

Tarea 01

Modelos de Supervivencia y Series de Tiempo

Aaron Mauricio Gómez Carlos Brito Mejía Diego Luna
Francisca Vilca Sánchez

Usted invierte dinero en la compra de una acción. Para saber qué tan viable es su inversión, ha decidido hacer un análisis de serie de tiempo. Este tipo de análisis se utiliza para examinar los datos históricos y predecir futuros cambios en el valor de una inversión. Antes de invertir, es importante informarse y evaluar su disposición al riesgo y objetivos, así como hacer su perfil de inversionista. También es fundamental conocer los fundamentos de las inversiones y la rentabilidad esperada. Además, existen diversas metodologías que pueden ayudar a desarrollar estrategias efectivas para la inversión.

Para su tarea busquen en Yahoo finanzas una acción que les llame su atención para su análisis. Yahoo finanzas es una de las principales páginas en la que usted podrá consultar información financiera, en ella además le permitirá ver gráficamente los datos y lo más importante descargar los datos en formato csv. Descargue los datos históricos del **1 de enero de 2001** al **30 Junio de 2020**, y en frecuencia indique periodicidad mensual, guarde estos datos en su carpeta de trabajo e impórtelos en R.

Con base a dichos valores realice:

1. Descargue e importe los datos en R y con ellos convierta el vector de precios a un objeto de series de tiempo y vea el comportamiento de sus datos, es decir calcule la media, moda, cuartiles, máximos, mínimos y varianza. Haga una gráfica de caja para ver visualmente estos resultados.

Nuestro interés está en ver la posibilidad de invertir en las acciones de [CEMEX](#) una empresa ligada al ámbito de la construcción que se autodenomina ser líder en la venta y producción de Cemento y otros materiales. Para saber si es una buena opción hacer esta inversión revisaremos los datos históricos desde el 1 de enero de 2001 al 30 de junio de 2020, usando algunos comandos y librerías de R extraeremos los datos desde [Yahoo! Finanzas](#), tal como se muestra en el siguiente código:

```
getSymbols("CX", src = "yahoo", from = "2001-01-01", to = "2020-06-30",  
           periodicity= "monthly")  
base <- CX[,6] # para solo usar la columna de los precios de cierre ajustados  
ts_base <- ts(base, start=2001, frequency = 12) # creamos el objeto de series de tiempo
```

Una vez que tenemos la base y el objeto series de tiempo comenzamos a hacer un análisis del comportamiento de los datos obteniendo algunos datos interesantes, que se pueden ver resumidos en la Tabla 1:

	Valores de la acción
Media	9.703
Moda*	7.463
Mínimo	2.120
Cuartil 0.25	6.322
Mediana	7.961
Cuartil 0.75	10.284
Máximo	27.697
Varianza	31.298

Note:

* = Es el valor de la locación estimada de la moda

Tabla 1: Resumen del comportamiento de los datos

De esto se puede notar que AÑADIR ANÁLISIS

Además a través de un gráfico de caja se puede ver el resumen de estas medidas, tal como se muestra en la Figura 1:

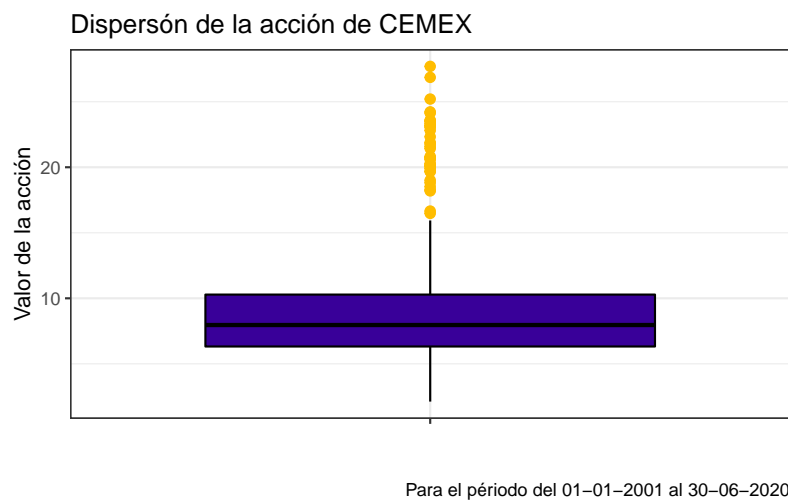


Figura 1: Boxplot de la acción de CEMEX

- Haga un análisis de sus datos, determinar si hay tendencias crecientes, decrecientes y periodicidad.

ESCRIBIR ARGUMENTOS

Enero	Febrero	Marzo	Abril
Min. : 4.040	Min. : 3.280	Min. : 2.120	Min. : 2.120
1st Qu.: 5.845	1st Qu.: 5.887	1st Qu.: 6.270	1st Qu.: 6.544
Median : 7.865	Median : 7.393	Median : 8.133	Median : 8.760
Mean : 9.653	Mean : 9.427	Mean : 9.545	Mean : 9.872
3rd Qu.: 9.554	3rd Qu.: 9.819	3rd Qu.: 10.088	3rd Qu.: 9.980
Max. : 25.204	Max. : 24.228	Max. : 23.330	Max. : 23.595

Mayo	Junio	Julio	Agosto
Min. : 2.400	Min. : 2.880	Min. : 0.000	Min. : 0.000
1st Qu.: 6.543	1st Qu.: 6.647	1st Qu.: 7.305	1st Qu.: 6.760
Median : 8.301	Median : 8.080	Median : 7.728	Median : 8.011
Mean : 9.882	Mean : 9.917	Mean : 9.556	Mean : 9.561
3rd Qu.:10.095	3rd Qu.:10.294	3rd Qu.:10.034	3rd Qu.:10.667
Max. :27.697	Max. :26.860	Max. :23.541	Max. :23.504
Septiembre	Octubre	Noviembre	Diciembre
Min. : 0.000	Min. : 0.000	Min. : 0.000	Min. : 0.000
1st Qu.: 6.457	1st Qu.: 5.734	1st Qu.: 5.612	1st Qu.: 6.288
Median : 7.837	Median : 7.581	Median : 7.448	Median : 7.810
Mean : 9.259	Mean : 8.751	Mean : 8.958	Mean : 9.140
3rd Qu.:10.414	3rd Qu.: 9.146	3rd Qu.: 9.641	3rd Qu.: 9.406
Max. :21.779	Max. :22.325	Max. :23.181	Max. :24.135

3. Realice una gráfica de a serie de tiempo usando las metodologías vistas en clases (ggplot2,plot y xts)

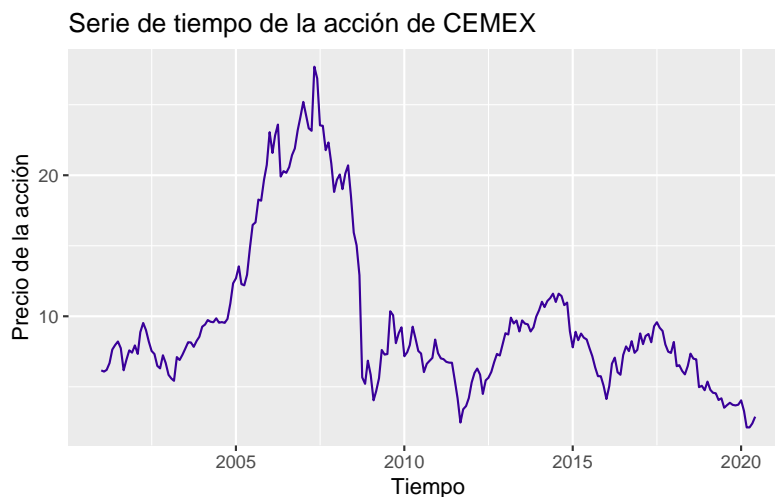


Figura 2: Serie de Tiempo de la acción de CEMEX

4. Realiza las siguientes suavizaciones:

- Suavización de promedios móviles simple (PS) de orden 2 y 10.

```
p.mov.2 = ma(ts_base, order = 2)
p.mov.10 = ma(ts_base, order = 10)
```

- Suavización de promedios móviles ponderados (PPM) de orden 4 con los pesos (0.5,0.2,0.2,0.1) para $(x_t, x_{t-1}, x_{t-2}, x_{t-3})$ y otra modelo (PPM) de orden 12 seleccionando ustedes mismos los pesos y mencionado el porque de esos pesos.

```

pesos.b1 <- c(0.5,0.2,0.2,0.1)
p.mov.pond.4 <- stats::filter(ts_base, pesos.b1, sides=1)

ts <- c(ts_base, rep(0,6))
ts_matrix <- matrix(ts, ncol = 12, byrow = T)
suma <- sum(colMeans(ts_matrix))
pesos.b2 <- round(colMeans(ts_matrix)/suma,4)

p.mov.pond.12 <- stats::filter(ts_base, pesos.b2, sides=1)

```

- Suavización de promedios móviles simples centrados (PSC) de orden 2 y 10.

```

p.mov.2.c = ma(ts_base, order = 2, centre = TRUE)
p.mov.10.c = ma(ts_base, order = 10, centre = TRUE)

```

- Suavización exponencial simple con $\alpha = 0.01, \alpha = 0.5, \alpha = 0.99$, ¿Cuál es el papel de α en la estimación de valores pronostico?

```

ts.base.sua_001 = ses(ts_base, alpha=0.01, initial = 'simple')
ts.base.sua_05 = ses(ts_base, alpha=0.5, initial = 'simple')
ts.base.sua_099 = ses(ts_base, alpha=0.99, initial = 'simple')

```

ESCRIBIR RESPUESTA

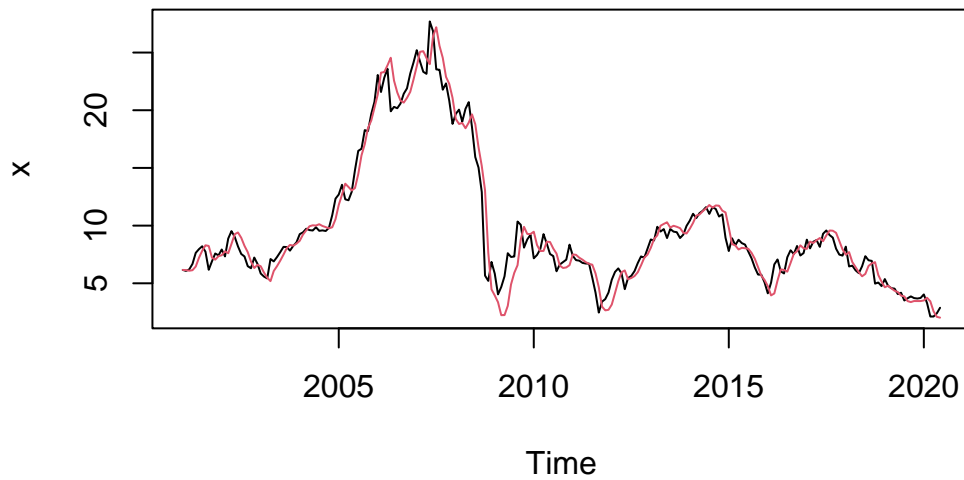
- Suavización Holt, con $\alpha = 0.5$ y $\beta = 0.2, \beta = 0.8$ ¿Cuál es el papel de β en la estimación de valores pronostico?

```

ts.holt_0502 = Holt(ts_base, alpha = 0.5, beta=0.2)

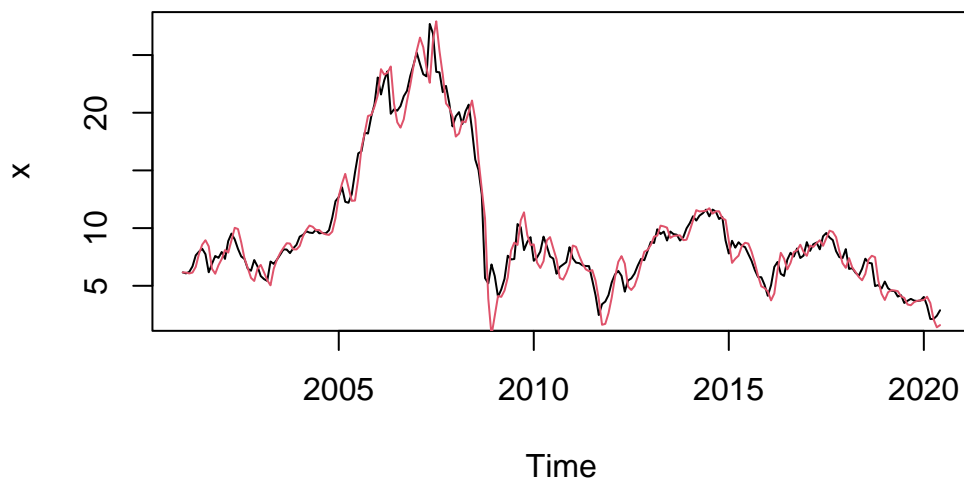
```

original v.s smoothed data



```
ts.holt_0508 = Holt(ts_base, alpha = 0.5, beta=0.8)
```

original v.s smoothed data



ESCRIBIR RESPUESTA

- Suavización Holt-Winters, con $\alpha = 0.5$, $\beta = 0.2$, $\gamma = 0.8$ y otro modelo Holt-Winters con una selección automática de R.

```
hw.ts.1 = hw(ts_base, h=12, optim.start = c(0.5, 0.2, 0.8)) # :c
hw.ts.2 = hw(ts_base, h=12, seasonal='additive')
```

5. Cada inciso del punto anterior tiene dos o tres modelos del mismo tipo, así que una misma gráfica muestra los resultados junto con la serie original por modelo.

HACER GRÁFICOS

6. Seleccione el mejor modelo y mencione el por qué de su elección.

RESPONDER MODELO

7. Utilizar la autocorrelación para determinar si hay patrones en los residuos de una serie de tiempo.

HACER GRÁFICO ACF

8. Desea proyectar la información un año, ¿Cuál sería el mejor modelo para la proyección? realice la proyección y concluya si la proyección tiene sentido o no, detallando sus observaciones

VER LOS DATOS Y ELEGIR

9. Demuestra que si una serie de tiempo es estacionaria en media y varianza, entonces su autocorrelación es

$$\rho(h) = \text{Corr}(x_t, x_{t+h}) = \frac{E[x_t, x_{t+h}] - \mu}{\sigma^2}$$

10. Demuestra mediante inducción que en el suavizado exponencial esta expresión es válida

$$\hat{x}_{t+1} = \alpha x_t + (1 - \alpha) \hat{x}_t$$

$$\bar{X}_{t+1} = \alpha X_t + (1 - \alpha) \bar{X}_t$$