UNAP

FINESI

Series de Tiempo

VI Semestre

Grupo: B

EJEMPLOS

Modelos ARIMA

Mtr. Alcides Ramos Calcina

FINESI

Ejemplo 1

La empresa Cameron Consulting Corporation se especializa en los servicios de inversión de cartera. El analista de la compañía, tiene la obligación de encontrar un modelo lineal univariado para pronosticar los promedios de cierre diario de Transportación. La Tabla 5.1 representa 65 promedio de cierre diario del Índice de Transportación durante el verano.

Tabla 5.1
Promedio de cierre diario del Índice de Transportaciones.

t	Pt	t	Pt	t	Pt	t	Pt	t	Pt
1	222.34	14	223.56	27	241.14	40	251.80	53	258.62
2	222.24	15	223.07	28	241.48	41	251.07	54	259.25
3	221.17	16	225.36	29	246.74	42	248.05	55	261.49
4	218.88	17	227.60	30	248.73	43	249.76	56	264.95
5	220.05	18	226.82	31	248.83	44	251.66	57	268.25
6	219.61	19	229.69	32	248.78	45	253.41	58	272.16
7	216.40	20	229.30	33	249.61	46	252.04	59	272.79
8	217.33	21	228.96	34	249.90	47	248.78	60	275.03
9	219.69	22	229.99	35	246.45	48	247.76	61	278.49
10	219.32	23	233.05	36	247.57	49	249.27	62	281.75
11	218.25	24	235.00	37	247.76	50	247.95	63	285.70
12	220.30	25	236.17	38	247.81	51	251.41	64	286.33
13	222.54	26	238.31	39	250.68	52	254.67	65	288.57

Fuente: Hanke, J.E. - Pronóstico en los negocios.

Ejemplo 1

- **Librería e importación de datos**

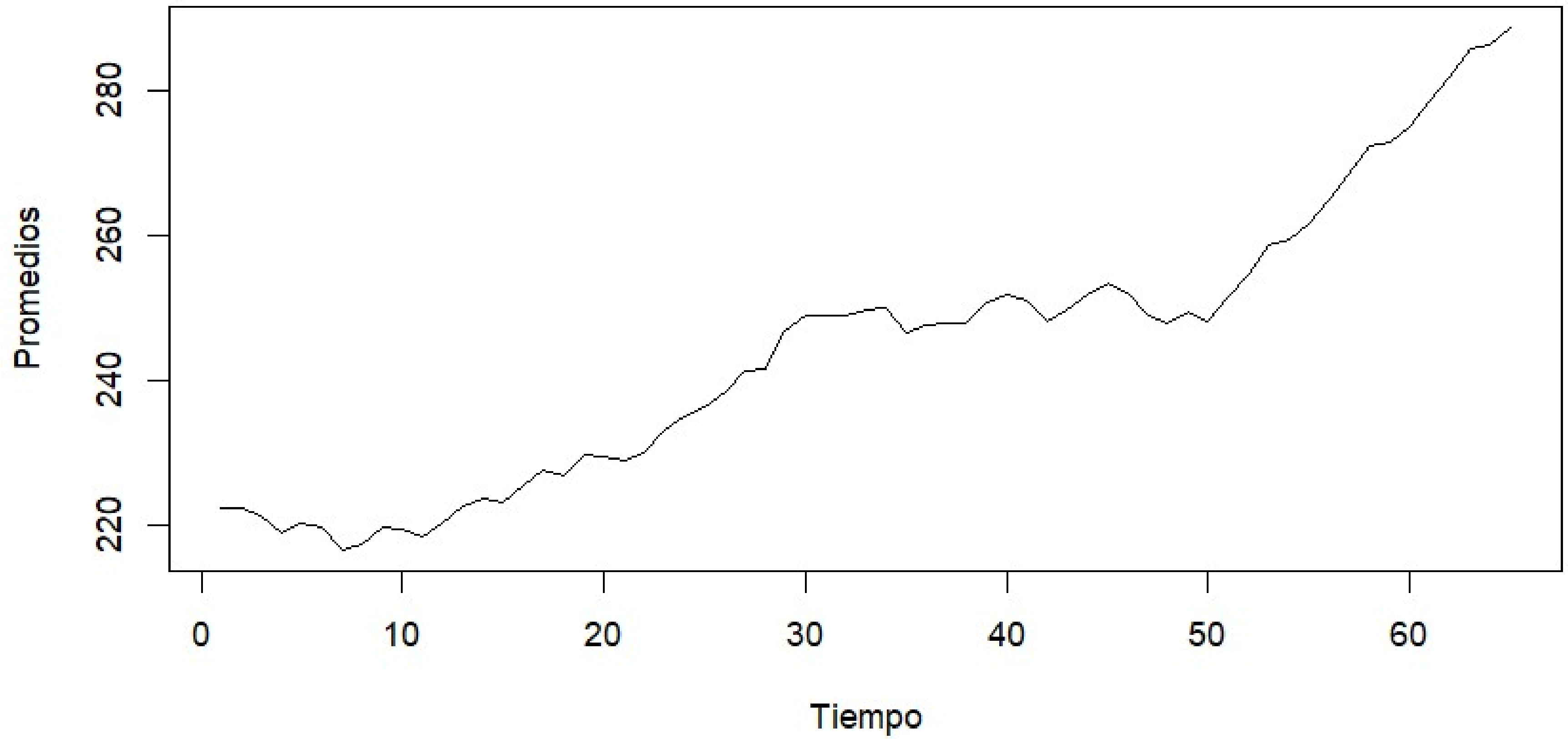
```
# Libreria necesaria
library(forecast) # Modelo ARIMA
library(tseries)  # Para series de tiempo
library(TSA)      # Para series de tiempo
library(urca)     # Para hacer el Test de Raiz Unitaria (detectar si hay o no estacionariedad)
library(ggplot2)  # Para hacer gráficos
library(dplyr)    # Para la manipulación de datos (filtrar, seleccionar, agregar, transformar)

#Importar datos
library(readxl)
datos <- read_excel("D:/... /Ejm_5_1-Indice.xlsx")
View(datos)
```

Gráfica de la serie

```
Ypt <- ts(datos$Pt, start = 1, frequency = 1)
plot(Ypt, xlab="Tiempo", ylab="Promedios")
```

Ejemplo 1

Promedios de cierre del Índice de Transportaciones



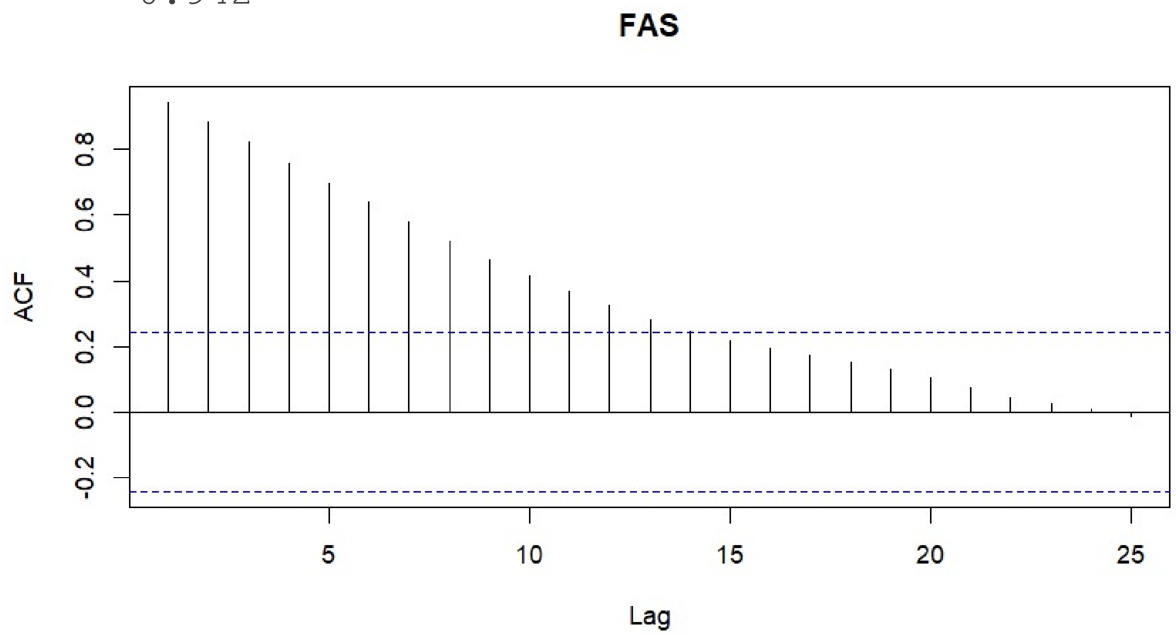
En apariencia, se observa una tendencia creciente en la serie. Por consiguiente, estaríamos en el caso de una serie no estacionaria en media.

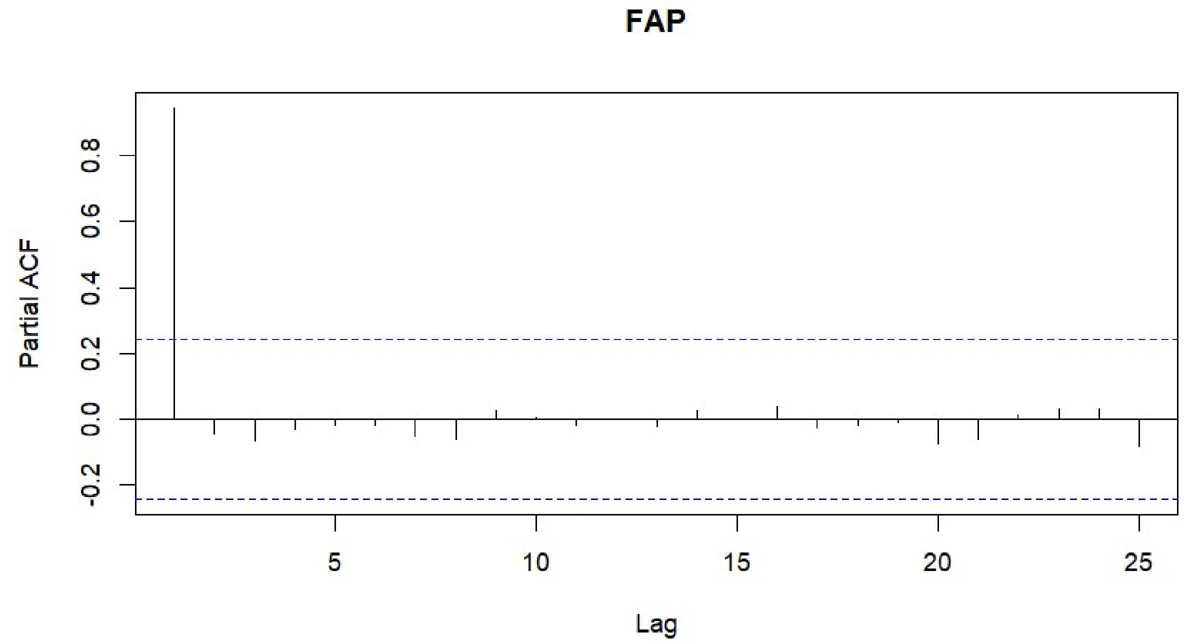
Ejemplo 1

- **Estacionariedad en media**

Los correlograma FAS y FAP de la muestra.

```
FAS <- acf(Ypt, lag.max = 25, main="FAS", level = 0.95)
FAP <- pacf(Ypt, lag.max = 25, main="FAP", level = 0.95)
FAP[1]

Partial autocorrelations of series ‘Ypt’, by lag
    1
0.942
```





Ejemplo 1

Verificación a través de la prueba de Raíz Unitaria de Dickey-Fuller Aumentada para la serie con tendencia.

```
Yp_adf <- ur.df(Ypt, type="trend", selectlags = "AIC", lags = 1)
summary(Yp_adf)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
...
Residual standard error: 1.839 on 59 degrees of freedom
Multiple R-squared:  0.1584,       Adjusted R-squared:  0.1156
F-statistic: 3.701 on 3 and 59 DF,  p-value: 0.0165

Value of test-statistic is: -1.2392 4.6721 2.7329

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

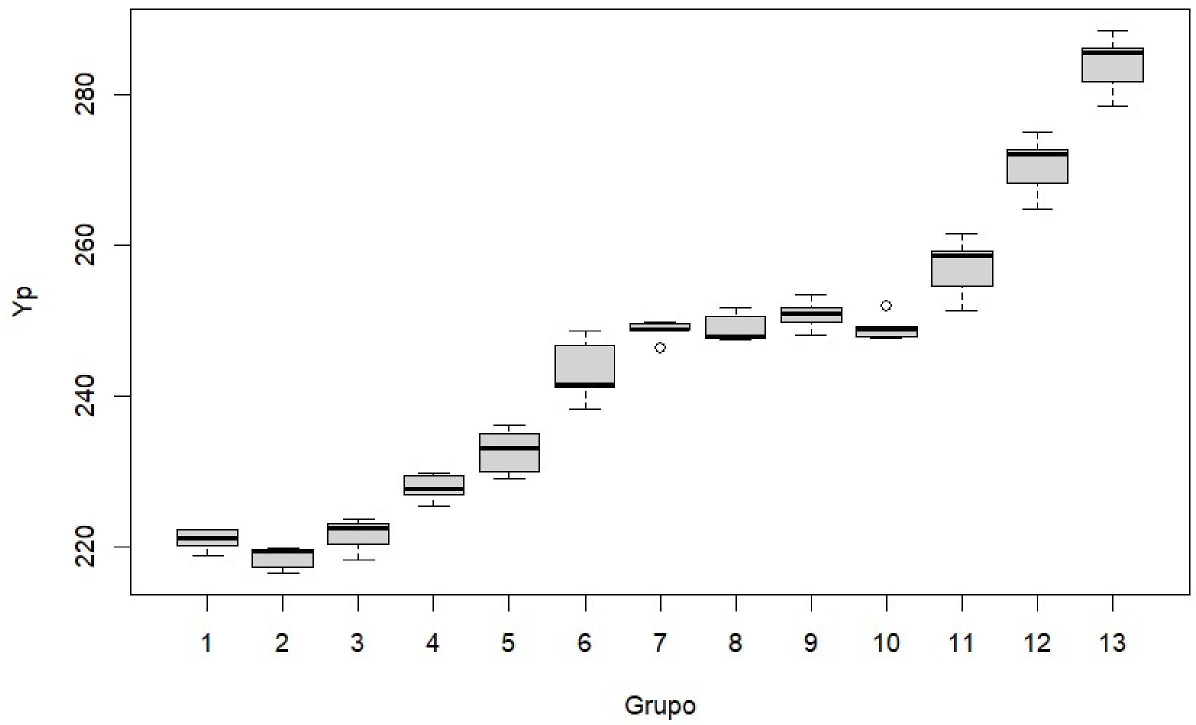
De los resultados, el valor de $\tau_{calculado}$ = -1.2392 > $\tau_{crítico}$ = -3.45, entonces aceptamos la hipótesis nula de **no estacionariedad**.

Ejemplo 1

- **Estacionariedad en varianza**

En este caso tomar un buen número de grupos para estudiar el comportamiento de la varianza, tomamos 13 grupos cuya representación a través de diagrama de cajas es:

```
Grupo <- c(rep(1:13,rep(5,13)))
boxplot(datos$Pt ~ Grupo, xlab = "Grupo", ylab="Yp")
```



Se puede concluir que los grupos de muestras tienen similar variación a lo largo de la evolución de la serie, entonces podemos suponer que la serie es estacionaria en varianza.

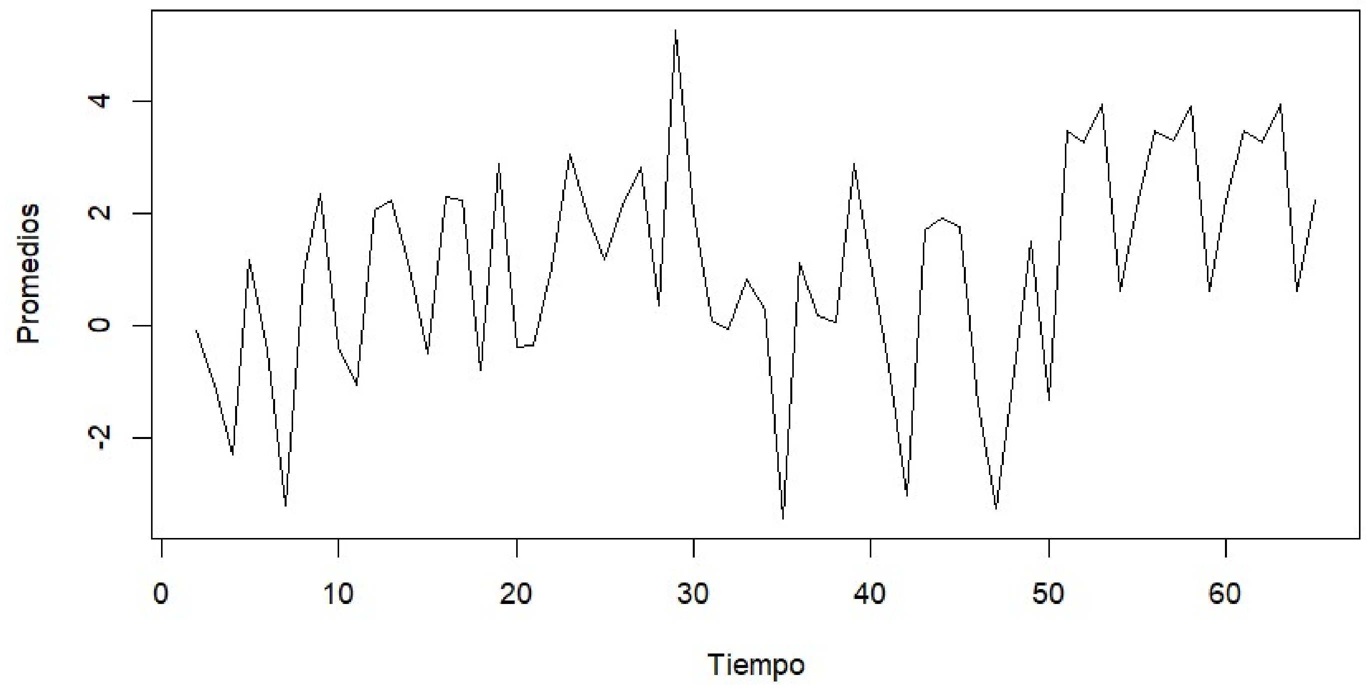
Ejemplo 1

- **Corrección de la estacionariedad en media**

Transformación de la primera diferencia: $\Delta Y_t = Y_t - Y_{t-1} = (1-L)Y_t$

Se utiliza la función **`diff( )`** para calcular la diferencia de una serie.

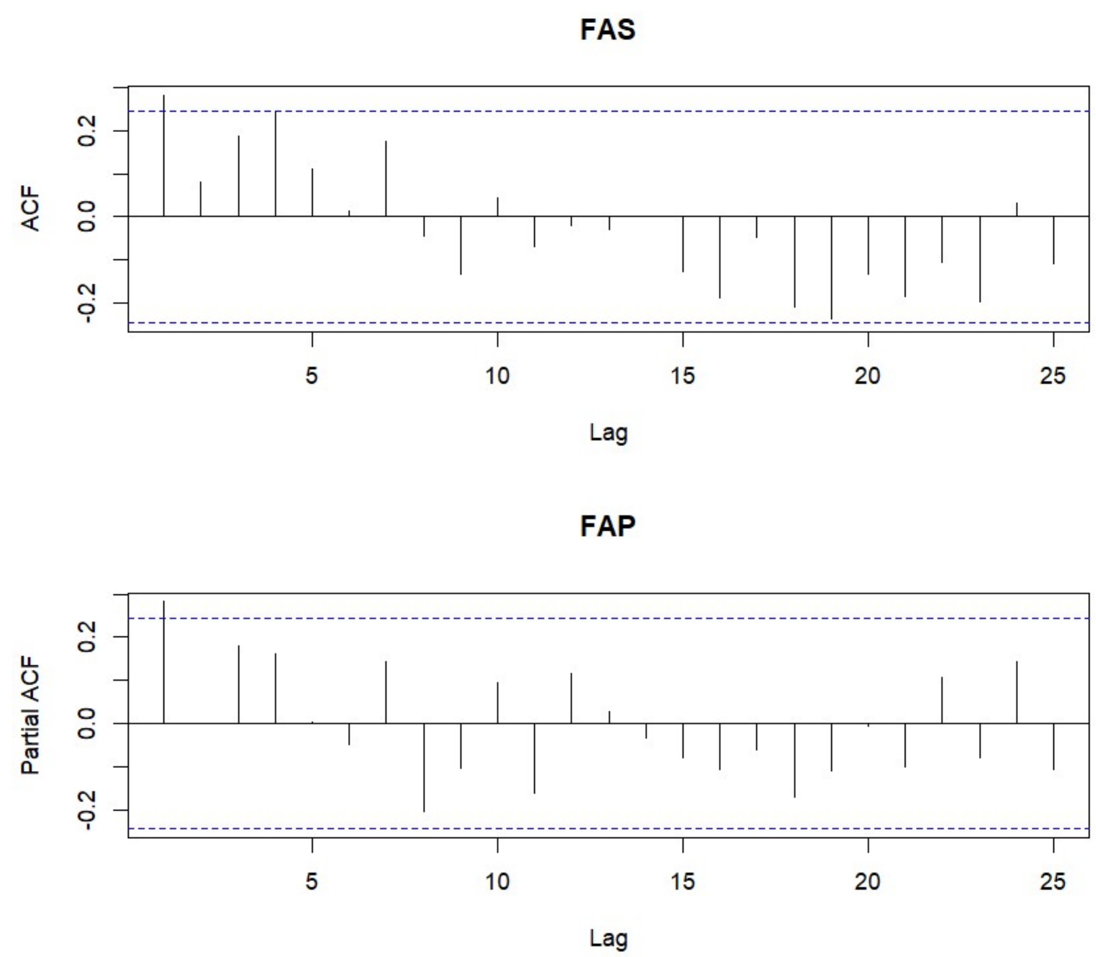
```
dif1_Ypt <- diff(Ypt)
plot(dif1_Ypt, xlab="Tiempo", ylab="Promedios")
```



Ejemplo 1

Ahora obtendremos los correlogramas FAS y FAP de la serie diferenciada.

```
par(mfrow = c(2,1))
FAS <- acf(dif1_Ypt, lag.max = 25, main="FAS", level
= 0.95)
FAP <- pacf(dif1_Ypt, lag.max = 25, main="FAP", level
= 0.95)
FAP[1]

Partial autocorrelations of series ‘dif1_Ypt’, by lag

    1
0.282
```



Los coeficientes de la FAS del primer al segundo descienden rápidamente hacia cero, así mismo, el primer coeficiente de la FAP es no significativo (0.282 < 0.9), serían altos indicios de que la serie es **estacionaria**.

Ejemplo 1

A continuación, realizamos la prueba de la raíz unitaria de Dickey-Fuller Aumentada la serie **dif1_Ypt**.

```
d1_Yp_adf <- ur.df(dif1_Ypt, type="drift", selectlags = "AIC", lags = 1)
summary(d1_Yp_adf)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

---
Residual standard error: 1.907 on 59 degrees of freedom
Multiple R-squared:  0.3659,	Adjusted R-squared:  0.3444
F-statistic: 17.02 on 2 and 59 DF,  p-value: 1.459e-06


Value of test-statistic is: -4.6778 10.9693

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

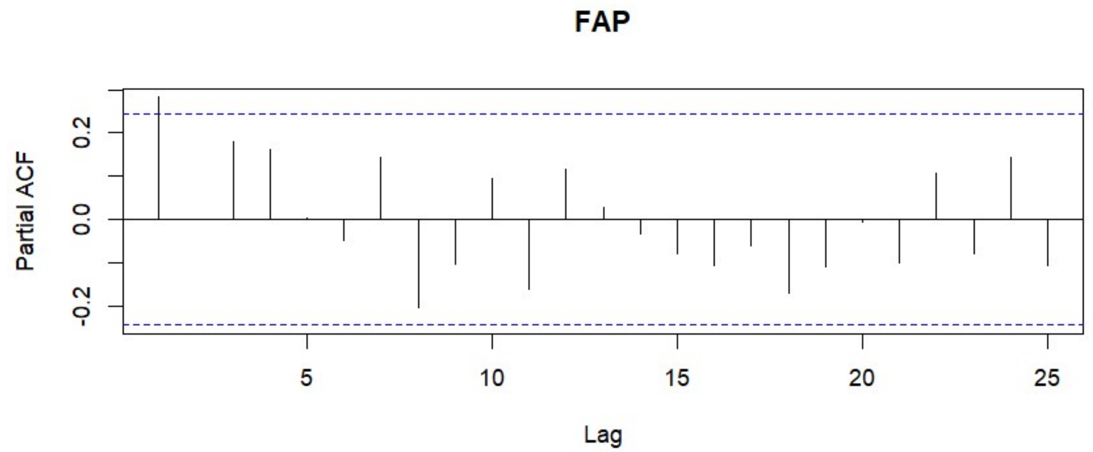
De los resultados, el valor de $\tau_{calculado} = -4.6778 < \tau_{crítico} = -2.89$, por tanto, se rechaza la hipótesis nula (Ho) de la existencia de raíz unitaria, es decir, que la serie de promedios Ypt ahora si es ESTACIONARIA.

Ejemplo 1

- **Identificación del orden de un proceso AR**

Los límites de confianza al nivel del 95% se determinan: $I_{0.95} \approx \left[-1.96 \cdot \sqrt{1/65}\ ,\ 1.96 \cdot \sqrt{1/65}\right]$

$$I_{0.95} \approx [-0.243\ ,\ 0.243]$$



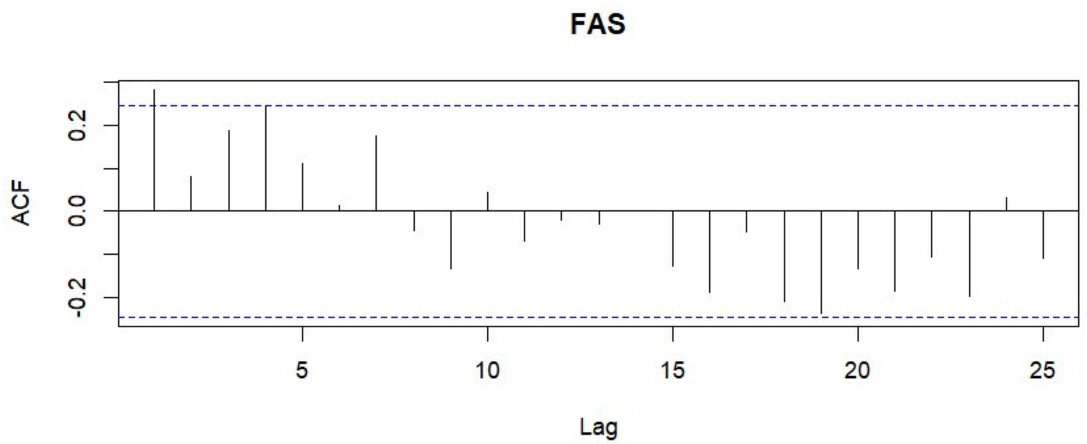
- El primer coeficiente de autocorrelación parcial (FAP) es significativo, es decir, distinto de cero (0.286 $\notin$ [-0.243, 0.243]).

- Por tanto, este patrón de autocorrelaciones parciales determina un modelo **AR(1)**.

Ejemplo 1

- **Identificación del orden de un proceso MA**

Observamos el FAS:



- El primer coeficiente de autocorrelación simple (FAS) es significativamente distintos de cero, y los FAC en general a partir del segundo coeficiente decrecen lentamente en forma sinusoidal esto se asemeja al comportamiento de un modelo teórico **MA(1)**.

Por consiguiente, planteamos los siguientes modelos al Índice de Transportación:

✓ ARIMA(1, 1, 0)
✓ ARIMA(0, 1, 1)
✓ ARIMA(1, 1, 1)

Ejemplo 1

- **Estimación de parámetros**

Modelo **ARIMA(1, 1, 0)**: $\Delta Y_t = \delta + \phi_1 \Delta Y_{t-1} + a_t$

```
modelo1 =  arma(dif1_Ypt, order = c(1,0))
summary(modelo1)

Call:
arma(x = dif1_Ypt, order = c(1, 0))

Model:
ARMA(1,0)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.2970 -1.2553  0.3172  1.4859  4.3989 
Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)   
ar1          0.2837      0.1199    2.366  0.01798 * 
intercept    0.7647      0.2625    2.913  0.00358 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 3.516,  Conditional Sum-of-Squares = 217.97,  AIC = 266.08
```

Los parámetros estimados son significativos ($p < \alpha$) , por tanto, el modelo es adecuado para los pronósticos.

El primer modelo estimado es:

ARIMA(1,1,0): $\Delta Y_t = 0.7647 + 0.2837 \Delta Y_{t-1} + a_t$

También

$$Y_t = Y_{t-1} + 0.7647 + 0.2837\left(Y_{t-1} - Y_{t-2}\right) + a_t$$

Ejemplo 1

- **Estimación de parámetros**

Modelo **ARIMA(0, 1, 1)**: $\Delta Y_t = \mu + a_t - \theta_1 a_{t-1}$

```
modelo2 =  arma(dif1_Ypt, order = c(0,1))
summary(modelo2)

Call:
arma(x = dif1_Ypt, order = c(0, 1))

Model:
ARMA(0,1)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.2780 -1.2018  0.2011  1.5141  4.5394 

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)    
ma1          0.2915      0.1228    2.373 0.017629 *  
intercept    1.0525      0.2995    3.514 0.000441 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Fit:
sigma^2 estimated as 3.527,  Conditional Sum-of-Squares = 218.7,  AIC = 266.3
sigma^2 estimated as 3.516,  Conditional Sum-of-Squares = 217.97,  AIC = 266.08
```

Los parámetros estimados son significativos (p < α) , por tanto, el modelo es adecuado para los pronósticos.

El primer modelo estimado es:

ARIMA(0,1,1): $\Delta Y_t = 1.0525 + a_t - 0.2915 a_{t-1}$

También

$$Y_t = Y_{t-1} + 1.0525 + a_t - 0.2915 a_{t-1}$$

Ejemplo 1

- **Estimación de parámetros**

Modelo **ARIMA(1, 1, 1)**: $\Delta Y_t = \delta + \phi_1 \Delta Y_{t-1} + a_t - \theta_1 a_{t-1}$

```
modelo3 = arma(dif1_Ypt, order = c(1,1))
summary(modelo3)

Call:
arma(x = dif1_Ypt, order = c(1, 1))

Model:
ARMA(1,1)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5723 -1.4971  0.1096  1.3600  3.8244 

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)    
ar1          0.8740      0.1122    7.786 6.88e-15 ***
ma1         -0.7291      0.1609   -4.530 5.89e-06 ***
intercept    0.1711      0.1304    1.312    0.189    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Fit:
sigma^2 estimated as 3.365, Conditional Sum-of-Squares = 208.72,  AIC = 265.29
```

Los parámetros estimados son significativos ($p < \alpha$) , excepto el intercepto, por tanto, el modelo es:

El primer modelo estimado es:

ARIMA(1,1,1):

$$\Delta Y_t = 0.1711 + 0.8740 \Delta Y_{t-1} + a_t - 0.7291 a_{t-1}$$

El modelo debería ser estimado sin intercepto:

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + a_t - \theta_1 a_{t-1}$$

Ejemplo 1

Modelos sin intercepto

Modelo **ARIMA(1, 1, 1)**: $\Delta Y_t = \phi_1 \Delta Y_{t-1} + a_t - \theta_1 a_{t-1}$

```
modelo3 = arma(dif1_Ypt, order = c(1,1), include.intercept = FALSE)
summary(modelo3)

Call:
arma(x = dif1_Ypt, order = c(1, 1), include.intercept = FALSE)

Model:
ARMA(1,1)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.3426 -1.1487  0.3751  1.8241  3.9517 

Coefficient(s):
     Estimate  Std. Error  t value Pr(>|t|)    
ar1   0.97233     0.06488   14.987  < 2e-16 ***
ma1  -0.76950     0.12664   -6.076 1.23e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 3.482,  Conditional Sum-of-Squares = 220.77,  AIC = 265.48
```

Finalmente, el modelo ARIMA(1, 1, 1) sin intercepto sería:

$$\Delta Y_t = 0.97233 \Delta Y_{t-1} + a_t - 0.76950 a_{t-1}$$

Ejemplo 2

Considere la tabla 5.2, en la cual se proporcionan 90 valores de las ventas semanales, en unidades de 1000 tubos de la pasta dental Ultra Shine.

Tabla 5.2
Ventas semanales de pasta dental Ultra Shine.

t	Yt	t	Yt	t	Yt	t	Yt	t	Yt
1	235.0	19	434.9	37	607.0	55	791.1	73	930.8
2	239.0	20	445.8	38	617.5	56	805.8	74	941,3
3	244.1	2	455.9	39	622.9	57	815.1	75	950.3
4	252.7	22	465.6	40	633.4	58	822.9	76	952.4
5	264.4	23	477.9	41	647.4	59	830.7	77	960.0
6	277.9	24	491.4	42	658.2	60	839.6	78	968.1
7	286.7	25	507.7	43	670.8	61	846.9	79	972.5
8	295.6	26	517.2	44	685.5	62	853.8	80	977.4
9	310.4	27	524.3	45	690.9	63	860.8	81	977.7
10	325.1	28	532.1	46	693.6	64	871.1	82	979.5
11	336.1	29	538.1	47	700.7	65	877.8	83	982.9
12	344.5	30	544.9	48	712.5	66	881.1	84	985.8
13	355.4	31	551.9	49	726.5	67	884.2	85	991.4
14	367.7	32	557.9	50	736.4	68	890.2	86	996.3
15	384.0	33	564.3	51	743.2	69	894.9	87	1003.1
16	398.0	34	572.2	52	751.2	70	901.3	88	1010.3
17	412.9	35	582.9	53	764.3	71	913.1	89	1018.4
18	422.9	36	595.3	54	777.9	72	922.5	90	1029.5

Fuente: Bowerman, B. L. – Pronósticos, series de tiempo y regresión.


GRACIAS
https://aulavirtual2.unap.edu.pe/