Tema 03.01: Procesos estocásticos estacionarios con R

@umh1465: Análisis estadístico de series económicas

Xavi Barber

Centro de Investigación Operativa
Universidad Miguel Hernández de Elche

2018-03-08

Procesos estocásticos estacionarios

Proceso estocástico y Series temporales

- El mecanismo generador de los datos de las series temporales no es fijo pero no cambia por completo entre una observación y otra.
- No son variables aleatorias independiente e identicamente distribuidas (iid).
- Por lo tanto, una serie temporal podría mostrar:
  - Evolución en los parámetros que rigen las propiedades estadísticas del proceso generador de datos.
  - Dependencia entre las observaciones.

Luego el fenómeno inflación mensual en un determinado país será una secuencia (posiblemente infinita) de variables aleatoria ordenadas en el tiempo

$$x(1), x(2), \dots, x(t), \dots, x(T)$$

A esta secuencia de variables aleatorias se le denomina **proceso estocástico**.

Una serie temporal es una realización finita de un proceso estocástico.

Proceso estocástico	$x(1), x(2), \ldots, x(t), \ldots, x(T)$
Serie temporal (observada)	$x_1, x_2, \ldots, x_t$

Debemos tener claro y diferenciar entre:

- Proceso estocástico: la inflación mensual en España.
- Variable aleatoria: la inflación mensual en el mes t en España.
- La inflación observada en el mes t es la realización de esta variable aleatoria, pero otras realizaciones podrían ser posibles.

Evolución de la ley probabilística de un proceso estocástico

La evolutividad podría referirse a distintos parámetros de la ley probabilística.

- En fenómenos económicos, los principales parámetros evolutivos son:
  - Media
  - Varianza (desviación típica).

La media de una serie

Evolución en la varianza

- En series financieras, la varianza condicional no es constante sino que depende del pasado.
- Esto requiere una modelización específica que no se cubrirá en este curso sino en una asignatura específica de econometría de series financieras.

Series con tendencia

- Lineal

$$X_t = a + bt + W_t$$

- Exponencial

$$X_t = e^{a+bt} e^{W_t}$$
$$log(Xt) = a + bt + W_t t$$

Series temporales con un claro límite superior en su evolución.

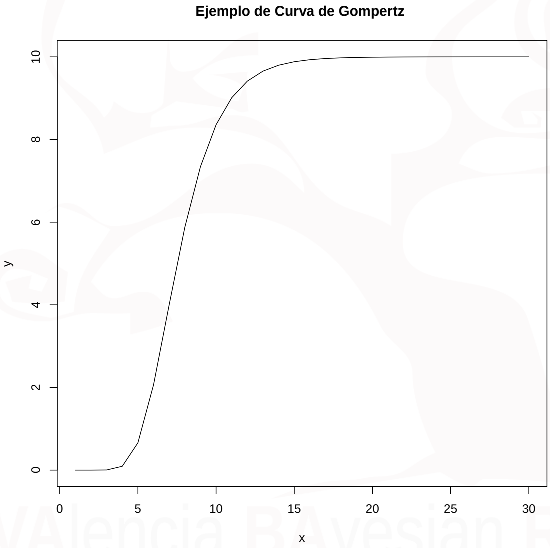
- Curva de Gompertz

$$log(T_t) = a + br^t, 0 < r < 1$$

- Curva logística

$$T_t = \frac{1}{a + br^t}, 0 < r < 1$$

- Posibles ejemplos:
    - Porcentaje de personas alfabetizadas en una sociedad desarrollada, porcentaje de casas con electricidad, etc.
    - En otros casos no está claro que exista un límite superior: número de teléfonos o automóviles por cada mil habitantes.


Ejemplo de Curva de Gompertz
y
0
2
4
6
8
10
0
5
10
15
20
25
30
x

Tendencia estocástica

Para ver como funciona el proceso $y_t =y_{t-1}+w_t

- Podemos realizar sustituciones recursivas. Si suponemos que existe un periodo inicial que denotamos por y0, la solución del sistema viene dada por

$$y_t = y_0 + \sum_i w_i$$

- Tomando expectativas obtenemos $E(y_t) = E(y_0)$. Es decir, se trata de un proceso con media constante.
- La varianza no es constante sino que depende del tiempo: $Var(y_t) = Var(y_0 + \sum_i w_i) = t\sigma^2$.

¿Cómo extraer la tendencia en la práctica?

- Si la tendencia es determinista mediante regresión. (Se puede eliminar también mediante diferenciación pero resulta una serie cuyas propiedades no son correctas).
- Si la tendencia es estocástica mediante diferenciación.

Así si X_t es un proceso con tendencia puramente estocástica. X_t es $I(2)$ ΔX_t es $I(1)$ $\Delta^2 X_t$ es $I(0)$, estacionario.

- Si la estacionalidad es puramente determinista mediante regresión.
- Si es estocástica tomando diferencias estacionales

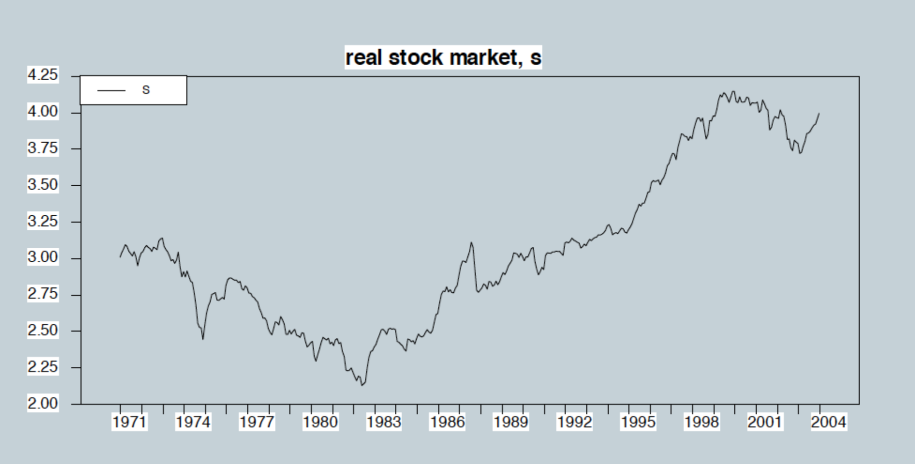

real stock market, s
s
4.25
4.00
3.75
3.50
3.25
3.00
2.75
2.50
2.25
2.00
1971
1974
1977
1980
1983
1986
1989
1992
1995
1998
2001
2004

Figure 1

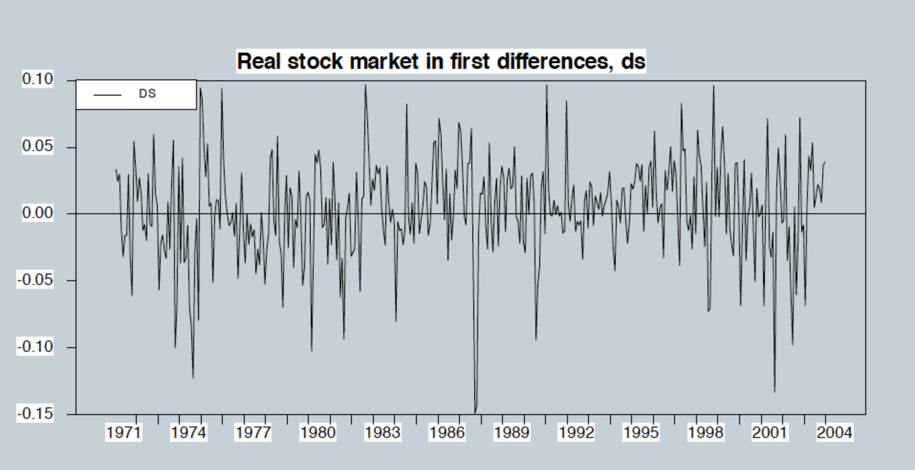

Real stock market in first differences, ds
DS
0.10
0.05
0.00
-0.05
-0.10
-0.15
1971
1974
1977
1980
1983
1986
1989
1992
1995
1998
2001
2004

Figure 2

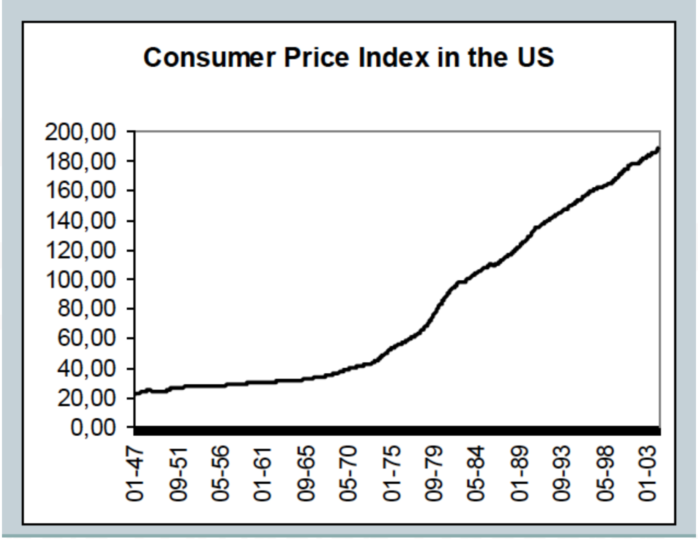

Consumer Price Index in the US
200,00
180,00
160,00
140,00
120,00
100,00
80,00
60,00
40,00
20,00
0,00
01-47
09-51
05-56
01-61
09-65
05-70
01-75
09-79
05-84
01-89
09-93
05-98
01-03

Figure 3

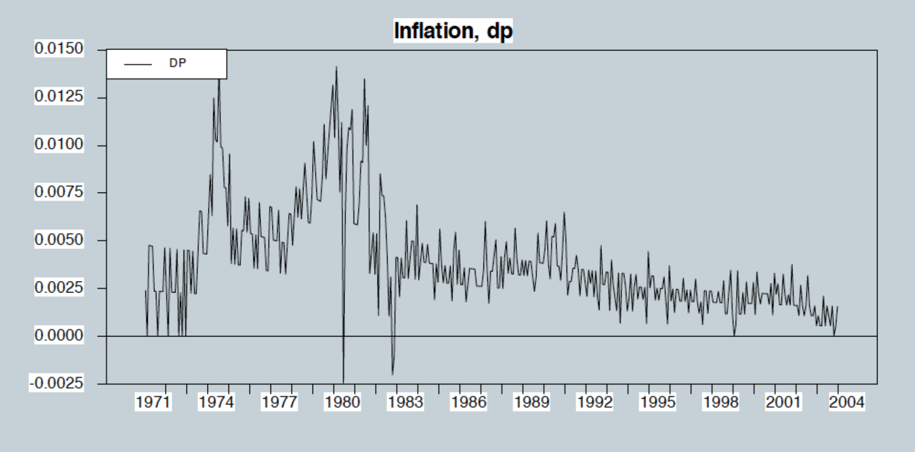

Inflation, dp
DP
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
-0.0025
1971
1974
1977
1980
1983
1986
1989
1992
1995
1998
2001
2004

Figure 4

Análisis de los modelos univariantes

- Estos modelos se caracterizan porque para conocer la dinámica de una variable se utiliza información sobre su pasado.
- No son modelos econométricos en el sentido que no utilizan formas funcionales que vienen dadas por la teoría económica. Sin embargo, la información que proporcionan estos modelos suele ser de mucha utilidad para el analista económico ya que proporciona información sobre las características del proceso.

Transformando las series de modo que en vez de considerar los datos originales se utilizan sus incrementos o los incrementos de los incrementos (si la serie presenta estacionalidad uno de dichos incrementos será estacional) tales transformaciones **no muestran evolución en el nivel y por lo tanto son estacionarias**.

La evolución tendencial de una serie que implica un determinado modelo se define a partir de su número de diferencias que incluye además del orden polinomial determinista

$$X_t = X_{t-1} + 0.5 + w_t$$

$$X_t = 2X_{t-1} - x_{t-2} + w_t$$

$$X_t = X_{t-1} + 0.5t + w_t$$

$$X_t = 2X_{t-1} - X_{t-2} + 0.5 + w_t$$

Los elementos determinantes en la definición de tendencia

- Los modelos del tipo

$$X_t = a + bt + W_t$$

$$X_t = X_{t-1} + b + W_t$$

- Pero si los parámetros a y b cambian con el periodo de predicción, los modelos integrados de orden 2 resultarán mejores para series con crecimiento sistemático.
- Asimismo un modelo integrado de orden 1 sería mejor que un modelo con un elevado número de rupturas deterministas para las series con oscilaciones locales de nivel.
- Box-Jenkins propone el uso del número máximo de diferencias.

Procesos estocásticos estacionarios

Supongamos que para una determinada serie eliminamos su tendencia con d diferencias regulares y D diferencias estacionales. Es importante tener en cuenta que en series económicas el número de diferencias totales, $d + D$, no debe ser nunca superior a 2.

Por ejemplo, veamos un caso extremo de crecimiento de precios en una serie mensual.

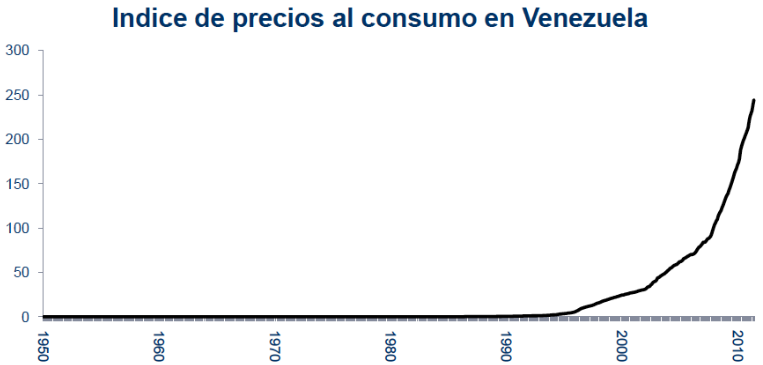


Figure 5

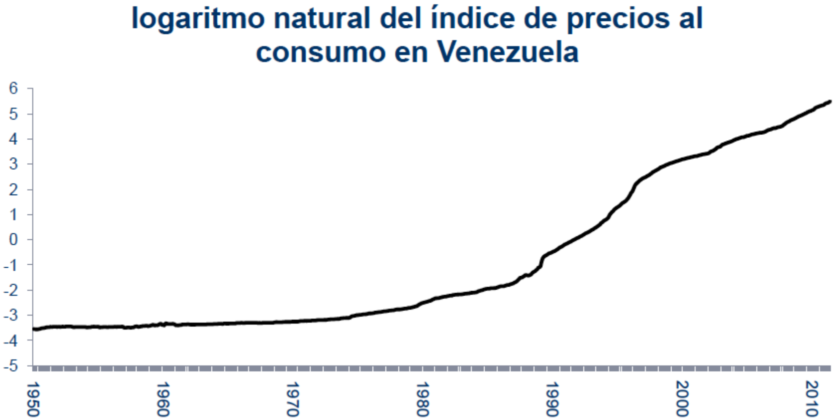


Figure 6

Incluso tras la transformación logarítmica vemos que la evolución de la serie es no lineal aproximadamente en el periodo que va desde 1980 hasta el 2000.

Procesos estocásticos estacionarios.

Si dispusiéramos de varias realizaciones de un proceso entonces la media para cada una de las variables podría ser estimada mediante:

$$\mu = \frac{\sum_{j=1}^{m} y_t^{(j)}}{m}$$

Sin embargo en la práctica sólo tenemos una realización por lo que no es posible estimar los momentos de las variables que componen el proceso.

Un **proceso estocástico es estacionario** si:

- La media es constante: $E(y_t) = \mu$ para todo t.
- La varianza es constante: $Var(y_t) = \sigma^2$
- La covarianza entre dos variables separadas h periodos sólo depende de h: $cov(y_t, y_{t+h}) = \gamma(h)$

En la práctica la restricción de estacionariedad nos permite estimar las covarianzas (medida de dependencia lineal entre observaciones separadas h periodos en el tiempo) mediante la covarianza muestral.

$$\hat{\gamma}(h) = \frac{\sum_{t=h+a}^{T}(y_t - \bar{y})(y_{t+h} - \bar{y})}{T}$$

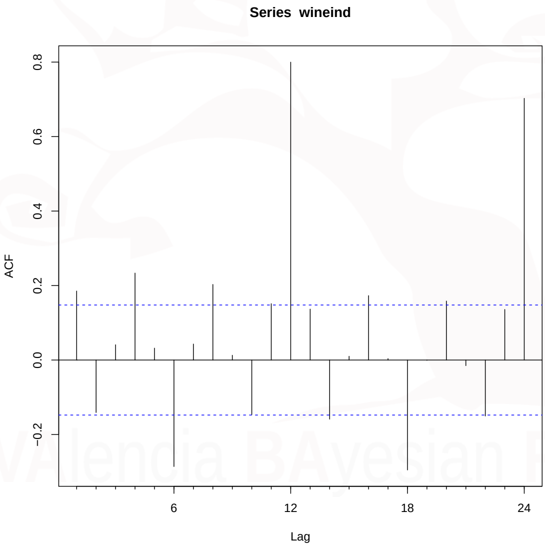
Correlación y Correlograma

- La estimacionde correlacionesesmas utilal no dependerde lasunidadesde medidade la serie.

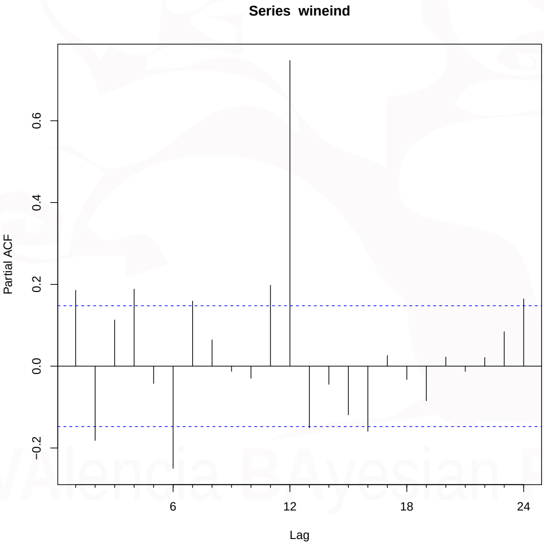
$$\hat{\rho} = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}$$

- El gráfico de las autocorrelaciones para diferentes valores de k constituyen el **correlograma**.

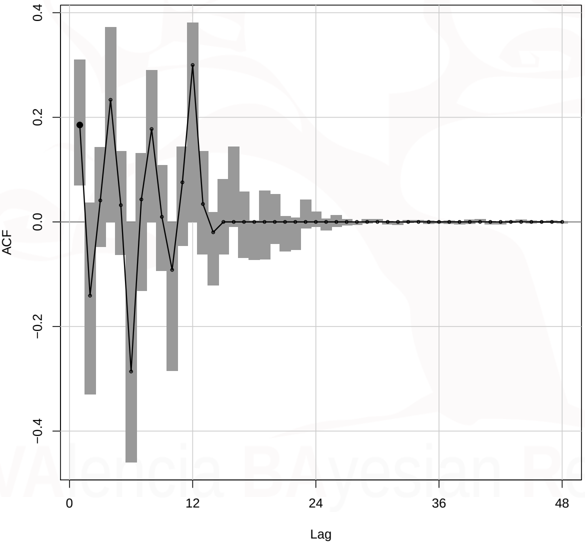
```r
library(forecast)
Acf(wineind)
```



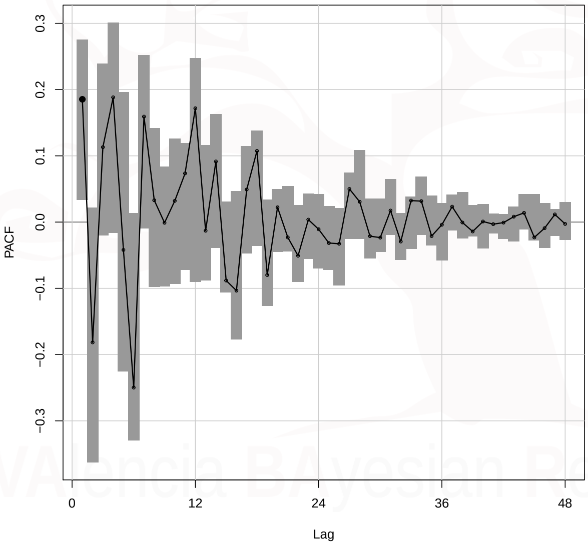
```r
Pacf(wineind)
```



```r
taperedacf(wineind, nsim = 50)
```



```r
taperedpacf(wineind, nsim = 50)
```



El proceso ruido blanco

Definición

El objetivo de las series temporales es descomponer la serie observada en una parte que es dependiente del pasado y otra que es impredecible:

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_1) + a_t$$

El proceso a_t se conoce como innovación y tiene la característica de ser **ruido blanco**:

- $E(a_t) = 0$
- $Var(a_t) = \sigma_a^2$
- $cov(a_t, a_{t+h}) = 0$

- El ruido blanco se puede interpretar como un elemento de innovación o sorpresa que vamos a incorporar en nuestro modelo.
- Si todas las series que observamos en la realidad fuesen ruido blanco serían impredecibles y no habría ningún modelo que proponer.

Teorema de Wold

Cualquier proceso estacionario y_t con media cero se puede representar de la forma

$$y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} + k_t$$

- El proceso ϵ_t es ruido blanco y representa el error resultante de predecir y_t con una función lineal de los retardos de y_t:

$$\epsilon_t = y_t - E(y_t | y_{t-1}, y_{t-2}, \ldots)$$

- A k_t se llama el componente linealmente determinístico.

Funciones de autocovarianzas y autocorrelaciones

- Funciones de autocorrelación miden la relación lineal entre variables aleatorias de procesos separadas de una cierta distancia en el tiempo.
- Estimación de estas funciones permiten determinar la forma del procesos estocástico.

- La función de autocovarianza

$$\gamma_{t,s} = Cov(x_t, x_{t+s}) = E[(x_t - \mu_t)(x_{t+s} - \mu_{t+s})]s = \ldots, -1, 0, 1, \ldots$$

- Si el proceso es estacionario, su esperanza es constante a largo del tiempo, y la función de autocovarianza no depende del momento en tiempo, sólo la distancia temporal.

$$\gamma_t = E\left[(x_t - \mu)(x_{t+s} - \mu)\right]$$

- Para cada retardo hay un valor diferente para la función de autocovarianzas, autocovarianza de orden s.
- Función de autocorrelación simple ACF (*autocorrelation function*)

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,s}\gamma_t}} = \frac{E[(x_t - \mu_t)(x_{t+s} - \mu_{t+s})]}{\sqrt{E[(x_t - \mu_t)]^2 E[(x_{t+s} - \mu_{t+s})]^2}}$$

Si el proceso es estacionario, los momentos de segundo orden no dependen de t.

$$\rho_s = \frac{\gamma_s}{\gamma_0} = \frac{E\left[(x_t - \mu)(x_{t+s} - \mu)\right]}{E\left[(x_t - \mu)\right]^2}$$

- Un correlograma muestra el ACF en función de s

Definimos el momento central de orden k como: $\mu^k = E\left[(x - E(x))^k\right]$.

Función de autocorrelación parcial

La función de autocorrelación parcial PACF (*partial autocorrelation function*) enseña la relación lineal cuando se ha eliminado la correlación que estas variables tienen con otras variables.

$$\phi_{kk} = Corr(x_t, x_{t-k} | x_{t-1}, \ldots, X_{t-k+1})$$

$$\phi_{kk} = \frac{Cov[(x_t - \hat{x}_t), (x_{t-k} - \hat{x}_{t-k})]}{\sqrt{Var(x_t - \hat{x}_t) Var(x_{t-k} - \hat{x}_{t-k})]}}$$

- Se pueden obtener los coeficientes de las ACF a traves de los modelos lineales:

$$
\begin{cases}
x_t = \phi_{11} x_{t-1} + \nu_t \\
x_t = \phi_{21} x_{t-1} + \phi_{22} x_{t-2} + \nu_t \\
\ldots \\
x_t = \phi_{k1} x_{t-1} + \phi_{k2} x_{t-2} + \ldots + \phi_{kk} x_{t-k} + \nu_t
\end{cases}
$$

Se puede demostrar que los coeficientes de FAS se pueden escribir como una función de coeficientes de PACF. Esta relación se llama el sistema de ecuaciones de Yule-Walker.

a (sistema d') **equacions de Yule-Walker**:[8]

$$\begin{cases} \rho_1 = \phi_{k1} + \phi_{k2}\rho_1 + \cdots + \phi_{kk}\rho_{k-1} \\ \rho_2 = \phi_{k1}\rho_1 + \phi_{k2} + \cdots + \phi_{kk}\rho_{k-2} \\ \cdots \\ \rho_k = \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \cdots + \phi_{k,k-1}\rho_1 + \phi_{kk} \end{cases} \tag{2}$$

per a $k = 1, 2, \ldots$ Aquest sistema d'equacions es pot escriure com:

$$\begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix},$$

del qual interessa conèixer únicament el valor de ϕ_{kk}. Aplicant la regla de Cramer s'obté aquest valor com a cocient de dos determinants:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

El primer coeficiente de la PACF coincidirá simepre con el del ACF.

Función de autocorrelación simple

$$\hat{\rho}_s = r_s = \frac{\gamma_s}{\gamma_0} = \frac{\sum_{t=1}^{T-s}(x_t - \bar{x})(x_{t+s} - \bar{x})}{\sum_{t=1}^{T}(x_t - \bar{x})^2}$$

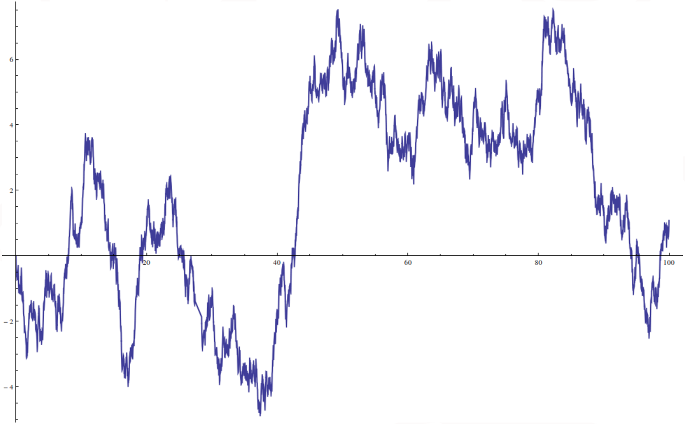
Random Walk

Procesos de paseo aleatorio

Un proceso estocástico sigue un paseo aleatorio si $x_t = x_{t-1} + \varepsilon_t$, donde ε_t es **ruido blanco**.

- El valor en un momento es el valor del periodo anterior más un efecto aleatorio ruido blanco.

$$x_t - x_{t-1} = \Delta x_t = (1 - L)x_t = \varepsilon_t$$


6
4
2
−2
−4
20
40
60
80
100

Random Walk con media

Se puede incorporar al modelo una cierta constante:

$$\Delta x_t = \delta + \varepsilon_t, \rightarrow x_t = x_{t-1} + \delta + \varepsilon_t$$

ACF y PACF con R

Atmospheric concentration of CO2

The record of the atmospheric concentration of CO2 collected at the Mauna Loa Observatory in Hawai

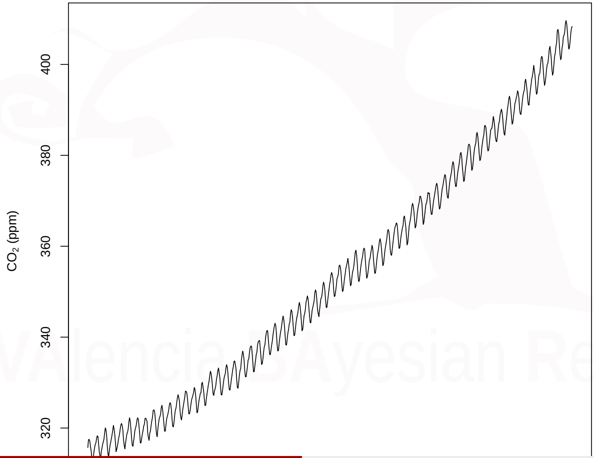
```
library(RCurl)
## get CO2 data from Mauna Loa observatory
ww1 <- "ftp://aftp.cmdl.noaa.gov/products/"
ww2 <- "trends/co2/co2_mm_mlo.txt"
CO2 <- read.table(text = getURL(paste0(ww1, ww2)))[, c(1, 2,
    5)]
## assign better column names
colnames(CO2) <- c("year", "month", "ppm")
```

Table 1: atmospheric concentration of CO2

year	month	ppm
1958	3	315.71
1958	4	317.45
1958	5	317.50

Creando el objeto 'Serie Temporal'

```
## create a time series (ts) object from the CO2 data
co2 <- ts(data=CO2$ppm, frequency=12,
              start=c(CO2[1,"year"],CO2[1,"month"]))
## plot the ts
plot.ts(co2, ylab=expression(paste("CO"[2]," (ppm)")))
```

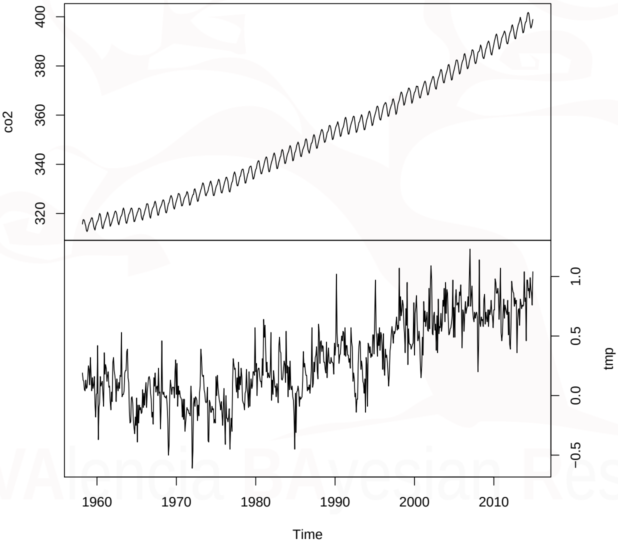


Combinado diferentes series

```
## get N Hemisphere land & ocean temperature anomalies from NOAA
Temp <- read_csv("https://www.ncdc.noaa.gov/cag/global/time-series
                  /nhem/land_ocean/p12/12/1880-2014.csv" , skip =
## create ts object
tmp <- ts(data=Temp$Value, frequency=12, start=c(1880,1))
```

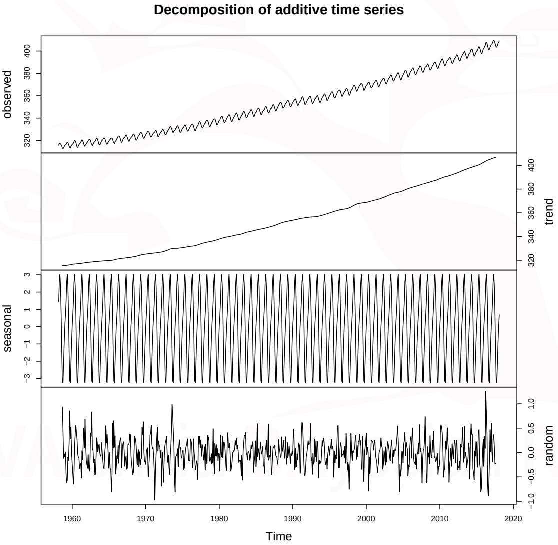
x
-0.01
-0.28
-0.30
-0.07
0.03
-0.06

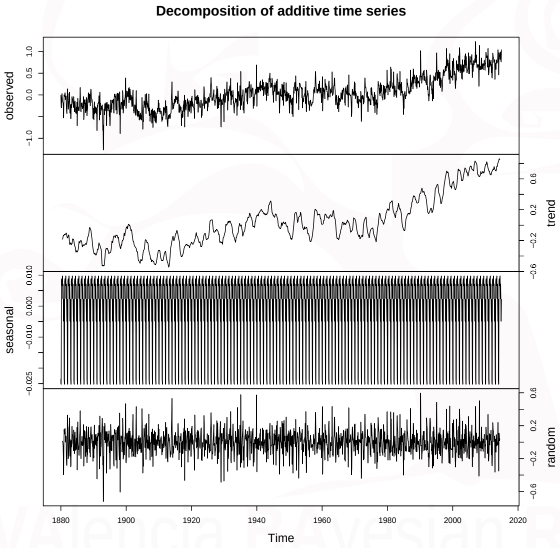
```
## [1] 682   2
```


co2
400
380
360
340
320
tmp
1.0
0.5
0.0
−0.5
1960
1970
1980
1990
2000
2010
Time

Descomponiendo las series

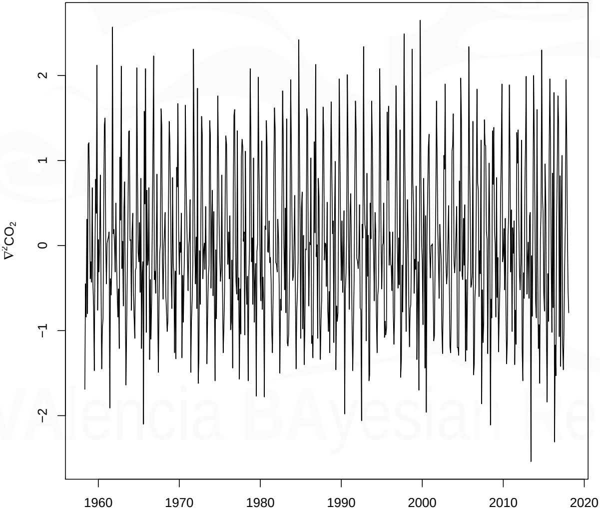
```r
co2.decomp <- decompose(co2)
plot(co2.decomp, yax.flip = TRUE)
```



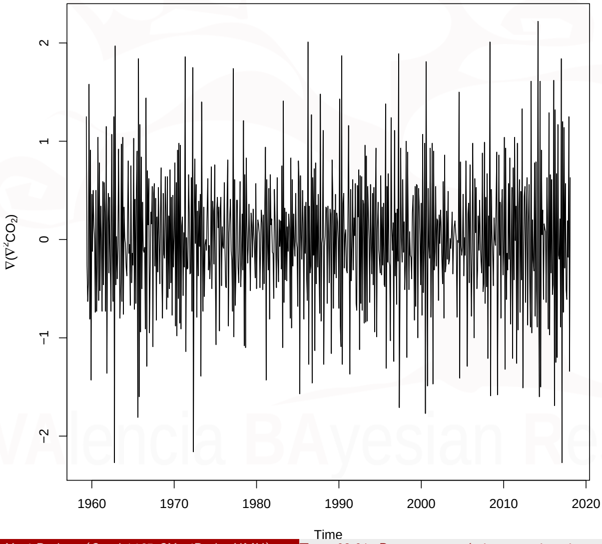

Decomposition of additive time series
observed
trend
seasonal
random
Time
1880
1900
1920
1940
1960
1980
2000
2020

Diferenciando la series con el comando *diff*

```r
## twice-difference the CO2 data
co2.D2 <- diff(co2, differences = 2)
## plot the differenced data
plot(co2.D2, ylab = expression(paste(nabla^2, "CO"[2])))
```

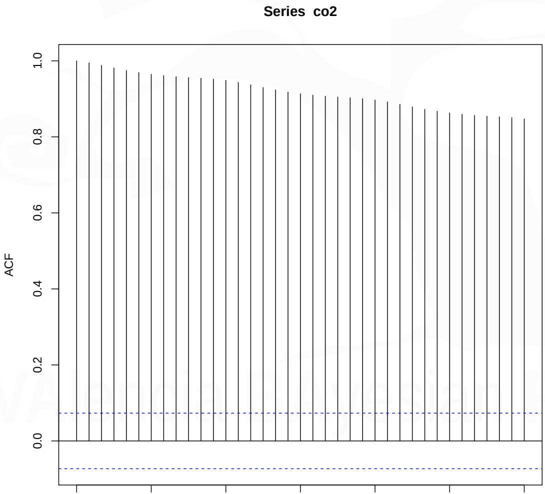


```
## difference the differenced CO2 data
co2.D2D12 <- diff(co2.D2, lag=12)
## plot the newly differenced data
plot(co2.D2D12,
     ylab=expression(paste(nabla,"(",nabla^2,"CO"[2],")")))
```

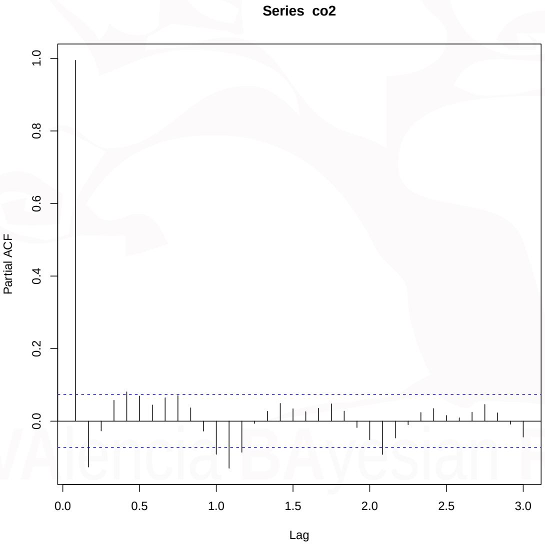


Correlogramas

```r
## correlogram of the CO2 data
acf(co2, lag.max = 36)
```



```
## correlogram of the CO2 data
pacf(co2, lag.max = 36)
```

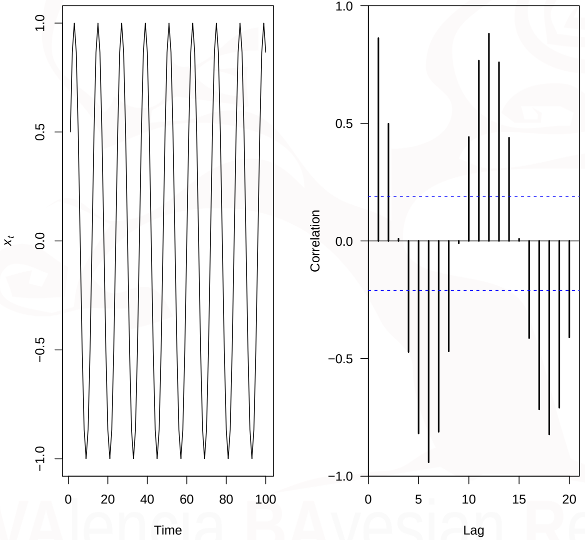


Ejemplos de correlogramas

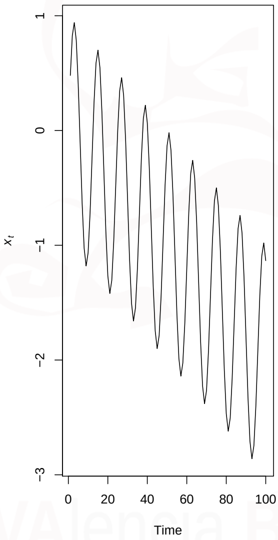
Creamos priemero una función más a nuestro gusto:

```
plot.acf <- function(ACFobj) {
    rr <- ACFobj$acf[-1]
    kk <- length(rr)
    nn <- ACFobj$n.used
    plot(seq(kk), rr, type = "h", lwd = 2, yaxs = "i", xaxs = "i",
        ylim = c(floor(min(rr)), 1), xlim = c(0, kk + 1), xlab = "Lag",
        ylab = "Correlation", las = 1)
    abline(h = -1/nn + c(-2, 2)/sqrt(nn), lty = "dashed", col = "blue")
    abline(h = 0)
}
```

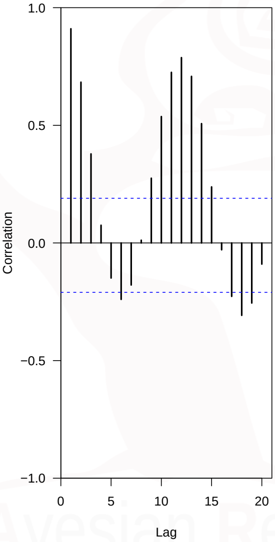
```r
## create sine wave
tt <- sin(2 * pi * seq(nn)/12)
## set up plot area
par(mfrow = c(1, 2))
## plot line
plot.ts(tt, ylab = expression(italic(x[t])))
## get ACF
sine.acf <- acf(tt, plot = FALSE)
## plot ACF
plot.acf(sine.acf)
```


1.0
0.5
0.0
−0.5
−1.0
x_t
0
20
40
60
80
100
Time
Correlation
0
5
10
15
20
Lag

```r
## create sine wave with trend
tt <- sin(2 * pi * seq(nn)/12) - seq(nn)/50
## set up plot area
par(mfrow = c(1, 2))
## plot line
plot.ts(tt, ylab = expression(italic(x[t])))
## get ACF
sili.acf <- acf(tt, plot = FALSE)
## plot ACF
plot.acf(sili.acf)
```



x_t
1
0
−1
−2
−3
0
20
40
60
80
100
Time



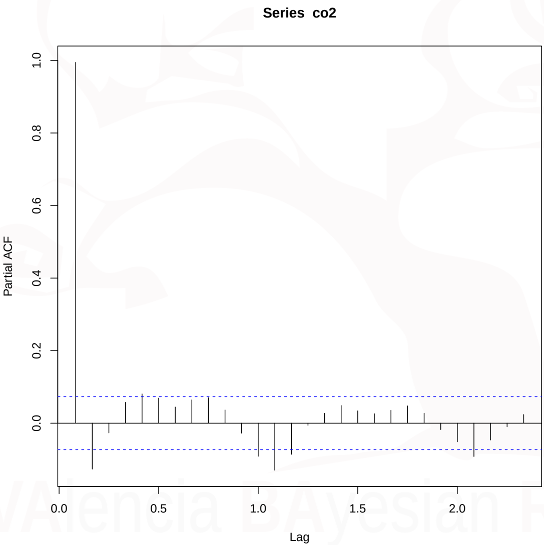
Correlation
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20
Lag

Autocorrelogramas parciales

Creemos primero una función retocada

```
plot.pacf <- function(PACFobj) {
    rr <- PACFobj$acf
    kk <- length(rr)
    nn <- PACFobj$n.used
    plot(seq(kk), rr, type = "h", lwd = 2, yaxs = "i", xaxs = "i",
        ylim = c(floor(min(rr)), 1), xlim = c(0, kk + 1), xlab = "Lag",
        ylab = "PACF", las = 1)
    abline(h = -1/nn + c(-2, 2)/sqrt(nn), lty = "dashed", col = "blue")
    abline(h = 0)
}
```

```
## PACF of the CO2 data
co2.pacf <- pacf(co2)
```




```
## correlogram of the CO2 data
plot.acf(co2.pacf)
```

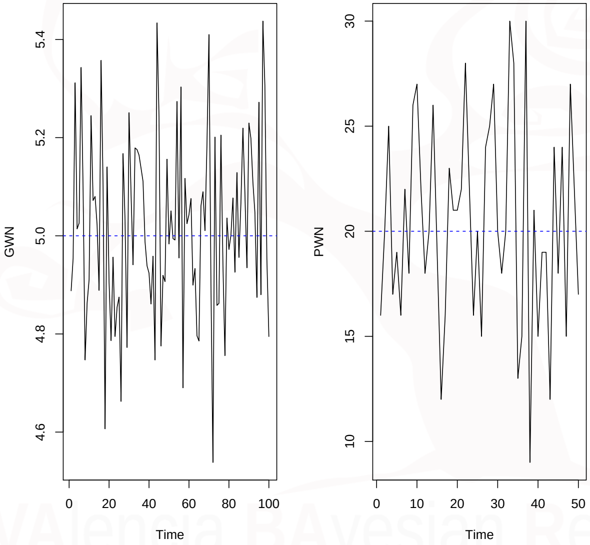
White Noise

Generamos un W-N.

```
set.seed(123)
## random normal variates
GWN <- rnorm(n = 100, mean = 5, sd = 0.2)
## random Poisson variates
PWN <- rpois(n = 50, lambda = 20)
```

Lo pintamos:

```
## set up plot region
par(mfrow = c(1, 2))
## plot normal variates with mean
plot.ts(GWN)
abline(h = 5, col = "blue", lty = "dashed")
## plot Poisson variates with mean
plot.ts(PWN)
abline(h = 20, col = "blue", lty = "dashed")
```


GWN
5.4
5.2
5.0
4.8
4.6
0
20
40
60
80
100
Time
PWN
30
25
20
15
10
0
10
20
30
40
50
Time

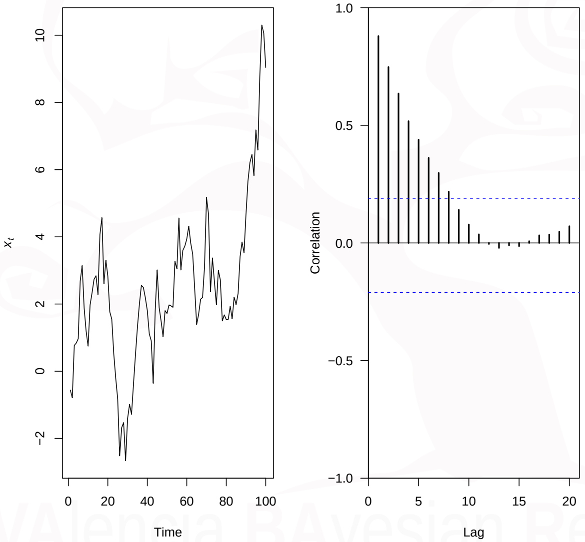
Y veamos sus autocorrelogramas

```r
## set up plot region
par(mfrow = c(1, 2))
## plot normal variates with mean
acf(GWN, main = "", lag.max = 20)
## plot Poisson variates with mean
acf(PWN, main = "", lag.max = 20)
```

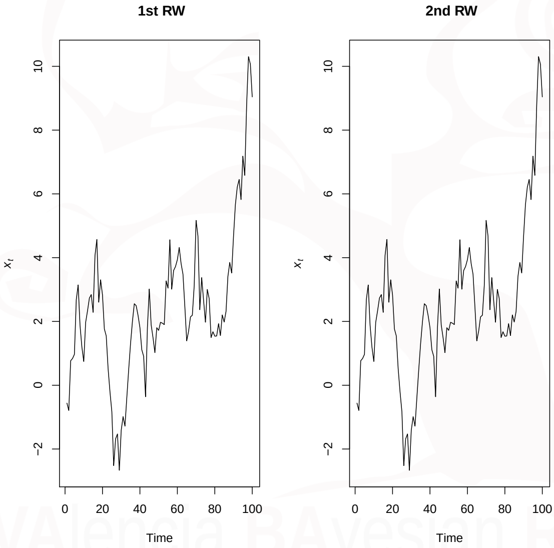
```r
## set up plot region
par(mfrow = c(1, 2))
## plot normal variates with mean
acf(GWN, main = "", lag.max = 20)
## plot Poisson variates with mean
acf(PWN, main = "", lag.max = 20)
```

Random Walk

```
## set random number seed
set.seed(123)
## length of time series
TT <- 100
## initialize {x_t} and {w_t}
xx <- ww <- rnorm(n = TT, mean = 0, sd = 1)
## compute values 2 thru TT
for (t in 2:TT) {
    xx[t] <- xx[t - 1] + ww[t]
}
```



x_t
Time
Correlation
Lag


1st RW
2nd RW
x_t
x_t
−2
0
2
4
6
8
10
−2
0
2
4
6
8
10
0
20
40
60
80
100
0
20
40
60
80
100
Time
Time