

Series de Tiempo 2018

Maestría en Estadística Aplicada, UNR
Unidad 0

Luis Damiano

damiano.luis@gmail.com

2018-05-04

- Curso **introductorio**.
- Conocimientos teóricos **básicos pero sólidos** que permitan un estudio independiente en temas más profundos.
- Foco en modelos probabilísticos (comparar contra modelos no probabilísticos).
- Visión aplicada con especial atención en las cuestiones metodológicas relacionadas con el análisis de datos.
- Familiaridad con las rutinas de R que más se emplean en el análisis de series de tiempo.
- ¿Interesados en modelos más complejos? Tesis de Maestría!

¿Qué temas veremos?

- Procesos estacionarios (autoregresivos y de media móvil).
- Procesos no estacionarios.
- Transformaciones.
- Estacionalidad.
- Identificación, estimación, diagnóstico y selección de modelos
- Pronósticos probabilísticos.
- Valores atípicos.
- Breve mención a modelos de espacio de estado.
- Breve mención a modelos de suavizados.
- Lenguaje de Programación R !!!

¿Qué temas no veremos?

- Enfoque de dominio.
- Datos con períodos de tiempo irregulares.
- Datos discretos, o continuos con cotas.
- Datos faltantes.
- Observaciones multivariadas.
- Heterocedasticidad condicional.
- Procesos de memoria larga.
- Modelos de tiempo continuo.
- Modelos de suavizados como espacio de estados.
- Modelos no paramétricos.
- Modelos no lineales.
- Modelos Bayesianos :(

Prácticos:

- Shumway and Stoffer (2017) Capítulos 1-3, 6*.
- Metcalfe and Cowpertwait (2009) Capítulos 1-3, 6-7, 12*.
- Cryer and Chan (2008) Capítulos 1-10.
- Tsay (2010) Capítulos 1-2, 11*.
- Hyndman and Athanasopoulos (2018) Capítulos 1-3, 6, 8.
- Peña (2010) Capítulos 1-13.
- Campagnoli, Petrone, and Petris (2009) Capítulos 2-3*.

Teóricos:

- Hamilton (1994) Capítulos 1-5, 15, 17.
- Brockwell and Davis (2016) Capítulos 1-3, 5-6, 8*.

* Mención a los modelos de espacio de estado.

R es un gran ecosistema para el análisis de series de tiempo.

- Gran variedad de paquetes pensados para leer, manipular, graficar, y modelar series de tiempo.
 - Vista de Time Series en CRAN.
 - Econometría.
 - Finanzas.

Algunas tareas básicas:

- ¿Cómo leer una serie de tiempo desde diferentes formatos de archivos?
- ¿Cómo representar una serie de tiempo?
- ¿Cómo manipular una serie de tiempo?

R: Lectura de datos

Desde un archivo de texto plano

```
# http://bit.ly/2tVI81B
```

```
df <- read.table(  
  file = "data//BCRABadlar.txt",  
  header = TRUE,  
  sep = "\t",  
  dec = ",",  
)  
  
head(df)
```

```
##      FECHA  VALOR  
## 1 14/03/2018 23.0625  
## 2 13/03/2018 23.0000  
## 3 12/03/2018 22.5625  
## 4 09/03/2018 22.7500  
## 5 08/03/2018 22.8125  
## 6 07/03/2018 22.6250
```

Desde una planilla de cálculo *xlsx*

```
# http://bit.ly/2IzBrp5
```

```
library(openxlsx)  
  
df <- read.xlsx(  
  xlsxFFile = "data//INDEXProyecciones.xlsx",  
  sheet = "Sheet1",  
  rows = c(6, 9:39)  
)  
  
head(df)
```

```
##      X1    Total Varones  Mujeres  
## 1 2010 40788453 19940704 20847749  
## 2 2011 41261490 20180791 21080699  
## 3 2012 41733271 20420391 21312880  
## 4 2013 42202935 20659037 21543898  
## 5 2014 42669500 20896203 21773297  
## 6 2015 43131966 21131346 22000620
```

R: Representación de los datos

- Los datos suelen leídos en un objeto `data.frame`.
 - Los timestamps suelen importarse como cadenas de caracteres. No son una verdadera representación del tiempo.
 - No trae herramientas pensadas para el análisis de series de tiempo.
- Representando una serie de tiempos como un objeto `ts`.
 - Indexa las observaciones en unidades discretas y equiespaciadas de tiempo.
 - Requiere definir la frecuencia como cantidad de observaciones por unidad de tiempo.
 - No emplea fechas.
- Representando una serie de tiempos como un objeto `xts`.
 - Basado en `zoo`.
 - Requiere primero convertir timestamps en formato tiempo.
 - Admite series con datos no equi-espaciados.
 - Trae herramientas gráficas diseñadas especialmente para series de tiempo.
 - Trae operadores típicos para series de tiempo (ej. `lag`, `diff`).
 - Trae funciones para manipular datos secuenciales (ej. `rollsum`, `rollapply`)
 - Permite manipular datos basados en el timestamp (ej. `x[2017/2018]`, `align.time`, `apply.weely`).
 - Permite unir conjuntos de datos basados en el timestamp (ej. `cbind`, `merge`).

R: Manipulación de datos (1)

Leer datos y convertir en formato ts.

```
# http://bit.ly/2tVI81B

df <- read.table(
  file = "data/BCRABadlar.txt",
  header = TRUE,
  sep = "\t",
  dec = ",",
)

z <- ts(data = df[, 2], frequency = 1)

head(z)

## Time Series:
## Start = 1
## End = 6
## Frequency = 1
## [1] 23.0625 23.0000 22.5625 22.7500 22.8125 22.6250
```


R: Manipulación de datos (2)

Leer datos y convertir en formato xts.

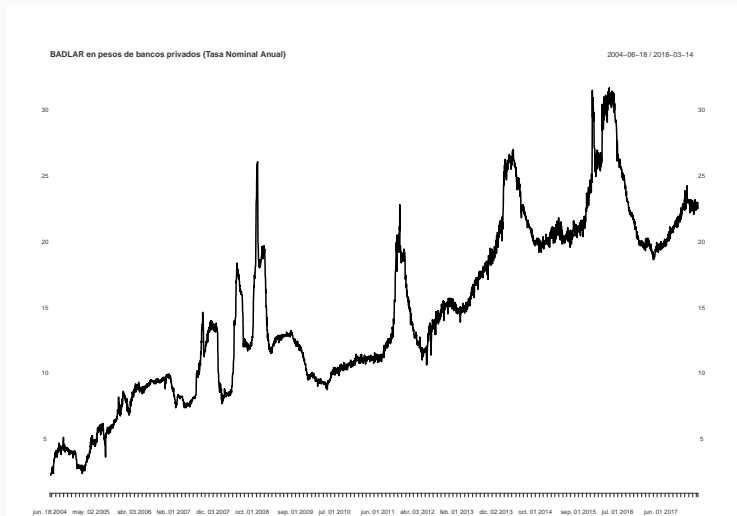
```
# http://bit.ly/2tVI81B
```

```
df <- read.table(  
  file = "data/BCRABadlar.txt",  
  header = TRUE,  
  sep = "\t",  
  dec = ",",  
)  
  
df[, 1] <- as.Date(df[, 1], format = "%d/%m/%Y")  
  
z <- xts(x = df[, 2], order.by = df[, 1])  
  
head(z)
```

```
##           [,1]  
## 2004-06-18 2.3125  
## 2004-06-22 2.3125  
## 2004-06-23 2.3125  
## 2004-06-24 2.3750  
## 2004-06-25 2.3750  
## 2004-06-28 2.4375
```

R: Manipulación de datos (3)

Graficar series de tiempo automáticamente!



R: Manipulación de datos (4)

Manipulaciones típicas con datos secuenciales y/o indexados por el tiempo.

```
# Valor de hoy, valor del período anterior,  
# valor del quinto paso anterior,  
# valor del período siguiente,  
# primera diferencia
```

```
head(  
  cbind(  
    z, lag(z), lag(z, k = 5),  
    lag(z, k = -1), diff(z)  
  )  
)
```

```
##           ..1    ..2    ..3    ..4    ..5  
## 2004-06-18 2.3125    NA    NA 2.3125    NA  
## 2004-06-22 2.3125 2.3125    NA 2.3125 0.0000  
## 2004-06-23 2.3125 2.3125    NA 2.3750 0.0000  
## 2004-06-24 2.3750 2.3125    NA 2.3750 0.0625  
## 2004-06-25 2.3750 2.3750    NA 2.4375 0.0000  
## 2004-06-28 2.4375 2.3750 2.3125 2.5000 0.0625
```

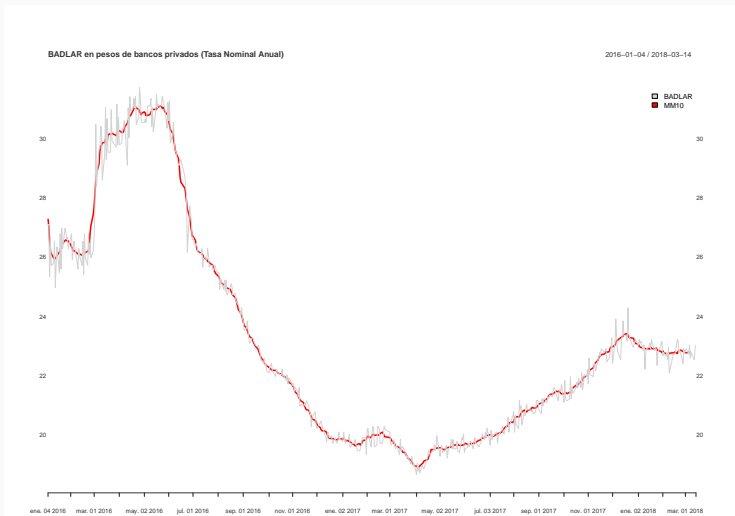
```
# Promedio por semana para el último año
```

```
head(  
  apply.weekly(z["2018"], mean)  
)
```

```
##           [,1]  
## 2018-01-05 23.15625  
## 2018-01-12 22.78750  
## 2018-01-19 23.08750  
## 2018-01-26 22.78750  
## 2018-02-02 22.96250  
## 2018-02-09 22.63750
```

R: Manipulación de datos (5)

Procesar datos usando una ventana móvil (*rolling window*).



¿Qué muestra este gráfico?

- T Tamaño de la muestra
- t Índice de tiempo (discreto)
- Z_t Observaciones
- a_t Error aleatorio
- μ Función de media
- σ^2 Función de varianza
- γ_k Función de autocovarianza (función de k)
- ρ_k Función de autocorrelación (función de k)
- Φ_{kk} Función de autocorrelación parcial (función de k)
- ϕ_k Coeficiente autorregresivo para el k -ésimo rezago
- θ_k Coeficiente de media móvil para el k -ésimo rezago
- k Rezago medido en cantidad de pasos
- h Horizonte medido en pasos hacia adelante
- s Subíndice de estacionalidad
- p Orden del proceso autorregresivo
- q Orden del proceso de media móvil

Notación específica para modelos de espacio de estados

- p Observaciones
- m Estados
- g Dimensión del error aleatorio
- y_t Vector $p \times 1$ de observaciones
- α_t Vector $m \times 1$ de estado (latente/no observado)
- Z_t Matriz $p \times m$ de salida
- T_t Matriz $m \times m$ de transición de estados
- R_t Matriz $m \times g$ de selección
- η_t Vector $g \times 1$ de error aleatorio
- H_t varianza del error de medida
- Q_t Matriz $g \times g$ de covariancias del error de transición

- Comandos básicos de R.
- Algunas funciones que emplearemos durante el curso:
 - base: ts, tsdiag, acf, pacf, arima, arima.sim, head, tail, fitted, residuals, coef, decompose.
 - xts: xts, plot, first, last, apply.*, rollapply, roll*.
 - forecast: Acf, Pacf, seasonplot, tsdisplay, Arima, ets, forecast.

Referencias

Brockwell, Peter J., and Richard A. Davis. 2016. *Introduction to Time Series and Forecasting*. Springer International Publishing. doi:[10.1007/978-3-319-29854-2](https://doi.org/10.1007/978-3-319-29854-2).

Campagnoli, Patrizia, Sonia Petrone, and Giovanni Petris. 2009. *Dynamic Linear Models with R*. Springer New York. doi:[10.1007/b135794](https://doi.org/10.1007/b135794).

Cryer, Jonathan D., and Kung-Sik Chan. 2008. *Time Series Analysis*. Springer New York. doi:[10.1007/978-0-387-75959-3](https://doi.org/10.1007/978-0-387-75959-3).

Hamilton, James D. 1994. *Time Series Analysis*. Princeton University Press.

Hyndman, Rob J, and George Athanasopoulos. 2018. *Forecasting: Principles and Practice*. <https://otexts.org/fpp2/>.

Metcalfe, Andrew V., and Paul S.P. Cowpertwait. 2009. *Introductory Time Series with R*. Springer New York. doi:[10.1007/978-0-387-88698-5](https://doi.org/10.1007/978-0-387-88698-5).

Peña, Daniel. 2010. *Análisis de Series Temporales*. Alianza.

Shumway, Robert H., and David S. Stoffer. 2017. *Time Series Analysis and Its Applications*. Springer International Publishing. doi:[10.1007/978-3-319-52452-8](https://doi.org/10.1007/978-3-319-52452-8).

Tsay, Ruey S. 2010. *Analysis of Financial Time Series*. John Wiley & Sons, Inc. doi:[10.1002/9780470644560](https://doi.org/10.1002/9780470644560).