

Trabajo No 1 de Técnicas de Pronósticos

Pronósticos de Series de Tiempo con modelos para la Tendencia y la Estacionalidad

Septiembre de 2022

Características del Trabajo

El trabajo para cada grupo consiste de 4 puntos enunciados en la sección siguiente. Los puntos se deben resolver con la serie asignada en Tabla 2. Las series se describen en la última sección y se indica en dónde se encuentran los datos. Tiene un valor de 33% de la nota definitiva.

- **Presentación.** Elaborar el reporte en formato Word o Latex. El documento final en pdf. Con el nombre de los integrantes del grupo y el número del grupo en el encabezado.
- **Detalles.** Desarrollar cada punto en secciones separadas. Las gráficas y las tablas deben contener una numeración y descripción resumida, que permita identificar rápidamente de qué se trata.
- **Entrega.** Cargar el informe pdf en el Moodle antes de la fecha de entrega. Identificar el archivo pdf con los apellidos de los autores. Tiene un plazo de entrega de 1 semana (en días hábiles) a partir de la fecha de publicación en Moodle.

Puntos del Trabajo

Defina la estrategia de validación cruzada escogiendo una parte de los datos para ajustar y otra para comparar los pronósticos. Sugerencia: en series mensuales use 12 meses, en trimestrales use 8 trimestres. En frecuencia diaria use 15 días.

1. (25/25) Estime el modelo No1 asignado en la Tabla 1. Reporte las tablas de parametros estimados, estadísticos t y F, valores p. Reporte la gráfica de los valores ajustados \hat{Y}_t , versus los datos observados Y_t , para el período de entrenamiento.
2. (25/25) Estime el modelo No2 asignado en la Tabla 1. Reporte la información según el procedimiento de estimación aplicado: parámetros, varianzas del modelo, tipo de red NNAR. Reporte la gráfica de los valores ajustados \hat{Y}_t , versus los datos observados Y_t , para el período de entrenamiento.
3. (25/25) Reporte MSE, AIC, BIC y R-cuadrado ajustado para los 2 modelos No1 y No2. Concluya cuál modelo ajustó mejor.
4. (25/25) Calcule los pronósticos para la validación cruzada con los 2 modelos No1 y No2. Reporte la gráfica de los pronósticos y de los datos observados, para el período de comparación. Reporte MAPE, RMSE, U-Theil para éstos. Concluya cuál modelo pronostica mejor.

No	Modelo
1	Exponencial lineal + indicadoras, Cap 3
2	Exponencial cuadrático + indicadoras, Cap 3
3	Descomposición STL, §5.1.2, pag. 97
4	Prophet §5.1.3, pag. 99
5	Holt-Winters espacio de estados §5.2.1, pag. 103
6	Holt-Winters componentes §5.2.2, pag. 105
7	Holt-Winters amortiguado §5.2.5, pag. 108
8	Red neuronal autoregresiva NNAR §5.4, pag. 115

Cuadro 1: Modelos para Series de Tiempo

Cuadro 2: Asignación de modelos y serie por grupo

Serie No	mod 1	mod 2	Estudiantes
1	1	5	Brandon López Avendaño – Carlos Andrés Duque Granda
2	1	6	Camilo Andrés Cabrera Meneses – John Mario Montoya Zapata
3	1	7	Melisa Tuiran Ruiz – Maria Rosales Silva – Daniel Ceballos González
4	1	8	Andrés Felipe Herrera Gómez – Estefany López Toro
5	2	5	Beatriz Cecilia Usta Díaz – Diana Catalina Peña Vasquez
6	2	6	Yuliany Rojas Londoño – Angel Mateo Ramírez Londoño
7	2	7	Jhair Santiago Martinez Osorio – Catalina Londoño Ceballos
8	2	8	Clara Mercedes Niño Perea – Laura Natalia Mesa Ruiz
9	3	5	José Luis Hoyos Cuadrado – Carolina González Pérez
10	3	6	Sara Cardona Gómez – Danilo Marcelo Osorio Copete
11	3	7	Maria Alejandra Patiño Restrepo – Jhon Jairo Gómez Cárdenas
12	3	8	Esteban Palacio Pérez
13	4	5	Brayan Valdez Arango
14	4	6	Luis Martínez Escobar
15	4	7	Pablo Andrés Jarava Guerra
16	4	8	Santiago Montaña Buitrago
17	3	7	Adiel Restrepo Gil
18	4	5	Maria Camila Arango Home
19	4	7	Estiven Garda Luna
20	3	7	Kevin Bravo Bolaño

Series para el Trabajo

La serie asignada a cada integrante está en la lista siguiente. Los archivos están en Moodle en la carpeta del tema del trabajo ó se pueden cargar desde las librerías indicadas en cada serie.

(1) producción vino tinto

Instrucciones de lectura. Archivo en Moodle

Wine sales

#Australia, monthly

#Jan 1980-Jun 1994

#Total --- Total Wine sales (th. of litres)

```
#by wine makers in bottles <= 1L.
#ABS (Australian Bureau of Statistics)
# Time Series Data Library
Vino = read.table("Wine.dat", header = TRUE, stringsAsFactors = FALSE)
attach(Vino)
y = ts(Vino$Red,frequency=12,start=c(1980,01))
```

(2) producción cerveza Australia,

en la librería fpp, se describe como:
 Producción total de cerveza en Australia: ‘‘
 Total quarterly beer production in Australia
 (in megalitres) from 1956:Q1 to 2008:Q3.’’.
 Se carga con las instrucciones

```
require(fpp)
y = ts(ausbeer,frequency=4,start=c(1956,01))
ts.plot(y)
```

(3) Demanda diaria energía eléctrica en Bogotá,

Empresa Codensa S.A., en Gwh,
 entre 1995-07-20 y 1997-10-23.
 Archivo en Moodle

```
S = read.table("codensa.diaria.dat",
header = TRUE, stringsAsFactors = FALSE)
```

(4) precio de carne pavo

turkey.price.ts: Monthly Average Turkey Price, January 2001 to April 2008
 his time series shows the average retail price of turkey in the United States bet

```
S = read.table("turkey.price.dat",
header = TRUE, stringsAsFactors = FALSE)
attach(S)
```

(5) ventas antidiabético

en la librería fpp, se describe como
 ‘‘Monthly anti-diabetic drug sales
 in Australia from 1992 to 2008’’.

```
require(fpp)
```

```

data(a10)
y = ts(a10,frequency=12,start=c(1992,01))
# generar un vector de fechas, clase 'Date'
fechas = seq(as.Date("1992/1/1"),
length.out = length(y), by = "months")
plot(fechas,y,type='l')

```

(6) Empleo hombres Suiza.

```

Time series from http://www.dataserie.org,
a comprehensive and up-to-date
collection of open data from Switzerland
require(dataseries)
D=dataseries::ds("ch_comb_jobs.596.tot.1.0")
y = D$ch_comb_jobs.596.tot.1.0
names(y) = "empleo total hombres trimestral Suiza"
fechas = D$time
y=ts(y,frequency=4)
ts.plot(y)

```

(7) La serie h02,

```

en la librería fpp,
se describe como: ‘‘Monthly cortecosteroid
drug sales in Australia from 1992 to 2008’’.
Se carga con las instrucciones
require(fpp)
y = h02
# generar un vector de fechas, clase 'Date'
fechas = seq(as.Date("1992/1/1"),
length.out = length(y), by = "months")

```

(8) serie turismo

```

# descripcion
The Tcomp package provides data from the tourism
forecasting competition described in George Athanasopolous,
Rob J. Hyndman, Haiyan Song, Doris C. Wu (2011)
‘‘The tourism forecasting competition’’,
International Journal of Forecasting 27 (2011) 822-844.

```

```
# instalacion
require(Tcomp)
data(package='Tcomp')
# lectura
y1 = tourism$M82$x
ts.plot(y,type='b')
```

(9) precio carne hamburguesa

```
A time series objects consiting of average monthly
retail prices per pound of ham in the United States
between January 2001 and April 2008
S = read.table("ham.price.dat",
header = TRUE, stringsAsFactors = FALSE)
attach(S)
```

(10) turismo Chipre

```
# leer turismo chipre: 1 serie
tu.chipre = read.table("turismo.chipre.dat",
header = TRUE, stringsAsFactors = FALSE)
y = tu.chipre$y
# generar un vector de fechas, clase 'Date'
fechas = seq(as.Date("1979/1/1"),
length.out = length(y), by = "month")
```

(11) Consumo mensual de papel en periódicos EUA

```
Series: M649
Type of series: INDUSTRY
Period of series: MONTHLY
Series description: Newsprint , United States, consumption by publishers

library(Mcomp)
M3[[2050]]
yi = ts(M3[[2050]]$x,frequency=12,star=c(1983,1),end=c(1992,06))
yf = ts(M3[[2050]]$xx,frequency=12,star=c(1992,07),end=c(1994,01))
m = 18
```

(12) generacion energía eléctrica

en la librería fpp, se describe como:

```
'' Electricity monthly total net
generation. January 1973-October 2010.''
require(fpp)
data(usmelec)
y = ts(usmelec,frequency=12, start=c(1973,01)
# generar un vector de fechas, clase 'Date'
fechas = seq(as.Date("1973/1/1"),
length.out = length(y), by = "months")
plot(fechas,y,type='l')
```

(13) producción total vino

```
# Instrucciones de lectura. Archivo en Moodle
Wine sales
#Australia, monthly
#Jan 1980-Jun 1994
#Total --- Total Wine sales (th. of litres)
#by wine makers in bottles <= 1L.
#ABS (Australian Bureau of Statistics)
# Time Series Data Library
Vino = read.table("Wine.dat", header = TRUE, stringsAsFactors = FALSE)
attach(Vino)
y = ts(Vino$Total,frequency=12,start=c(1980,01))
```

(14) accidentes fatales auto en UK,

```
# leer datos de la serie Seatbelts:
UKDriverDeaths is a time series giving the monthly
totals of car drivers in Great Britain killed or
seriously injured Jan 1969 to Dec 1984.
Compulsory wearing of seat belts was
introduced on 31 Jan 1983.
data(Seatbelts)
# car drivers killed.
y = Seatbelts[,2]
ts.plot(y)
```

(15) La serie IPC,

```
Variación porcentual mensual del Índice de precios
al consumidor, 2000-2012.
library(readxl)
res <- read_excel("IPCporcentual.xlsx", 1) # lee el primer libro
attach(res)
y = ts(IPC, frequency=12,start=c(2000,01))
ts.plot(y)
```

(16) Producción industrial España

```
# descripcion
Time-Series [1:61] from 1978 to 1983:
Brockwell and Davis (1991, Series E, p. 556)"
"Industrial production, Spain: monthly"
# instalacion
devtools::install_github("FinYang/tsdl")
require(tsdl)
# lectura
D <- subset(tsdl,12,"Production")
y = ts(D[[17]],frequency=12,start=c(1978,01))
par(mfrow=c(1,1))
t= seq(1,length(y))
ts.plot(y,type='b')
```

(17) Captación mensual Cía de Crédito.

```
#-----datos en el archivo
Serie de captación mensual de dinero de una empresa comercial de créditos,
entre 1991-01-01 y 2003-02-01. El trabajo es sobre
la variación porcentual calculada como rendimiento logarítmico.
G = read.table("captacion promedio mensual.dat", header = TRUE)
# la serie original es el acumulado mensual
cy = ts(G$y,frequency=12,start=c(1995,1))
# la serie del trabajo es la variacion porcentual mensual
y = diff(log(cy),1,1)
y = ts(y,frequency=12,start=c(1995,2))
```

(18) La serie PIB Colombia,

PIB Trimestral Colombia, Sin desestacionalizar


```

a precios constantes 2005.
G = read.table("PIBtrimestral.dat", header = TRUE,
stringsAsFactors=FALSE)
attach(G)
fechas = seq(as.Date("2000/1/1"),
length.out = length(pib), by = "quarter")
y = ts(pib,frequency=4,start=c(2000,01))

```

(19) turismo Australia

```

# descripcion
Quarterly visitor nights (in millions) spent by
international tourists to Australia. 1999-2015
# instalacion
require(fpp2)
# lectura
data(austourists )
y = ts(austourists,frequency=4,start=c(1999,01))
ts.plot(y,type='b')

```

(20) Turistas Suiza.

```

Time series from http://www.dataserries.org,
a comprehensive and up-to-date
collection of open data from Switzerland
require(dataseries)
D=dataseries::ds("ch_comb_jobs.596.tot.1.0")
y = D$ch_comb_jobs.596.tot.1.0
names(y) = "empleo total hombres trimestral Suiza"
fechas = D$time
y=ts(y,frequency=4)
ts.plot(y)

```