreto se dice que es una **cadena de Markov** (de tiempo discreto) si se verifica:

- 1) Cada X_n toma valores en un conjunto numerable (es decir, finito o infinito numerable) \mathcal{S} llamado **espacio de estados**.
- 2) Se cumple que

$$\mathbb{P}(X_{n+1} = a_{n+1} | X_n = a_n, X_{n-1} = a_{n-1}, \dots, X_0 = a_0) = \mathbb{P}(X_{n+1} = a_{n+1} | X_n = a_n), \quad (1)$$

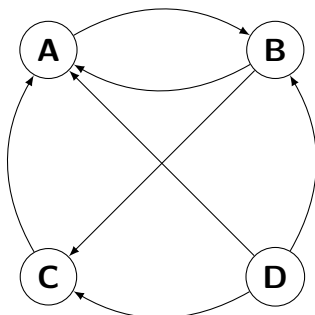
donde $a_0, a_1, \dots, a_n, a_{n+1} \in \mathcal{S}$.

La condición (1) arriba se llama **propiedad de Markov**. La interpretación es que la probabilidad de cualquier valor futuro del proceso, dado el valor actual, no está influenciada por ningún valor pasado. Se dice que las cadenas de Markov son procesos estocásticos sin memoria.

Ejemplo. (PageRank) es un algoritmo creado y desarrollado por la compañía tecnológica estadounidense Google para ordenar las apariciones de las páginas en cada búsqueda, dando preferencia a aquellas páginas que sean más "importantes" o "populares". Para medir esto, se analiza la cadena de Markov resultante de un individuo o "surfeador de la web" que va pulsando links al azar en un conjunto de páginas de internet. Por ejemplo, supongamos que el surfeador de la web navega haciendo clics al azar en las páginas A,B,C,D donde:

- A tiene enlace a B.
- B tiene enlaces a A y C.
- C tiene enlace a A.
- D tiene enlaces a las otras tres páginas.

Este proceso da lugar a una cadena de Markov con espacio de estados $\{A, B, C, D\}$, el cuál puede ser descrito mediante el siguiente grafo.



Además, tenemos las siguientes probabilidades de transición entre estados.

	A	B	C	D
A	0%	100%	0%	0%
B	50%	0%	50%	0%
C	100%	0%	0%	0%
D	33.3333%	33.3333%	33.3333%	0%

Claramente este proceso da lugar a una cadena de Markov, ya que, en cada paso, las probabilidades de visitar una página u otra, sólo depende de en qué página se encuentre el surfedor, sin importar qué páginas haya visitado anteriormente.

El algoritmo PageRank trata de determinar la probabilidad con la que una página es visitada a medida que el surfedor hace más y más clics, considerando como más importantes aquellas páginas para las que esta probabilidad. Éste es el criterio usado para ordenar las páginas en cada búsqueda en Google. Analizaremos este problema en detalle al final del tema.

Ejemplo. El paseo aleatorio simétrico $(X_n)_{n=0,1,2,\dots}$ es una cadena de Markov con espacio de estados infinito

$$\mathcal{S} = \mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

En este caso, para todo $i \in \mathcal{S}$

$$\begin{aligned}\mathbb{P}(X_{n+1} = i + 1 | X_n = i) &= \mathbb{P}(X_{n+1} = i - 1 | X_n = i) = \frac{1}{2}, \\ \mathbb{P}(X_{n+1} = j | X_n = i) &= 0 \text{ si } j \neq i \pm 1.\end{aligned}$$

Definición 2.1.2 Supongamos que $(X_n)_{n=0,1,2,\dots}$ es una cadena de Markov con espacio de estados \mathcal{S} . Las **probabilidades de transición** son las probabilidades

$$p_{x,y}(n) = \mathbb{P}(X_n = y | X_{n-1} = x),$$

donde $x, y \in \mathcal{S}$ y $n = 1, 2, 3, \dots$

Definición 2.1.3 Decimos que la cadena de Markov $(X_n)_{n=0,1,2,\dots}$ es **homogénea** si las probabilidades de transición no dependen del tiempo. En tal caso, definimos

$$\begin{aligned}p_{x,y} &= \mathbb{P}(X_n = y | X_{n-1} = x) \\ &= \mathbb{P}(X_1 = y | X_0 = x)\end{aligned}$$

donde hemos eliminado n de la notación.

En el resto de este tema vamos a trabajar con cadenas de Markov homogéneas con un número de estados **finito**. Por lo que a partir de ahora, cada vez que nos refiramos a una **cadena** estaremos hablando de una cadena de Markov con esas características.

En general, usaremos que $(X_n)_{n=0,1,2,\dots}$ es una cadena con espacio de estados finito que denotaremos por $\mathcal{S} = \{1, 2, 3, \dots, N\}$.

Definición 2.1.4 Definimos la **matriz de transición** de $(X_n)_{n=0,1,2,\dots}$ como

$$P = (p_{i,j})_{i,j=1,2,\dots,N} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N} \\ \cdots & \cdots & \cdots & \cdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,N} \end{bmatrix}$$

Obsérvese que las entradas de la fila 1 describen todas las probabilidades posibles condicionadas a empezar en el estado

$i = 1$. Por tanto, la suma de todas ellas debe ser igual a 1. Obviamente, lo mismo es cierto para el resto de filas, es decir, $p_{i,1} + p_{i,2} + \dots + p_{i,N} = 1$ para cada i . Por tanto, en una matriz de transición, la suma de todos los elementos de cualquier fila debe de ser 1. Sin embargo, esta condición no tiene por qué verificarse para las columnas de una matriz de transición.

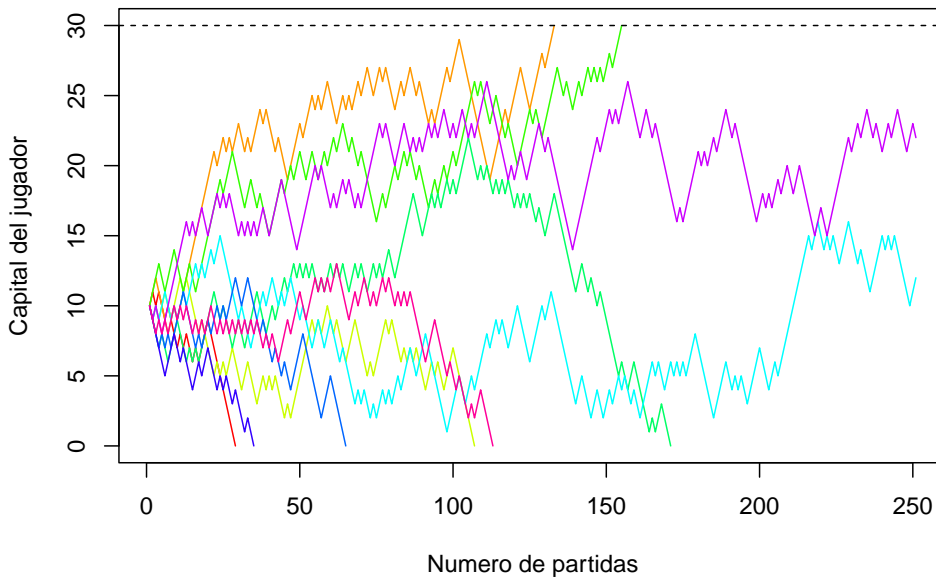
Ejemplo. (Urnas de Ehrenfest) Supongamos que tenemos dos urnas, U_1 y U_2 . En ellas están distribuidas N bolas numeradas. En cada paso, se elige un número al azar entre $\{1, 2, \dots, N\}$. A continuación se observa en qué urna está la bola con el número elegido y se cambia de urna. Denotemos por X_n el número de bolas contenidas en la urna U_1 en tiempo n . De esta manera, definimos una cadena $(X_n)_{n=0,1,2,\dots}$ con espacio de estados $\mathcal{S} = \{0, 1, 2, \dots, N\}$ y probabilidades de transición

$$\begin{aligned}\mathbb{P}(X_n = i + 1 | X_{n-1} = i) &= \frac{N - i}{N}, \\ \mathbb{P}(X_n = i - 1 | X_{n-1} = i) &= \frac{i}{N}, \\ \mathbb{P}(X_n = j | X_{n-1} = i) &= 0 \quad \text{si } j \neq i \pm 1.\end{aligned}$$

Por ejemplo, si $N = 3$ tenemos espacio de estados $\{0, 1, 2, 3\}$ y una matriz de transición

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Ejemplo. (Ruina del jugador) Consideremos un individuo que juega a la ruleta, que posee una riqueza inicial de X_0 euros, y que apuesta 1 euro al rojo en cada jugada con probabilidad de ganar p . El jugador seguirá apostando hasta que, o bien alcance una riqueza objetivo M , o bien hasta que se arruine. El proceso $(X_n)_{n=0,1,2,\dots}$ de la riqueza acumulada hasta la jugada n es una cadena de Markov con espacio de estados finito $\mathcal{S} = \{0, 1, \dots, M\}$.



En este caso, el proceso de la fortuna acumulada $(X_n)_{n=0,1,2,\dots}$ es una cadena de estados $\{0, 1, 2, \dots, M\}$ y transición

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1-p & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & p & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

2.2) Matriz de transición de n pasos

En esta sección, sea $(X_n)_{n=0,1,2,\dots}$ una cadena con espacio de estados $\mathcal{S} = \{1, 2, \dots, N\}$. En general, definimos lo siguiente.

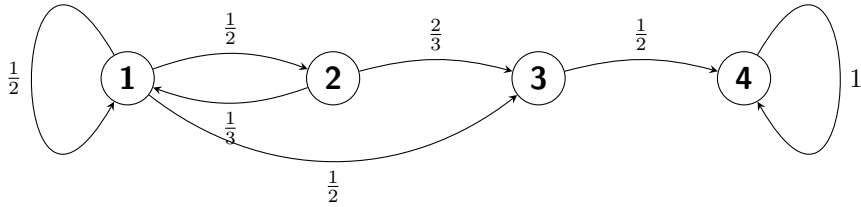
Definición 2.2.1 Dados dos estados $i, j \in \mathcal{S}$, definimos la probabilidad de transición de n pasos

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i).$$

Definimos la **matriz de transición de n pasos** de $(X_n)_{n=0,1,2,\dots}$ como

$$P^{(n)} = \left(p_{i,j}^{(n)} \right)_{i,j=1,2,\dots,N} = \begin{bmatrix} p_{1,1}^{(n)} & p_{1,2}^{(n)} & \cdots & p_{1,N}^{(n)} \\ p_{2,1}^{(n)} & p_{2,2}^{(n)} & \cdots & p_{2,N}^{(n)} \\ \cdots & \cdots & \cdots & \cdots \\ p_{N,1}^{(n)} & p_{N,2}^{(n)} & \cdots & p_{N,N}^{(n)} \end{bmatrix}.$$

Para entender cómo calcular la matriz de transición de n pasos analicemos el siguiente ejemplo. Consideremos la cadena $(X_n)_{n=0,1,2,\dots}$ dada por el grafo:



Esta cadena tiene matriz de transición

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Queremos ver lo que ocurre tras dos pasos del proceso.

Por ejemplo, veamos la probabilidad de llegar al estado 1 desde el estado 1 en dos pasos

$$p_{1,1}^{(2)} = \mathbb{P}(111) + \mathbb{P}(121) = p_{1,1} \cdot p_{1,1} + p_{1,2} \cdot p_{2,1} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} = \frac{5}{12}$$

Incluyendo todos los posibles caminos, incluso aquellos que no existen en el grafo, podemos poner

$$\begin{aligned} p_{1,1}^{(2)} &= \mathbb{P}(111) + \mathbb{P}(121) + \mathbb{P}(131) + \mathbb{P}(141) \\ &= p_{1,1} \cdot p_{1,1} + p_{1,2} \cdot p_{2,1} + p_{1,3} \cdot p_{3,1} + p_{1,4} \cdot p_{4,1} \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} + 0 + 0 = \frac{5}{12}. \end{aligned}$$

Escrito de esta manera, vemos que $p_{1,1}^{(2)}$ es la primera entrada de la matriz $P^2 = P \cdot P$. Es más, tenemos que $P^2 = (p_{i,j}^2)_{i,j \in \{1,2,3,4\}}$.

En general, tenemos lo siguiente.

Teorema 2.2.1 *La matriz de transición de n pasos vendrá dada por n -ésima potencia de P . Es decir,*

$$P^n = (p_{i,j}^{(n)})_{i,j=1,2,\dots,N}.$$

Como consecuencia, se tiene la relación:

$$\pi^{(n)} = \pi^{(0)} \cdot P^n$$

donde:

$$\pi^{(0)} = (\pi_1^{(0)}, \dots, \pi_N^{(0)})$$

denota el vector de probabilidades iniciales de cada estado, es decir, $\pi_j^{(0)} = \mathbb{P}(X_0 = j)$ y

$$\pi^{(n)} = (\pi_1^{(n)}, \dots, \pi_N^{(n)})$$

denota el vector de probabilidades iniciales de cada estado en el instante n , es decir, $\pi_j^{(n)} = \mathbb{P}(X_n = j)$, para todo $j = 1, 2, \dots, N$.

Teniendo en cuenta el teorema anterior junto al hecho de que $P^{m+n} = P^m P^n$, tenemos lo siguiente.

Corolario 2.2.1 *(Ecuaciones de Chapman-Kolmogorov)*

$$p_{i,j}^{(n+m)} = \sum_{k=1}^N p_{i,k}^{(n)} p_{k,j}^{(m)}.$$

2.3) Clasificación de los estados

Sea $(X_n)_{n=0,1,2,\dots}$ una cadena con espacio de estados \mathcal{S} . Dados dos estados $x, y \in \mathcal{S}$, definimos

$$r_{x,y} = \mathbb{P}(X_n = y \text{ para algún } n \geq 1 | X_0 = x).$$

Esto es, $r_{x,y}$ es la probabilidad de que la cadena alcance el estado y (en algún tiempo futuro) si la cadena se inicia en el estado x .

Definición 2.3.1 Sean $x, y \in \mathcal{S}$ con $x \neq y$.

- 1) Decimos que y es **accesible** desde x si $r_{x,y} > 0$. En tal caso escribimos $x \rightarrow y$.
- 2) Decimos que x se **comunica** con y si son accesibles entre sí (es decir, $x \rightarrow y, y \rightarrow x$). En tal caso escribimos $x \leftrightarrow y$.

Por convenio, se considera que cualquier estado x es accesible desde sí mismo ($x \rightarrow y$), y que se comunica consigo mismo ($x \leftrightarrow x$), incluyendo el caso de que $r_{x,x} = 0$.

Cada cadena puede ser representada con un grado. Podemos ver la accesibilidad simplemente observando si en el grafo existe un camino desde x hasta y respetando la dirección de las flechas. Cuando haya caminos en ambas direcciones, significará que ambos estados se comunican.

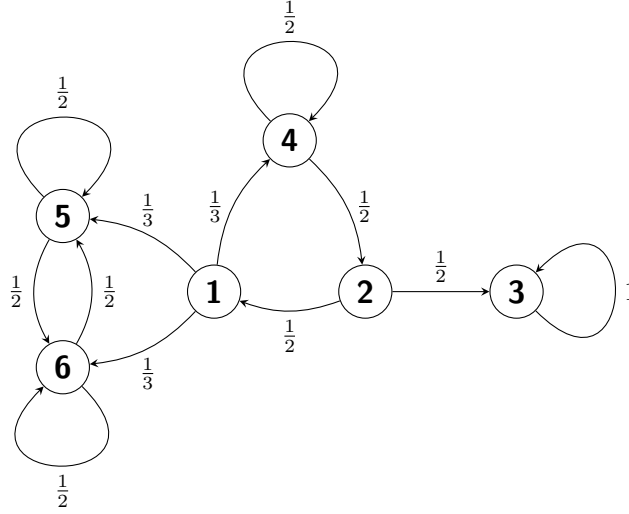


Figure 3: Observando las flechas podemos analizar la accesibilidad y la conexión entre estados

Por ejemplo, consideremos la cadena dada por el grafo de la Figura 3. Vemos que por ejemplo que $1 \rightarrow 3$ ya que podemos ir desde 1 hasta 3, siguiendo la dirección de las flechas, pasando por 4 y 2. Sin embargo, $3 \nrightarrow 1$, ya que de 3 sólo se puede volver a saltar a sí mismo. Por otro lado, $1 \leftrightarrow 2$ ya que podemos conectar ambos estados por caminos tanto empezando en 1 como empezando en 2.

En general, si un estado y es accesible desde un estado x , siempre será posible encontrar un camino desde x hasta y de modo que dicho camino no pase por un mismo estado dos veces. Para ello, basta considerar cualquier camino desde x hasta y , y si algún estado z aparece dos veces en el camino, eliminamos el tramos del camino desde la primera aparición de dicho estado hasta la última aparición del mismo, de modo que aparezca una sola vez. Dicho camino, al no pasar dos veces por un mismo estado, tendrá como mucho tantos pasos como estados tenga la cadena. En definitiva, tenemos lo siguiente.

Proposición 2.3.1 *Supongamos que el espacio de estados \mathcal{S} tiene N elementos. Entonces,*

$$x \rightarrow y \quad \text{si, y sólo si,} \quad p_{x,y}^{(n)} > 0 \quad \text{para algún } n \leq N.$$

La relación entre estados definida por la propiedad de estar comunicados es una relación de equivalencia. Es decir, tenemos lo siguiente.

Proposición 2.3.2 *Dado el espacio de estados \mathcal{S} , siempre es posible dividirlo en clases disjuntas*

$$\mathcal{S} = C_1 \cup C_2 \cup \dots \cup C_n \quad (2)$$

donde para todo $x, y \in \mathcal{S}$ se verifica

$$\begin{aligned} x \leftrightarrow y & \quad \text{si } x, y \in C_i, \\ x \nleftrightarrow y & \quad \text{si } x \in C_i, y \in C_j, i \neq j. \end{aligned}$$

Las clases en (2) se llaman *clases irreducibles*. La proposición nos dice que cada cadena admite siempre una descomposición en clases irreducibles. Una cadena se dice que es *irreducible* si sólo posee una clase irreducible, es decir, todos sus estados se comunican entre sí.

Las clases irreducibles pueden ser fácilmente identificadas mirando el grafo de la cadena. Observando de nuevo el grado de la

cadena de la Figura 3, vemos que $1 \rightarrow 5$ pero $5 \nrightarrow 1$, por lo que están en clases diferentes. Por otro lado, $1 \rightarrow 4$, $4 \rightarrow 2$, $2 \rightarrow 1$, por lo que están en la misma clase. Finalmente, 3 no está comunicado con nadie, luego forma él solo una clase. En la Figura 4, podemos ver las tres clases irreducibles de la cadena que hemos identificado.

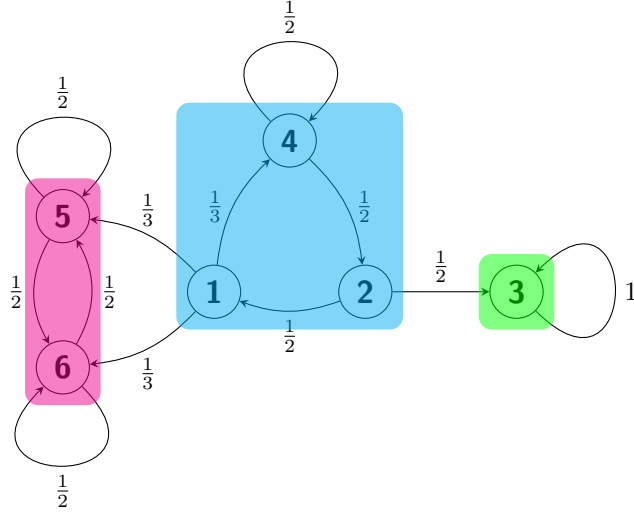


Figure 4: Tenemos en distinto color las tres clases irreducibles

A continuación vamos a introducir un criterio para clasificar los distintos estados de una cadena.

Definición 2.3.2 Sea $x \in S$.

- Decimos que x es **recurrente** si $r_{x,x} = 1$.
- Decimos que x es **transitorio** si $r_{x,x} < 1$.
- Decimos que x es **absorbente** si $p_{x,x} = 1$.

Si x es recurrente, tendremos que una vez que la cadena alcanza el estado x entonces tendremos total certeza de que volverá a alcanzar el estado x en el futuro.

Si x es transitorio, tendremos que una vez que la cadena alcanza el estado x entonces no tendremos certeza de que la cadena vuelva a alcanzar el estado x de nuevo.

Si x es absorbente, tendremos que una vez que la cadena alcanza el estado x con toda certeza permanezca en el estado x en el futuro. En particular, todo estado absorbente es también recurrente.

Mirando de nuevo al grafo de la Figura 4 arriba, vemos que el estado 3 es absorbente (y por tanto recurrente), ya que si comenzamos en él solo llegaremos al él mismo. Los estados 5 y 6 son recurrentes. Por ejemplo, vemos que si empezamos en 5 no podremos llegar a ningún otro estado excepto 6 y él mismo. Además, la única manera para que, empezando en 5, no se vuelva a visitar 5 es que la cadena se cambie al estado 6 y permanezca en ese estado en lo sucesivo. Pero la probabilidad de que eso ocurra es $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdots = 0$. Así que con probabilidad 1 se volverá en algún momento a 5, y por lo tanto es recurrente. El mismo argumento es aplicable a 6. Finalmente, los estados 1, 2 y 4 son transitorios ya que desde esos estados se accede a zonas de las que no se puede volver.