Introducción a las Series de Tiempo

Gustavo A. Caffaro

Instituto Tecnológico de Santo Domingo

Nov 2021 - Ene 2022

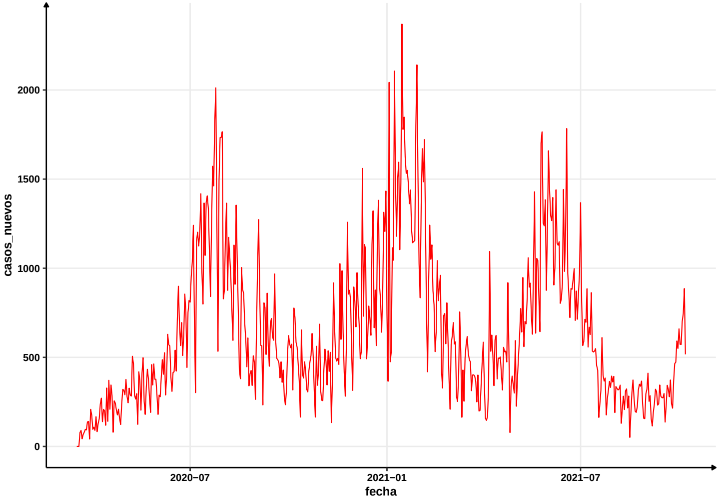
Introducción

En este curso aprenderemos y aplicaremos métodos estadísticos para analizar datos temporales, los cuales llamamos *series de tiempo* o series temporales.

La clase tendrá un formato aplicado pero riguroso.

Como veremos más adelante, el desafío principal de este curso se encuentra en que las observaciones temporales suelen estar muy correlacionadas.

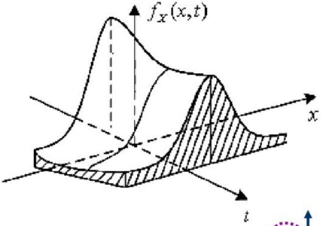
Casos de Covid en RD

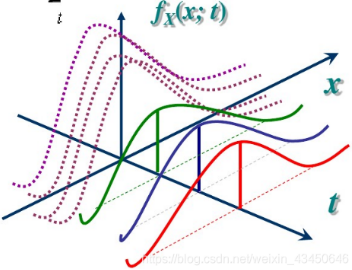


Contenido de la clase

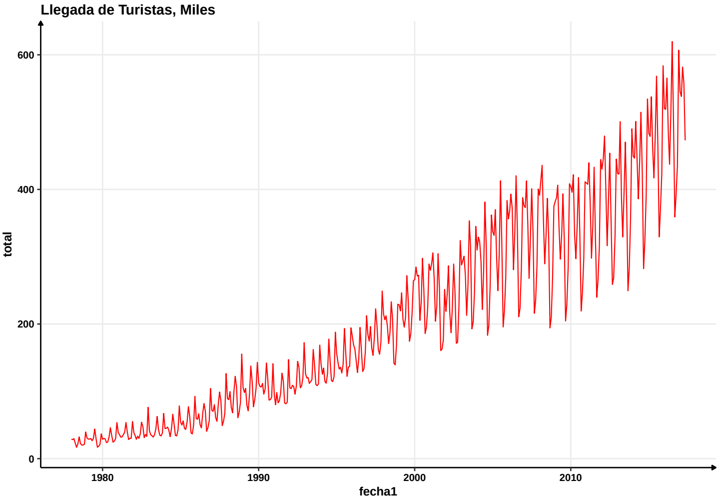
- Modelos univariados de series de tiempo: Modelos autoregresivos y de media móvil (ARIMA)
- Herramientas de identificación, estimación, y evaluación de modelos.
- Usar modelos para proyectar el futuro
- Métodos de suavización y descomposición de tendencias
- Relación entre distintas series de tiempo
- Vectores autoregresivos
- Modelos de estado-espacio
- Modelos financieros econométricos

Series de tiempo como procesos estocásticos

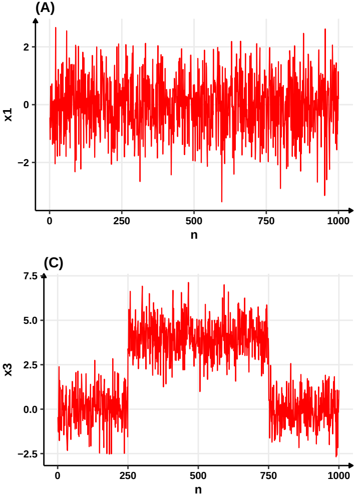




Llegada de turistas



Algunos ejemplos de series de tiempo



Algunos ejemplos de series de tiempo

(A):

- $x_t \overset{iid}{\sim} N(0,1)$

(B):

- $x_t \sim N(0,\sigma^2)$
  - $\sigma = 5$ si $250 \leq t \leq 750$
  - $\sigma = 1$ si no

(C):

- $x_t = a + \epsilon_t, \quad \epsilon_t \sim N(0,1)$
  - $a = 4$ si $250 \leq t \leq 750$
  - $a = 0$ si no

(D):

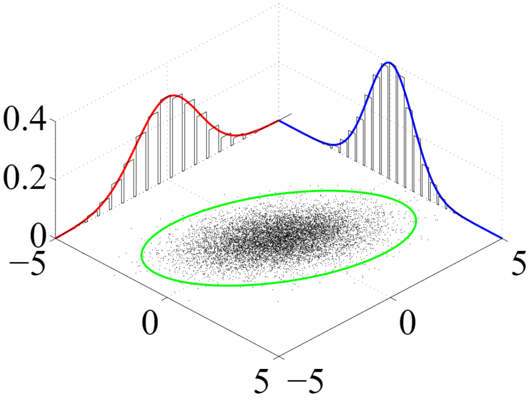
- $x_t = 0.1 + x_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0,1)$

Definición de una serie de tiempo

Una serie de tiempo es una colección de n variables aleatorias en puntos arbitrarios $t_1, t_2, \ldots, t_n$, para cualquier entero positivo n, donde su probabilidad conjunta viene dada por

$$F_{t_1,t_2,\ldots,t_n}(c_1, c_2, \ldots, c_n) = \Pr(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \ldots, , x_{t_n} \leq c_n)$$

Distribución bidimensional



Definición

A pesar de que la función de distribución conjunta puede describir la data completamente, no es la herramienta más directa para mostrar ni analizar las series de tiempo.

La **función marginal de distribución**:

$$F_t(x) = P(x_t \leq x)$$

La **densidad marginal**:

$$f_t(x) = \frac{\partial F_t(x)}{\partial x}$$

La media

La **función de la media** se define como

$$\mu_{xt} = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx$$

en el caso de variables contínuas, y

$$\mu_{xt} = E(x_t) = \sum_{i=1}^{n} x_i f_t(x_i)$$

en el caso de variables discretas.

Si no existe confusión respecto a la serie de tiempo que nos referimos, podemos escribir μ_t en vez de μ_{xt}.

La media (media móvil)

Ej.: Media móvil

Si $w_t \sim N(0, \sigma^2)$, y definimos su media móvil, v_t, como el promedio entre el valor actual de w_t y los períodos adyacentes, tenemos que

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

Por tanto, la función de la media de v_t es

$$\mu_{vt} = E(v_t) = \frac{1}{3}\Big[E(w_{t-1}) + E(w_t) + E(w_{t+1})\Big] = 0$$

La media (RW)

Ej. La caminata aleatoria (RW)

Una caminata aleatoria es una suma de ruidos blancos gaussianos, $w_t \sim N(0, \sigma^2)$:

$$x_t = \sum_{j=1}^{t} w_j$$

Por lo que su media es

$$E(x_t) = E\left(\sum_{j=1}^{t} w_j\right) = \sum_{j=1}^{t} E(w_j) = 0$$

La media (caminata aleatoria con tendencia)

Ej. La caminata aleatoria con tendencia

Podíamos incluir una tendencia en el modelo de caminata aleatoria de la siguiente manera:

$$x_t = \delta t + \sum_{j=1}^{t} w_j,$$

donde $\delta \in \mathbb{R}$. Por lo tanto, su media es

$$E(x_t) = E\left(\delta t + \sum_{j=1}^{t} w_j\right) = \delta t + \sum_{j=1}^{t} E(w_j) = \delta t$$

ya que t es conocido en cada período.

La Media

```
n <- 5e2
set.seed(n)
x1 <- rnorm(n) # ruido blanco gaussiano
x2 <- stats::filter(x1, sides=2, filter=c(1,1,1)/3) # media movil
x3 <- cumsum(x1) # caminata aleat.
mean(x1);mean(x2,na.rm=T);mean(x3)
```

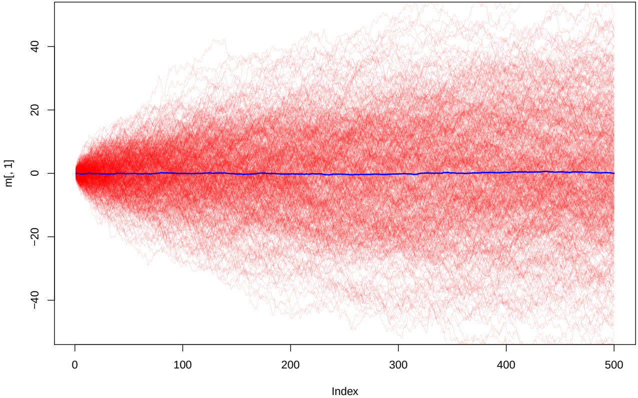
```
## [1] -0.04556147

## [1] -0.04865267

## [1] -14.92419
```

```
m <- matrix(n^2,n,n)
medias <- numeric(n)
for (i in 1:n){
  m[,i] <- cumsum(rnorm(n)) # caminata aleat.
  medias[i] <- mean(m[,i])
}
mean(medias)
```

```
## [1] 0.006638905
```

La Media (RW)

```
plot(m[,1], type="l", col=rgb(1,0,0, alpha = 0.05),ylim=c(-50,50))
for (i in 2:n){
  lines(m[,i], type="l", col=rgb(1,0,0, alpha = 0.1))
}
lines(rowMeans(m),col="blue",lwd=2)
```



La autocovarianza

Si tenemos una serie x_t y queremos tener una medida de cómo se relacionan dos variables aleatorias de la misma serie, podemos utilizar la **función de autocovarianza**, la cual se define como:

$$\gamma_x(s,t) = cov(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)],$$

para todo s y t, por lo que es obvio que $\gamma_x(s,t) = \gamma_x(t,s)$.

La autocovarianza mide la dependencia lineal entre dos puntos de la misma serie observado en tiempos distintos.

Usualmente, para s y t lejanos, las series muy suaves tienden a tener autocovarianzas elevadas, mientras que las series entrecortadas tienden a tener autocovarianzas cercanas a cero.

La autocovarianza

Ej. Ruido blanco gaussiano

Recordemos que $w_t \overset{iid}{\sim} N(0, \sigma^2)$. Por tanto:

$$\gamma_w(s,t) = cov(w_s, w_t) = \begin{cases} \sigma^2, & \text{si } s = t \\ 0, & \text{si no} \end{cases}$$

Si X y Y son dos variables aleatorias y $a, b \in \mathbb{R}$, entonces

$$cov(aX, bY) = ab \cdot cov(X, Y)$$

La autocovarianza

Ej. Media móvil

Si aplicamos la media móvil de tres puntos a una serie de ruido blanco, $w_t \sim N(0, \sigma_w^2)$, entonces su autocovarianza viene dada por:

$$\gamma_v(s,t) = cov(v_s, v_t) = cov\left[\frac{1}{3}(w_{s-1} + w_s + w_{s+1}), \frac{1}{3}(w_{t-1} + w_t + w_{t+1})\right].$$

Cuando $s = t$, entonces

$$\begin{aligned}
\gamma_v(t,t) &= \frac{1}{9} cov\left[(w_{t-1} + w_t + w_{t+1}), (w_{t-1} + w_t + w_{t+1})\right] \\
&= \frac{1}{9} E\left\{\left(w_{t-1} + w_t + w_{t+1}\right)^2\right\} \\
&= \frac{1}{9} E(w_{t-1}^2) + E(w_t^2) + E(w_{t+1}^2) \\
&= \frac{1}{3}\sigma_w^2
\end{aligned}$$

La autocovarianza

Cuando $s = t + 1$,

$$\begin{aligned}\gamma_v(t+1,t) &= \frac{1}{9} cov\left[(w_t + w_{t+1} + w_{t+2}), (w_{t-1} + w_t + w_{t+1})\right] \\ &= \frac{1}{9} E\left\{\left(w_t + w_{t+1}\right)^2\right\} \\ &= \frac{2}{9}\sigma_w^2\end{aligned}.$$

De manera similar, si $s = t - 1$, $\gamma_v(t-1,t) = 2\sigma_w^2/9$,

$$\gamma_v(t-2,t) = \gamma_v(t+2,t) = \sigma_w^2/9,$$

y

$$\gamma_v(t \pm h, t) = 0$$

para $h > 2$.

La autocovarianza cruzada

En el caso en el que queremos medir la relación entre dos puntos de dos series x_t y y_t, podemos utilizar la **autocovarianza cruzada**, la cual se define como:

$$\gamma_{xy}(s,t) = cov(x_s, y_t) = E[(x_s - \mu_{xs})(y_t - \mu_{yt})]$$

La desigualdad de Cauchy-Schwarz

Digamos que X y Y son dos variables aleatorias. La **desigualdad de Cauchy-Schwarz** mantiene que

$$cov(X, Y)^2 \leq var(X)var(Y).$$

Prueba

Consideremos la varianza de la siguiente combinación lineal de X y Y:

$$0 \leq var(kX + Y) = var(X)k^2 + 2cov(X, Y)k + var(Y)$$

Observa que esta es una ecuación cuadrática en k.

Reescribiendo, tenemos

$$0 \leq Ak^2 + Bk + C,$$

una función cuadrática no negativa, por lo que su discriminante es no positivo:

$$B^2 - 4AC \leq 0$$

La desigualdad de Cauchy-Schwarz

Si resustituímos, tenemos que

$$4cov(X, Y)^2 - 4var(X)var(Y) \leq 0$$

y

$$cov(X, Y)^2 \leq var(X)var(Y).$$

Si consideramos $X = x_s$ y $Y = x_t$, podemos escribir la desigualdad de Cauchy-Schwarz en términos de la autocovarianza:

$$\gamma(s, t)^2 \leq \gamma(s, s)\gamma(t, t)$$

La función de autocorrelación

La **función de autocorrelación (FAC)**, se define como

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}.$$

La FAC mide la predictibilidad lineal de la serie, en el tiempo t, usando sólo el valor x_s.

De manera similar, la **función de autocorrelación cruzada** viene dada por:

$$\rho_{xy}(s,t) = \frac{\gamma_{xy}(s,t)}{\sqrt{\gamma_x(s,s)\gamma_y(t,t)}}.$$

Nota que cuando utilizamos la palabra *cruzada*, nos referimos a dos series distintas.

Estacionariedad estricta

Un proceso estocástico $\{x_t \mid t = 1, 2, \dots\}$ es **estríctamente estacionario** si, para cada colección de índices temporales $1 \leq t_1 \leq t_2 \leq \cdots \leq t_m$, la distibución conjunta de

$$(x_{t_1}, x_{t_2}, \dots, x_{t_m})$$

es idéntica a la de

$$(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_m+h}).$$

Por lo que

$$\Pr(x_{t_1} \leq c_1, \dots, x_{t_m} \leq c_k) = \Pr(x_{t_1+h} \leq c_1, \dots, x_{t_m+h} \leq c_k)$$

para todo $k \in \mathbb{N}$, todo $t_1, t_2, ..., t_k$, todo $c_1, c_2, \dots, c_k$ y todo $h \in \mathbb{Z}$.

Estacionariedad estricta

Intuitivamente, la estacionariedad estricta nos dice que si tomamos una colección de términos en una serie, y nos desplazamos h términos, entonces la distribución de la colección original debe ser igual que la distribución de la colección desplazada.

Ej.: Si la serie x_t es estríctamente estacionaria y $h = 1$, entonces

$$\Pr(x_1 \leq a) = \Pr(x_2 \leq a) = \Pr(x_3 \leq a) = \cdots$$

$$\Pr(x_1 \leq a_1, x_{h+1} \leq a_{h+1}) = \Pr(x_t \leq a_t, x_{t+h} \leq a_{t+h})$$

Estacionariedad débil

Una serie es **(débilmente) estacionaria** sí y sólo sí

1. Tiene media constante e independiente del tiempo.
2. La función de autocovarianza, $\gamma(s, t)$ sólo depende de la diferencia entre s y t, $h = |s - t|$.

Una conclusión directa de la definición de estacionariedad débil es que $\gamma(s, t) = \gamma(s + j, t + j)$ para cualquier entero j.

En esta clase utilizaremos el término *estacionario* para referirnos a esta definición.

Autocovarianza y autocorrelación (series estacionarias)

La **función de autocovarianza de una serie estacionaria** viene dada por

$$\gamma(h) = cov(x_{t+h}, x_t) = E[(x_{t+h} - \mu)(x_t - \mu)].$$

De manera similar, la **función de autocorrelación de una serie estacionaria** viene dada por

$$\rho(h) = \frac{\gamma(t+h, t)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} = \frac{\gamma(h)}{\gamma(0)}$$

La desigualdad de Cauchy-Schwarz muestra que $-1 \leq \rho(h) \leq 1$ para todo h.

Autocovarianza y autocorrelación (RBG)

Ej.: Ruido Blanco Gaussiano (RBG)

Si $w_t \sim N(0, \sigma_w^2)$, entonces

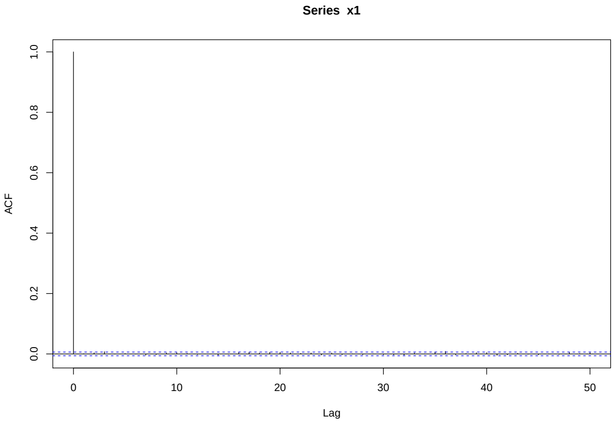
$$\gamma_w(h) = cov(w_{t+h}, w_t) = \begin{cases} \sigma_w^2, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

Como $\gamma_w(0) = \sigma_w^2$, entonces $\rho(h) = 1$ si $h = 0$ y $\rho(h) = 0$ si $h \neq 1$.

Adicionalmente, como $\mu_{wt} = \mu_w = 0$, entonces el ruido blanco gaussiano cumple con la definición de estacionariedad débil (también con la definición estricta!).

FAC (RBG)

```r
n <- 1e3
set.seed(n)
x1 <- rnorm(1e5)
acf(x1)
```



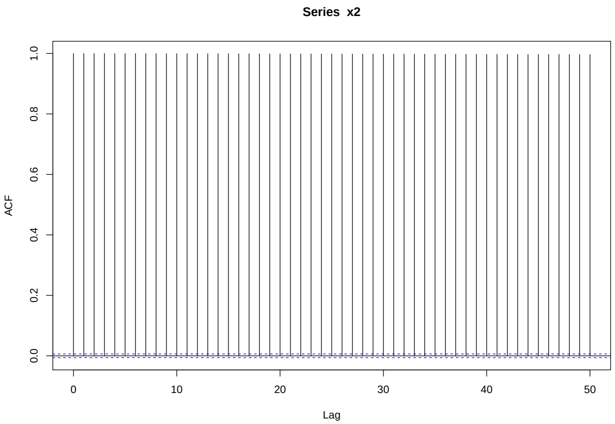
Autocovarianza y autocorrelación (RW)

Ej.: Caminata Aleatoria (RW)

Calculemos la función de covarianza, la f. de correlación y determinemos si es estacionaria.

FAC (RW)

```r
x2 <- cumsum(x1)
acf(x2)
```



Autocovarianza y autocorrelación (media móvil)

Ej.: Media móvil

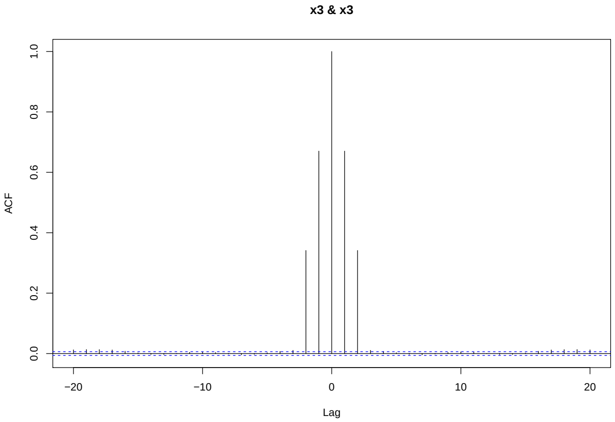
Recuerda que la autocovarianza de la media móvil de tres puntos viene dada por

$$\gamma_v(h) = \begin{cases} \frac{3}{9}\sigma_w^2, & h = 0 \\ \frac{2}{9}\sigma_w^2, & h = \pm 1 \\ \frac{1}{9}\sigma_w^2, & h = \pm 2 \\ 0, & |h| > 2. \end{cases}$$

Como la media $\mu_{vt} = \mu_v = 0$ y autocovarianza no dependen de t sino del número de pasos h, satisfacen la definición de débilmente estacionaria.

FAC (media móvil)

```
library(forecast)
x3 <- ma(x1,order=3,centre=T)
x3 <- x3[complete.cases(x3)] # ignorar NAs prod. por media movil
ccf(x3,x3,lag.max = 20) # funcion de correlacion cruzada
```



Propiedades de la función de autocovarianza

Si la serie x_t es estacionaria, entonces

1. $\gamma(h) = \gamma(-h)$
2. $|\gamma(h)| \leq \gamma(0)$

Prueba 1

$$\gamma(h) = cov(x_{t+h}, x_t) = cov(x_t, x_{t-h}) = \gamma(-h)$$

Prueba 2

Cuando $h = 0$, tenemos que $\gamma(0) = E[(x_t - \mu)^2]$.

La desigualdad de Cauchy-Schwarz implica que

$$|\gamma(h)| \leq \gamma(0)$$

Estacionariedad tendencial

Definamos una serie $x_t = \alpha + \beta t + y_t$, donde y_t es estacionaria. Entonces

$$\mu_{xt} = E(x_t) = \alpha + \beta t + \mu_y,$$

la cual no es independiente del tiempo. Por lo tanto, x_t no es estacionaria.

Sin embargo, la función de autocovarianza es independiente del tiempo, porque

$$\begin{aligned}\gamma_x(h) &= E[(x_{t+h} - \mu_{x,t+h})(x_t - \mu_{x,t})] \\ &= E[(y_{t+h} - \mu_y)(y_t - \mu_y)] \\ &= \gamma(h).\end{aligned}$$

Por lo tanto, se considera que el modeo tiene estacionariedad al rededor de su tendencia. A este comportamiento le llamamos *estacionariedad tendencial*.