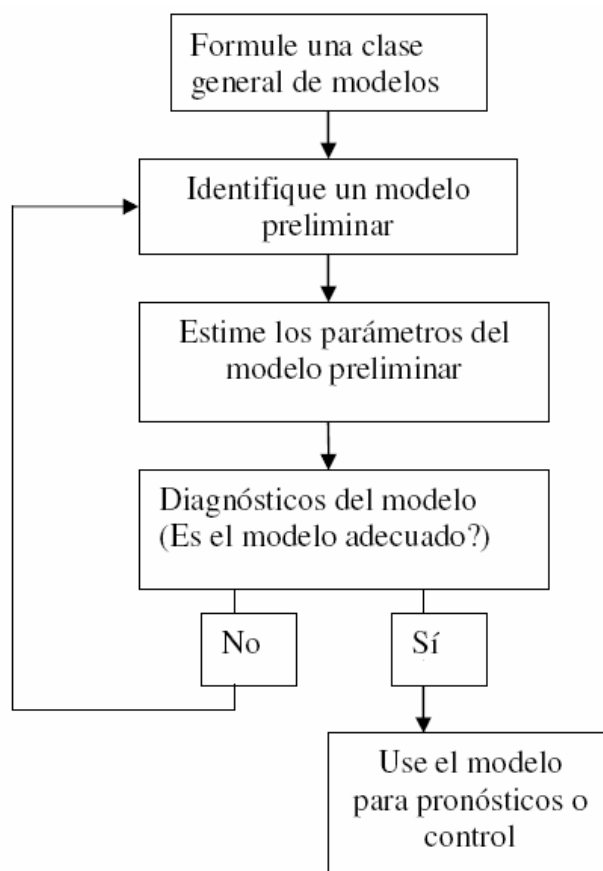


4. Construcción de modelos para series de tiempo univariadas: Metodología de Box-Jenkins. Identificación, estimación y validación de un modelo.

- Para la construcción de un modelo de series de tiempo, Box y Jenkins sugieren seguir las etapas del siguiente procedimiento iterativo.



- **Identificación.** Considere el modelo ARIMA(p,d,q)

$$\phi(B)(1-B)^d Z_t^{(\lambda)} = \theta_0 + \theta(B)a_t$$

bajo las condiciones de regularidad anteriormente definidas. La identificación del modelo se refiere a la metodología para identificar:

- ✓ La transformación estabilizadora de varianza, λ
 - ✓ El orden de la diferenciación, d .
 - ✓ Los órdenes apropiados para p y q .
 - ✓ La decisión de incluir el parámetro θ_0 .
- Pasos útiles para la identificación del modelo tentativo:

Paso 1. Grafique la serie de tiempo y elija la transformación adecuada para estabilizar la varianza de la serie usando la transformación de Box y Cox; si $\lambda = 1$ o está cerca de 1, no es necesario transformar la serie para estabilizar varianza. Para el empleo de la transformación de Box y Cox es necesario tener un número moderado o grande de observaciones.

Paso 2. Calcule y examine la ACF y PACF de la serie obtenida en el Paso 1. Del análisis del gráfico de la serie y del examen de la ACF y PACF determine si es necesario diferenciar la serie. Algunas reglas generales son:

- Observe si la serie tiene tendencias.
- Si la ACF muestral decae muy lentamente en forma aritmética (los valores individuales de la ACF pueden no ser grandes) y la PACF presenta un corte después del rezago 1, esto es indicativo de que posiblemente se necesita diferenciar la serie. Use la primera diferencia $(1-B)Z_t$. Alternativamente, es conveniente usar **pruebas de raíces unitarias** (Más tarde se verán distintas prueba de raíces unitarias). En casos dudosos, generalmente es recomendable diferenciar.
- Verifique si la serie diferenciada es estacionaria graficándola y calculando su ACF y PACF como antes o usando el contraste de raíces unitarias sobre la serie diferenciada. En la práctica, generalmente d toma los valores de 0, 1 o 2. Se debe tener cuidado de no sobrediferenciar la serie:

esto puede conducir a modelos sobreparametrizados y a la no invertibilidad.

PASO 3. Calcule y examine la ACF y PACF muestral de la serie adecuadamente transformada para identificar a p y q. Para esta identificación recuerde los siguientes resultados teóricos:

PROCESO	ACF	PACF
AR(p)	Decaimiento exponencial y/o onda sinusoidal amortiguada	Corte después del rezago p
MA(q)	Corte después del rezago q	Decaimiento exponencial y/o onda sinusoidal amortiguada
ARMA(p,q)	Decaimiento exponencial y/o onda sinusoidal amortiguada	Decaimiento exponencial y/o onda sinusoidal amortiguada

Para construir adecuadamente el modelo se necesitan un mínimo de 50 observaciones, si no existen componentes estacionales. El número de elementos estimados en la ACF y PACF debe estar alrededor de $n/4$.

Las varianzas estimadas para los estimadores de los elementos de la ACF y PACF:

- Bajo la hipótesis de que $\rho_k=0$ para $k>m$

$$Var(\hat{\rho}_k) \approx \frac{1}{n} (1 + 2\hat{\rho}_1^2 + 2\hat{\rho}_2^2 + \dots + 2\hat{\rho}_m^2)$$

- Bajo la hipótesis de que $\phi_{kk}=0$ para todo k

$$Var(\hat{\phi}_{kk}) \approx 1/n$$

- Si el proceso es identificado como AR (ACF decae exponencialmente) determine su orden contrastando en la PACF muestral, cuál es el máximo rezago cuyo coeficiente de correlación parcial es significativamente diferente de cero. El correlograma parcial muestra el gráfico de la PACF muestral junto con una banda. Si el coeficiente cae al interior de la banda, quiere decir que el coeficiente no es estadísticamente diferente de cero.
- Si el proceso es identificado como un MA (PACF decae exponencialmente) determine su orden contrastando en la ACF muestral, cual es el máximo rezago cuyo coeficiente de correlación es significativamente diferente de cero. El correlograma muestra el gráfico de la ACF muestral junto con una banda. Si el coeficiente cae al interior de la banda, quiere decir que el coeficiente no es estadísticamente diferente de cero.
- Si el proceso es identificado como un ARMA (ACF y PACF con decaimiento exponencial) no podemos determinar su orden usando contrastes de hipótesis sobre la ACF y PACF muestrales. En estos casos puede usarse la Función de Autocorrelación Extendida muestral, ESACF (Tsay y Tiao, 1984).

PASO 4: Finalmente, verifique la existencia de θ_0 el cual corresponde a una constante en el modelo (si $d=0$) o a un término de tendencia determinística, (si $d>0$). Para contrastar la hipótesis $H_0: \theta_0=0$, use el estadístico $t = \bar{W} / S_{\bar{W}}$, donde \bar{W} es la media de la serie diferenciada y $S_{\bar{W}}$ es su desviación estándar si $d>0$, o el estadístico $t = \bar{Z} / S_{\bar{Z}}$, si $d=0$. La desviación estándar $S_{\bar{W}}$ está dada por

$$S\bar{w} = \left(\frac{\hat{\gamma}_0}{n} (1 + 2\hat{\rho}_1^2 + 2\hat{\rho}_2^2 + \dots + 2\hat{\rho}_s^2) \right)^{1/2}$$

donde $\hat{\gamma}_0$ es la varianza muestral y $\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_s$ son las primeras s autocorrelaciones muestrales significativas del proceso diferenciado W_t (si no hay que diferenciar use $S_{\bar{z}}$ calculada en forma similar usando la serie original Z_t). Rechace H_0 si $|t| > t(\alpha/2, n-1)$, donde $t(\alpha/2, n-1)$ es el percentil $\alpha/2$ superior de una distribución t con $n-1$ grados de libertad.

Alternativamente, θ_0 puede ser descartado si después de haber sido introducido al modelo, su estimación no es estadísticamente significativa.

Prueba de la hipótesis de la existencia de raíces unitarias en una serie de tiempo.

Considere el proceso

$$z_t = \beta_0 + \beta_1 t + e_t$$

donde t es el tiempo y e_t sigue un proceso estacionario ARMA. Este proceso posee una tendencia determinística lineal y es llamado un proceso de tendencia-estacionario (Trend Stationary o TS). Observe que el proceso $z_t - \beta_0 - \beta_1 t = e_t$ es estacionario, y es el proceso adecuadamente *destendencializado* (detrended).

Considere ahora el proceso

$$z_t = \beta_0 + z_{t-1} + e_t$$

donde e_t sigue un proceso estacionario ARMA. La tendencia de este proceso es una mezcla de tendencia aleatoria y determinística lineal y es llamado un

proceso de diferencia-estacionario (Difference Stationary o DS). En ese caso, el proceso $z_t - z_{t-1} = \gamma_0 + e_t$ es estacionario y es el proceso adecuadamente destendencializado.

Uno de los problemas de interés en medición económica tiene que ver con la determinación del tipo de tendencia presente en una serie de tiempo: se debe decidir si una tendencia aleatoria o una determinística está presente en una serie de tiempo dada. Esta distinción es importante puesto que la naturaleza de la permanencia de las innovaciones o 'shocks' económicos sobre la serie depende del tipo de tendencia que posea: mientras que en una serie con tendencia determinística el 'shock' solamente tiene una permanencia *temporal*, en una con tendencia aleatoria dichos impactos son de naturaleza *permanente*.

La prueba de raíces unitarias de Dickey-Fuller

Dickey y Fuller (1979) consideran tres pruebas distintas para detectar la presencia de una raíz unitaria en el polinomio autorregresivo del proceso, bajo el supuesto de normalidad. Las pruebas consideran los siguientes modelos:

$$\begin{aligned} \checkmark \quad z_t &= \rho z_{t-1} + a_t, \\ \checkmark \quad z_t &= \beta_0 + \rho z_{t-1} + a_t, \\ \checkmark \quad z_t &= \beta_0 + \beta_0 t + \rho z_{t-1} + a_t, \end{aligned}$$

donde a_t es $\text{RBN}(0, \sigma_a^2)$.

Prueba 1: Utiliza el modelo $\Delta z_t = \gamma z_{t-1} + a_t$, el cual se obtiene restando z_{t-1} a ambos lados de la primera ecuación $z_t = \rho z_{t-1} + a_t$. El parámetro

$\gamma = \rho - 1$. En este caso, para probar la existencia de una raíz unitaria, se deben contrastar las hipótesis

$H_0 : \gamma = 0$, lo que implica que $\rho = 1$ y la serie tiene una raíz unitaria y por tanto la serie es no estacionaria.

$H_1 : \gamma < 0$, lo que implica que la serie no tiene una raíz unitaria y por tanto la serie es estacionaria.

Observe que bajo H_0 , el proceso $\Delta z_t = \gamma z_{t-1} + a_t$ es estacionario. El estadístico de la prueba es

$$\tau = \hat{\gamma} / se(\hat{\gamma})$$

donde $\hat{\gamma}$ es el estimador de mínimos cuadrados de γ y $se(\hat{\gamma})$ es su error estándar en el modelo $\Delta z_t = \gamma z_{t-1} + a_t$ donde a_t es $RBN(0, \sigma_a^2)$.

La distribución de τ bajo H_0 **no es la t de student tradicional** sino que pertenece a una clase de distribuciones no estándar y se encuentra tabulada para diferentes valores de n y niveles de significancia α . Si $\tau(\alpha, n)$ es el percentil α -inferior de dicha distribución, se debe rechazar H_0 si el valor observado de $\tau < \tau(\alpha, n)$.

Prueba 2: Utiliza el modelo $\Delta z_t = \beta_0 + \gamma z_{t-1} + a_t$, el cual se obtiene restando z_{t-1} a ambos lados de la segunda ecuación $z_t = \beta_0 + \rho z_{t-1} + a_t$. Para probar la existencia de una raíz unitaria, se deben contrastar las hipótesis

$H_0 : \gamma = 0$, lo que implica que la serie tiene una raíz unitaria y por tanto la serie es no estacionaria

$H_1 : \gamma < 0$, lo que implica que la serie no tiene una raíz unitaria y por tanto la serie es estacionaria

Observe que bajo H_0 , el proceso $\Delta z_t = \beta_0 + \gamma z_{t-1} + a_t$ es estacionario. El estadístico de la prueba es

$$\tau_\mu = \hat{\gamma} / se(\hat{\gamma})$$

Donde $\hat{\gamma}$ es el estimador de mínimos cuadrados de γ y $se(\hat{\gamma})$ es su error estándar en el modelo $\Delta z_t = \beta_0 + \gamma z_{t-1} + a_t$ donde a_t es $RBN(0, \sigma_a^2)$.

La distribución de τ_μ bajo H_0 no es la t de student tradicional sino que pertenece a una clase de distribuciones no estándar y es diferente a la distribución del estadístico de la prueba anterior. Se encuentra tabulada para diferentes valores de n y niveles de significancia α . Si $\tau_\mu(\alpha, n)$ es el percentil α -inferior de dicha distribución, se debe rechazar H_0 si el valor observado de $\tau_\mu < \tau_\mu(\alpha, n)$

Prueba 3: Utiliza el modelo $\Delta z_t = \beta_0 + \beta_1 t + \gamma z_{t-1} + a_t$, el cual se obtiene restando z_{t-1} a ambos lados de la segunda ecuación $z_t = \beta_0 + \beta_0 t + \rho z_{t-1} + a_t$. Para probar la existencia de una raíz unitaria, se deben contrastar las hipótesis

$H_0 : \gamma = 0$, lo que implica que la serie tiene una raíz unitaria y por tanto la serie es no estacionaria, $I(1)$.

$H_1 : \gamma < 0$, lo que implica que la serie no tiene una raíz unitaria y por tanto la serie es estacionaria en tendencia.

Observe que bajo H_0 , el proceso $\Delta z_t = \beta_0 + \beta_1 t + \gamma z_{t-1} + a_t$ es estacionario en tendencia. Estadístico de prueba es

$$\tau_\tau = \hat{\gamma} / se(\hat{\gamma})$$

Donde $\hat{\gamma}$ es el estimador de mínimos cuadrados de γ y $se(\hat{\gamma})$ es su error estándar en el modelo $\Delta z_t = \beta_0 + \beta_1 t + \gamma z_{t-1} + a_t$ donde a_t es $RBN(0, \sigma_a^2)$.

La distribución de τ_τ bajo H_0 no es la t de student tradicional sino que pertenece a una clase de distribuciones no estándar y es diferente a la de las pruebas anteriores. Se encuentra tabulada para diferentes valores de n y niveles de significancia α . Si $\tau_\tau(\alpha, n)$ es el percentil α -inferior de dicha distribución, se debe rechazar H_0 si el valor observado de $\tau_\tau < \tau_\tau(\alpha, n)$.

La prueba Aumentada de Dickey-Fuller, ADF

No siempre es posible representar cualquier serie de tiempo usando uno de los modelos anteriores donde a_t es $RBN(0, \sigma_a^2)$. Puede ocurrir que el término de error del modelo esté autocorrelacionado y entonces los estimadores no serán eficientes. En este caso, Dickey-Fuller proponen usar las pruebas anteriores extendiendo cada modelo de la siguiente manera:

Prueba 1: Utiliza el modelo $\Delta z_t = \gamma z_{t-1} + \sum_{i=1}^P \delta_i \Delta z_{t-i} + a_t$ con $a_t \sim RBN(0, \sigma_a^2)$

Prueba 2: Utiliza el modelo $\Delta z_t = \beta_0 + \gamma z_{t-1} + \sum_{i=1}^P \delta_i \Delta z_{t-i} + a_t$,

con $a_t \sim RBN(0, \sigma_a^2)$

Prueba 3: Utiliza el modelo $\Delta z_t = \beta_0 + \beta_1 t + \gamma z_{t-1} + \sum_{i=1}^P \delta_i \Delta z_{t-i} + a_t$,

con $a_t \sim RBN(0, \sigma_a^2)$

El coeficiente de interés sigue siendo γ . Los percentiles para las distribuciones nulas de los estadísticos de prueba siguen siendo las mismas. Usando análisis de superficies de respuesta, MacKinnon (1991) proporciona los valores críticos para muestras finitas de la prueba ADF.

- ✓ El verdadero proceso generador de datos puede contener componentes tanto AR como MA. En este caso se puede emplear el resultado de Said-Dickey (1984) de que todo proceso $ARIMA(p,1,q)$ puede ser aproximado adecuadamente por un proceso $ARIMA(p^*,1,0)$ donde $p^* \leq T^{1/3}$.
- ✓ La estimación adecuada de γ y su error estándar depende de la selección adecuada de P , el máximo orden de los rezagos de Δz_t en el lado derecho de la ecuación a estimar. Si P es desconocido, una aproximación es la siguiente, Schwert (1989):

Defina un máximo valor de P como $P = \text{Parte entera de } 12(T/100)^{1/4}$

Estime todos los modelos cuyos órdenes sean $\leq P$

Elija el modelo con menor criterio de Información de Schwartz (u otro criterio de información) para realizar la prueba. Verifique que el modelo elegido cumple con los supuestos de normalidad, homocedasticidad y no autocorrelación.

- ✓ Cuando existe una componente MA, el tamaño de la prueba, α , puede sufrir fuertes distorsiones.
- ✓ La prueba de ADF considera la existencia de una raíz unitaria. Sin embargo, pueden existir más de una.
- ✓ El proceso puede ser estacional, lo que implicaría la posible existencia de raíces estacionales.
- ✓ Pueden existir cambios estructurales los cuales pueden aparentar la existencia de tendencia en la serie.

- ✓ La estimación adecuada de γ y su error estándar depende de la **identificación correcta de las componentes exógenas dadas por la constantes y la tendencia determinística.**
- ✓ La potencia de la prueba es muy baja a medida que $\gamma \rightarrow 0$, es decir $\rho \rightarrow 1$. Considere el siguiente experimento: Para el modelo $z_t = \beta_0 + \beta_1 z_{t-1} + a_t$, donde $a_t \sim RBN(0,1)$, se obtuvieron 10000 realizaciones con $n=100$ observaciones cada una, para diferentes valores de β_1 . A continuación cada serie simulada es estimada como $\Delta z_t = \beta_0 + \gamma z_{t-1} + a_t$ y se usa la prueba de Dickey-Fuller para contrastar $H_0: \gamma = 0$ a un nivel de significancia de $\alpha = 0.10, 0.05, 0.01$. La siguiente tabla muestra el porcentaje de rechazo de H_0 (siendo falsa).

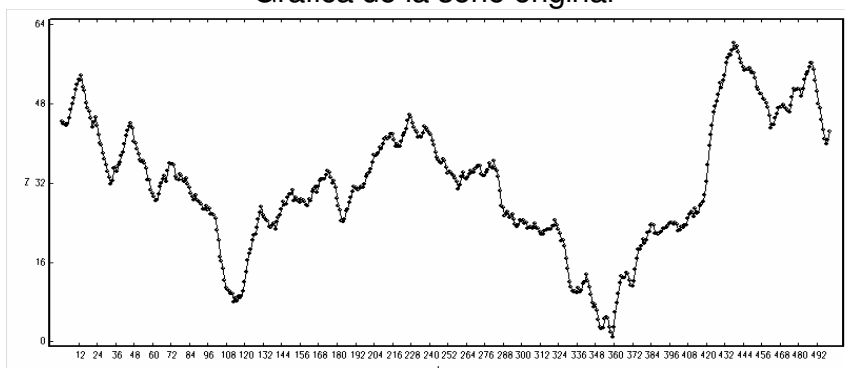
	α		
β_1	0.10	0.05	0.01
0.80	95.9	87.4	51.4
0.90	52.1	33.1	9.0
0.95	23.4	12.7	2.6
0.99	10.5	5.8	1.3

- ✓ Otras pruebas de raíz unitaria: Entre otras están las de Phillips-Perron (1988), Sargan-Bhargava (1983), Cochrane (1988), Perron y Ng (1996), Elliot, Rothenber y Stock (1996), Kwiatkowski, Phillips, Schmidt y Sin (1992). Para una descripción de ellos ver Maddala y Kim (2000). Ver también Castaño y Martínez (2006), para una prueba usando teoría tradicional.

Ejemplo. Serie simulada de un ARIMA(1,1,1), para $n=500$ datos, $\phi_1=0.7$ y $\theta_1=0.3$ y el término de error tiene una distribución $N(0,1)$. A la serie simulada se sumó 7 para volverla positiva.

Paso 1. Estabilización de la varianza de la serie.

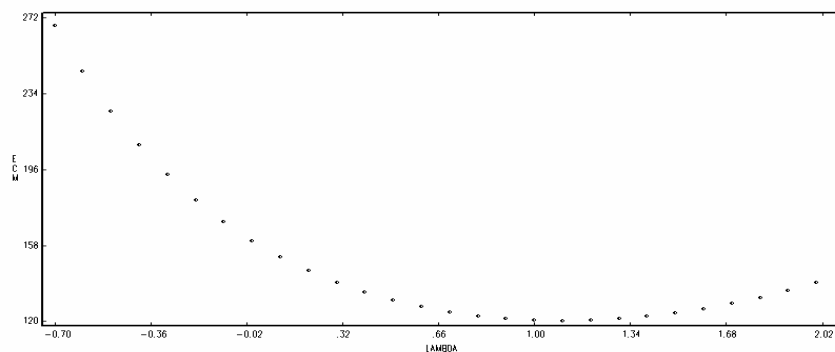
Gráfica de la serie original



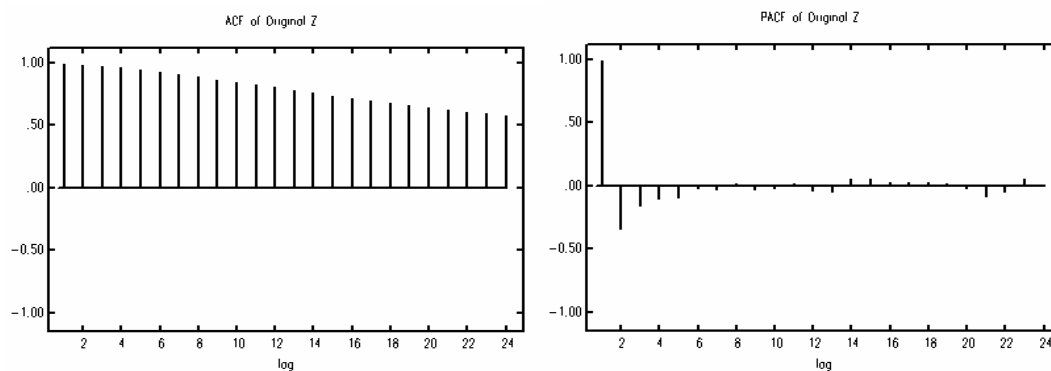
Para el análisis de la transformación estabilizadora de varianza de Box-Cox usando el procedimiento para transformaciones entre -2 y 2 con un incremento de 0.1, se elige aquella transformación para la cual se produce el menor error cuadrático medio, ECM.

VARIABLE	LAMBDA	ECM
1	-2.000	1284.871
2	-1.900	1113.120
3	-1.800	967.498
4	-1.700	843.810
5	-1.600	738.565
6	-1.500	648.852
⋮	⋮	⋮
23	.200	145.351
24	.300	139.505
25	.400	134.605
26	.500	130.555
27	.600	127.272
28	.700	124.689
29	.800	122.748
30	.900	121.403
31	1.000	120.614
32	1.100	120.350
33	1.200	120.588
34	1.300	121.309
35	1.400	122.500
36	1.500	124.152
37	1.600	126.264
38	1.700	128.835
39	1.800	131.871
40	1.900	135.382
41	2.000	139.380

Los resultados parecen indicar que no hay necesidad de transformar la serie para estabilizar su varianza. Gráficamente:

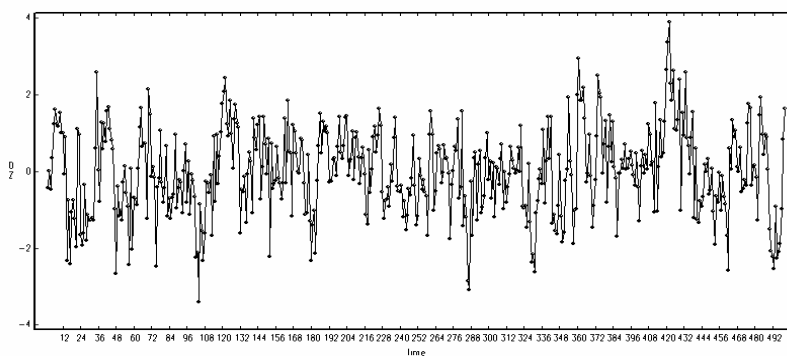


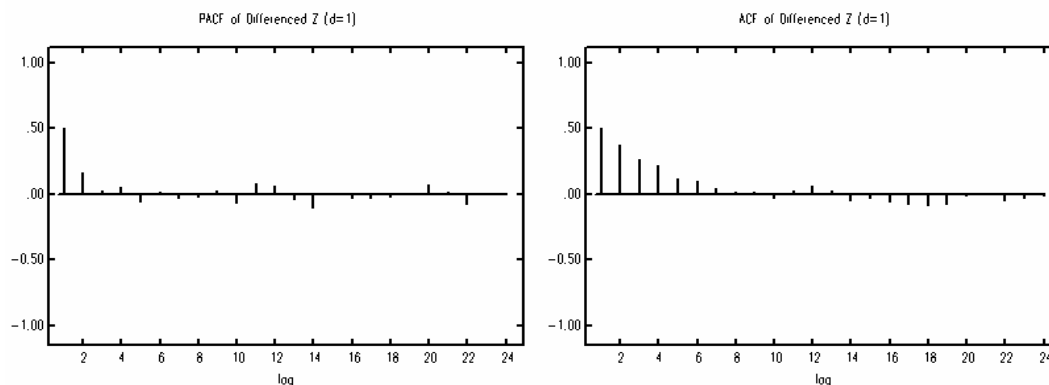
Paso 2. Examen de la ACF y PACF muestral de la serie original.



Los resultados muestran que la serie parece no ser estacionaria. Debemos diferenciar la serie una vez y observar de nuevo el gráfico de la serie diferenciada y la ACF y PACF de la serie diferenciada.

Gráfico de la serie diferenciada una vez, ΔZ_t .





Los resultados muestran que la serie $(1-B)Z_t$, parece ser estacionaria.

Prueba de la hipótesis de raíz unitaria usando la prueba de Dickey-Fuller

Se siguen los siguientes pasos:

- ✓ Búsqueda del modelo adecuado para hacer la prueba: inicialmente se estima el modelo (4)

$$\Delta z_t = \beta_0 + \beta_1 t + \gamma z_{t-1} + \sum_{i=1}^P \delta_i \Delta z_{t-i} + a_t$$

escogiendo $P = \text{Parte Entera}(12(n/100)^{1/4})$ y seleccionando el mejor modelo con el criterio SIC (se pueden usar otros criterios de información). Se obtienen los siguientes resultados:

Null Hypothesis: Z has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic based on SIC, MAXLAG=17)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.539658	0.8148
Test critical values: 1% level	-3.976629	
5% level	-3.418889	
10% level	-3.131986	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(Z)
 Method: Least Squares
 Date: 03/28/06 Time: 15:59
 Sample (adjusted): 4 500
 Included observations: 497 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Z(-1)	-0.003951	0.002566	-1.539658	0.1243
D(Z(-1))	0.429270	0.044415	9.664917	0.0000
D(Z(-2))	0.157814	0.044478	3.548131	0.0004
C	0.280658	0.147447	1.903447	0.0576
@TREND(1)	-2.23E-05	0.000358	-0.062204	0.9504
R-squared	0.284673	Mean dependent var		0.068922
Adjusted R-squared	0.278857	S.D. dependent var		1.170687
S.E. of regresión	0.994149	Akaike info criterion		2.836150
Sum squared resid	486.2595	Schwarz criterion		2.878490
Log likelihood	-699.7833	F-statistic		48.94924
Durbin-Watson stat	2.024279	Prob(F-statistic)		0.000000

De los resultados anteriores elegimos el siguiente modelo para hacer la prueba:

$$\Delta z_t = \beta_0 + \gamma z_{t-1} + \sum_{i=1}^2 \delta_i \Delta z_{t-i} + a_t$$

✓ Realización de la prueba ADF sobre el modelo elegido:

Null Hypothesis: Z has a unit root
 Exogenous: Constant
 Lag Length: 2 (Automatic based on SIC, MAXLAG=17)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.801180	0.3799
Test critical values:		
1% level	-3.443281	
5% level	-2.867136	
10% level	-2.569812	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(Z)
 Method: Least Squares
 Date: 03/28/06 Time: 16:05
 Sample (adjusted): 4 500
 Included observations: 497 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Z(-1)	-0.004029	0.002237	-1.801180	0.0723
D(Z(-1))	0.429406	0.044316	9.689550	0.0000
D(Z(-2))	0.158022	0.044307	3.566496	0.0004
C	0.279912	0.146810	1.906624	0.0572
R-squared	0.284667	Mean dependent var		0.068922
Adjusted R-squared	0.280314	S.D. dependent var		1.170687
S.E. of regression	0.993144	Akaike info criterion		2.832134
Sum squared resid	486.2634	Schwarz criterion		2.866006
Log likelihood	-699.7853	F-statistic		65.39650
Durbin-Watson stat	2.024420	Prob(F-statistic)		0.000000

Conclusión: no se puede rechazar H_0 , es decir, parece existir una raíz unitaria y por tanto la serie parece ser no estacionaria. Para seguir el análisis debemos diferenciar la serie y observar si ya es estacionaria, realizando de nuevo la prueba ADF pero sobre la serie diferenciada una vez, $Y_t = \Delta Z_t$. Para el ejemplo, el modelo apropiado para la prueba es (sin constante ni tendencia)

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + a_t$$

Resultados:

Null Hypothesis: D(Z) has a unit root
 Exogenous: None
 Lag Length: 1 (Automatic based on SIC, MAXLAG=17)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.374462	0.0000
Test critical values: 1% level	-2.569623	
5% level	-1.941462	
10% level	-1.616271	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(Z,2)

Method: Least Squares

Date: 03/28/06 Time: 16:19

Simple (adjusted): 4 500

Included observations: 497 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(Z(-1))	-0.410305	0.043768	-9.374462	0.0000
D(Z(-1),2)	-0.156948	0.044342	-3.539526	0.0004

R-squared	0.262320	Mean dependent var	-0.000741
Adjusted R-squared	0.260830	S.D. dependent var	1.157064
S.E. of regression	0.994786	Akaike info criterion	2.831437
Sum squared resid	489.8513	Schwarz criterion	2.848373
Log likelihood	-701.6122	Durbin-Watson stat	2.024014

Conclusión: El valor p de la prueba es aproximadamente 0, indicando una fuerte evidencia a favor de H_1 , es decir, las primeras diferencias parecen ser estacionarias. Este es el mismo resultado que habíamos obtenido antes usando la interpretación de la ACF y PACF.

Paso 3. Identificación de los órdenes p y q.

Para esto usamos el ACF y PACF muestral de la serie diferenciada una vez, que fueron obtenidos en el paso anterior. Los gráficos muestran que ambos parecen ser consistentes con funciones que decaen exponencialmente. Por tanto un modelo tentativo surge de suponer que los datos son generados por un proceso ARMA. Sin embargo, en este caso la ACF y PACF no son útiles para obtener a p y q. La información sobre estos valores puede ser obtenida usando la *función de autocorrelación extendida muestral* (EACF muestral de Tiao y Box (1981)) presentada a continuación.

La identificación de los órdenes p y q de un proceso ARMA: La función de autocorrelación muestral extendida (ESACF)

Para un proceso ARMA(p,q) la ACF y PACF decaen exponencialmente, lo cual hace difícil la identificación de los órdenes p y q del proceso.

Un método frecuentemente usado supone que si Z_t sigue el proceso ARMA(p, q)

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = \theta_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

o, equivalentemente,

$$Z_t = \theta_0 + \sum_{i=1}^p \phi_i Z_{t-i} + \sum_{i=1}^q \theta_i a_{t-i} + a_t$$

entonces,

$$Y_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = Z_t - \sum_{i=1}^p \phi_i Z_{t-i}$$

sigue un modelo MA(q)

$$Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

donde, sin pérdida de generalidad, $\theta_0 = 0$.

Por tanto, algunos autores como Tiao y Box (1981) sugieren usar la ACF de los residuales \hat{Y}_t resultantes de la estimación por mínimos cuadrados ordinarios (OLS) de un modelo AR(p), donde

$$\hat{Y}_t = (1 - \hat{\phi}_1 B - \hat{\phi}_2 B^2 - \dots - \hat{\phi}_p B^p) Z_t$$

para identificar a q y de ahí obtener los órdenes p y q del modelo ARMA.

Por ejemplo, un proceso MA(2) de los residuales generados por el ajuste de un AR(1), implica un modelo mixto ARMA(1,2).

Sin embargo, los estimadores OLS para los parámetros ϕ_i obtenidos cuando el término de error del modelo AR no es ruido blanco (que, en general no lo es, pues contienen la componente MA si el modelo adecuado es un ARMA(p,q)) **no son consistentes** y el procedimiento puede conducir a una identificación incorrecta.

Para derivar estimadores consistentes de los ϕ_i , suponga que se dispone de n observaciones de un proceso ARMA(p,q), donde, sin pérdida de generalidad, $\theta_0 = 0$.

Si se ajusta un modelo AR(p) a los datos, es decir si se ajusta el modelo

$$Z_t = \sum_{i=1}^p \phi_i Z_{t-i} + e_t$$

donde e_t representa el término de error, entonces los estimadores OLS $\hat{\phi}_i^0$ de los ϕ_i , $i=1, 2, \dots, p$, no son consistentes, y los residuales

$$\hat{e}_t^0 = Z_t - \sum_{i=1}^p \hat{\phi}_i^0 Z_{t-i}$$

no serán ruido blanco.

En efecto, si $q \geq 1$ los residuales rezagados \hat{e}_{t-i}^0 , $i=1, \dots, q$, contienen alguna información del proceso Z_t . Esto conduce a realizar las siguientes regresiones iteradas.

Primero, considere una regresión AR(p) más el término \hat{e}_{t-1}^0 , es decir, considere la regresión

$$Z_t = \sum_{i=1}^p \phi_i^{(1)} Z_{t-i} + \beta_1^{(1)} \hat{e}_{t-1}^{(0)} + e_t^{(1)}, \quad t = p+2, \dots, n$$

Donde el superíndice (1) hace referencia a la primera regresión iterada y $e_t^{(1)}$ representa el correspondiente término de error de la regresión. Los estimadores OLS $\hat{\phi}_i^{(1)}$ obtenidos serán consistentes si $q=1$. Sin embargo, si $q>1$ los estimadores $\hat{\phi}_i^{(1)}$ de nuevo serán inconsistentes, los residuales $\hat{e}_t^{(1)}$ no serán ruido blanco, y los valores rezagados $\hat{e}_{t-i}^{(1)}$ contendrán alguna información del proceso Z_t .

Entonces se debe considerar la segunda regresión AR(p) iterada

$$Z_t = \sum_{i=1}^p \phi_i^{(2)} Z_{t-i} + \beta_1^{(2)} \hat{e}_{t-1}^{(1)} + \beta_2^{(2)} \hat{e}_{t-2}^{(0)} + e_t^{(2)}, \quad t = p+3, \dots, n$$

Los estimadores OLS $\hat{\phi}_i^{(2)}$ serán consistentes si $q=2$. Pero si $q>2$ los estimadores $\hat{\phi}_i^{(2)}$ de nuevo serán inconsistentes. Sin embargo, se pueden obtener estimadores consistentes repitiendo la iteración anterior.

Es decir, los estimadores OLS $\hat{\phi}_i^{(q)}$ obtenidos de la q-ésima regresión AR(p) iterada serán consistentes:

$$Z_t = \sum_{i=1}^p \phi_i^{(q)} Z_{t-i} + \sum_{i=1}^q \beta_i^{(q)} \hat{e}_{t-i}^{(q-1)} + e_t^{(q)}, \quad t = p+q+1, \dots, n$$

donde $\hat{e}_t^{(j)} = Z_t - \sum_{i=1}^p \hat{\phi}_i^{(j)} Z_{t-i} - \sum_{i=1}^q \hat{\beta}_i^{(j)} \hat{e}_{t-i}^{(j-1)}$ es el residual de la j-ésima regresión AR(p) iterada y los $\hat{\phi}_i^{(j)}$ y $\hat{\beta}_i^{(j)}$ son los estimadores OLS correspondientes.

En la práctica los órdenes p y q son desconocidos y deben ser estimados. Basados en lo anterior, Tsay y Tiao (1984) sugieren un conjunto general de regresiones

iteradas e introducen el concepto de función de autocorrelación extendida muestral (ESACF) para estimar los órdenes p y q .

Específicamente, para $m=0, 1, 2, 3, \dots$, sea $\hat{\phi}_i^{(j)}$, $i=1, \dots, m$, los estimadores OLS de un regresión AR(m) iterada del proceso ARMA de Z_t . Tsay y Tiao definen la m -ésima ESACF en el rezago j , $\hat{\rho}_j^{(m)}$ como la autocorrelación para la serie transformada

$$\hat{Y}_{m,t}^{(j)} = (1 - \hat{\phi}_1^{(j)}B - \hat{\phi}_2^{(j)}B^2 - \dots - \hat{\phi}_m^{(j)}B^m)Z_t$$

Es decir, $\hat{\rho}_j^{(m)} = \hat{\rho}_j(\hat{Y}_{m,t}^{(j)})$.

En la práctica es útil colocar los $\hat{\rho}_j^{(m)}$ en una tabla de dos entradas como la siguiente.

AR	MA					
	0	1	2	3	4	...
0	$\hat{\rho}_1^{(0)}$	$\hat{\rho}_2^{(0)}$	$\hat{\rho}_3^{(0)}$	$\hat{\rho}_4^{(0)}$	$\hat{\rho}_5^{(0)}$...
1	$\hat{\rho}_1^{(1)}$	$\hat{\rho}_2^{(1)}$	$\hat{\rho}_3^{(1)}$	$\hat{\rho}_4^{(1)}$	$\hat{\rho}_5^{(1)}$...
2	$\hat{\rho}_1^{(2)}$	$\hat{\rho}_2^{(2)}$	$\hat{\rho}_3^{(2)}$	$\hat{\rho}_4^{(2)}$	$\hat{\rho}_5^{(2)}$...
3	$\hat{\rho}_1^{(3)}$	$\hat{\rho}_2^{(3)}$	$\hat{\rho}_3^{(3)}$	$\hat{\rho}_4^{(3)}$	$\hat{\rho}_5^{(3)}$...
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	...

La primera fila de la tabla anterior corresponde a la ACF muestral de la serie Z_t , la segunda fila da la primera ESACF $\hat{\rho}_j^{(1)}$, etc. Las filas son numeradas como 0, 1, 2, 3, ... para especificar el orden de la parte AR del modelo, y las columnas son numeradas en forma similar para el orden de la parte MA.

Tsay y Tiao (1984) mostraron que para un proceso ARMA(p,q), se da la siguiente convergencia en probabilidad,

$$\hat{\rho}_j^{(m)} \xrightarrow{p} \begin{cases} 0 & \text{si } 0 \leq m-p \leq j-q \\ X \neq 0 & \text{en otro caso} \end{cases}$$

Por tanto, la tabla de la ESACF asintótica para un modelo ARMA(1,1) es

AR	MA					
	0	1	2	3	4	...
0	X	X	X	X	X	...
1	X	0	0	0	0	...
2	X	X	0	0	0	...
3	X	X	X	0	0	...
4	X	X	X	X	0	...
⋮	⋮	⋮	⋮	⋮	⋮	...

Se puede observar que los ceros forman un triángulo con vértice en la posición (1,1) de la tabla.

En general, para un proceso ARMA(p,q) el vértice de un triángulo de ceros en la ESACF asintótica estará en la posición (p,q). Por tanto, la ESACF es una herramienta útil para identificar los órdenes del modelo ARMA(p,q).

Sin embargo, en la práctica, tenemos muestras finitas y los $\hat{\rho}_j^{(m)}$ para $0 \leq m-p \leq j-q$ pueden no ser exactamente cero. Para verificar las estimaciones no son estadísticamente significativas, la varianza asintótica de $\hat{\rho}_j^{(m)}$ puede ser aproximada por medio de $(n-m-j)^{-1}$ bajo la hipótesis de que la serie transformada $Y_t^{(j)}$ es ruido blanco. Usando este resultado, la ESACF puede ser construida usando

símbolos indicadores X cuando los valores de $|\hat{\rho}_j^{(m)}| > 2$ error estándar, y de 0 para $|\hat{\rho}_j^{(m)}| \leq 2$ error estándar.

Ejemplo.

Para los datos de la serie simulada anterior, su ESACF es la siguiente.

```

DIFFERENCE ORDERS. . . . . (1-B)
NAME OF THE SERIES . . . . . Z
TIME PERIOD ANALYZED . . . . . 1 TO 500
MEAN OF THE (DIFFERENCED) SERIES . . . -0.0041
STANDARD DEVIATION OF THE SERIES . . . 1.1378
T-VALUE OF MEAN (AGAINST ZERO) . . . -0.0806

```

THE EXTENDED ACF TABLE

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	.51	.38	.26	.22	.12	.10	.04	.02	.02	-.03	.03	.06	.03
(P= 1)	-.29	.03	-.06	.11	-.08	.06	.02	-.03	.02	-.01	.03	.06	.02
(P= 2)	-.16	-.15	.00	.07	.02	.07	.04	.01	.03	.02	.04	-.00	.02
(P= 3)	-.43	-.15	.01	.06	-.02	.07	.02	.00	.03	.02	.03	.02	.01
(P= 4)	.48	.33	.27	.16	.01	-.02	.01	.03	.03	-.05	.02	.02	.01
(P= 5)	.24	-.40	-.12	.18	.04	-.01	.00	.04	.03	-.07	.03	.02	.00
(P= 6)	.40	-.28	-.19	.06	.03	.00	.01	.03	.03	.02	.03	.03	-.00

SIMPLIFIED EXTENDED ACF TABLE (5% LEVEL)

(Q-->)	0	1	2	3	4	5	6	7	8	9	10	11	12
(P= 0)	X	X	X	X	O	O	O	O	O	O	O	O	O
(P= 1)	X	O	O	X	O	O	O	O	O	O	O	O	O
(P= 2)	X	X	O	O	O	O	O	O	O	O	O	O	O
(P= 3)	X	X	O	O	O	O	O	O	O	O	O	O	O
(P= 4)	X	X	X	X	O	O	O	O	O	O	O	O	O
(P= 5)	X	X	X	X	O	O	O	O	O	O	O	O	O
(P= 6)	X	X	X	O	O	O	O	O	O	O	O	O	O

Si una autocorrelación es significativa, en la segunda tabla su entrada es X. En otro caso, su entrada es 0. Para la identificación de los órdenes p y q, busque el vértice de un triángulo de ceros. Podemos identificar una región triangular que define a p y q, la cual emana del vértice p=1 y q=1. Como resultado, parece que el modelo ARIMA(1,1,1) es consistente con los datos analizados.

Paso 4.

Finalmente, contrastamos la existencia de un término de tendencia determinística, θ_0 . Para esto contrastamos la hipótesis $H_0: \theta_0=0$, usando el estadístico $t=\bar{W} / S_{\bar{W}}$. De la tabla anterior el valor de t es -0.0806, lo cual nos conduce a no rechazar H_0 , es decir parece que no existe dicho término de tendencia determinística. Otra forma de verificar la importancia θ_0 es introducir el parámetro en la especificación del modelo, estimarlo y verificar su significancia estadística. La observación de la serie original indica que no parece existir una tendencia lineal determinística, además de la tendencia aleatoria.

Otras Formas de identificación del Modelo ARMA

Existen otras técnicas que ayudan a la identificación del modelo adecuado para una serie de tiempo.

- ✓ Basadas en criterios de información: Suponga que un modelo con M parámetros es ajustado a los datos. Defina,

⇒ El Criterio de información de Akaike,

$$AIC(M)=n \ln \hat{\sigma}_a^2 + 2M$$

⇒ El Criterio de Bayesiano de información de Akaike,

$$BIC(M)=n \ln \hat{\sigma}_a^2 - (n - M) \ln \left(1 - \frac{M}{n}\right) + M \ln(n) + M \ln \left[\left(\frac{\hat{\sigma}_Z^2}{\hat{\sigma}_a^2} \right) / M \right]$$

donde $\hat{\sigma}_Z^2$ es la varianza muestral de la serie.

⇒ El criterio de información bayesiano de Schwartz,

$$SBC(M) = n \ln \sigma_a^2 + M \ln(n)$$

Usando los criterios anteriores, el orden óptimo del modelo es elegido como aquel donde se minimizan dichos criterios.

- ✓ Basada en la *función de correlación cruzada muestral* entre Z_t y los residuales de un modelo preliminar ajustado a los datos (Castaño y Martínez, 2008). Ver documento.

Ejemplos. Proceso de Identificación en las Series W1 a W7 del Texto de Wei. Vea macro de R, wei_data.r

- **Estimación**

Agregando el supuesto de **normalidad** en los errores a_t del modelo, los parámetros pueden ser estimados usando la técnica de máxima verosimilitud. La función de verosimilitud puede ser razonablemente aproximada por medio de una función de verosimilitud condicional (Box-Jenkins, 1970). Sin embargo, puede adoptarse una aproximación exacta a la función de verosimilitud (Hillmer y Tiao, 1979). Dadas las n observaciones Z_1, Z_2, \dots, Z_n , ambas aproximaciones calculan la función verosimilitud sobre la base de la estructura aleatoria de n observaciones,

$$\dot{Z}_t = \sum_{j=1}^p \phi_j \dot{Z}_{t-j} - \sum_{i=1}^q \theta_i a_{t-i} + a_t \quad \text{para } t=1, \dots, n.$$

donde $\dot{Z}_t = Z_t - \mu$ y $\{a_t\}$ es ruido blanco $N(0, \sigma_a^2)$.

Entonces, el logaritmo de la función de verosimilitud es:

$$\ln L(\phi, \mu, \theta, \sigma_a^2) = -n \ln(2\pi \sigma_a^2)/2 - \sum_{t=1}^n a_t^2$$

donde $a_t = \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} + \dot{Z}_t - \phi_1 \dot{Z}_{t-1} - \phi_2 \dot{Z}_{t-2} - \dots - \phi_p \dot{Z}_{t-p}$.

Asumimos las condiciones iniciales $Z^* = (Z_{1-p}, \dots, Z_{-1}, Z_0)'$ y $a^* = (a_{1-q}, \dots, a_{-1}, a_0)'$.

Los dos métodos difieren en que el algoritmo de la verosimilitud condicional asume que $a_p = a_{p-1} = \dots = a_{p+1-q} = 0$, y se calculan las a_t para $t \geq p+1$, mientras que el algoritmo de la verosimilitud exacta obtiene la verdadera función de verosimilitud. Otro método alternativo es llamado el método de Backcasting o de verosimilitud incondicional, el cual consiste en pronosticar hacia atrás los valores desconocidos $Z^* = (Z_{1-p}, \dots, Z_{-1}, Z_0)'$ y $a^* = (a_{1-q}, \dots, a_{-1}, a_0)'$ necesarios para calcular la función de verosimilitud.

La verosimilitud exacta e incondicional es recomendada cuando existen términos de medias móviles en el modelo, para modelos estacionales, para modelos que están cerca de la no estacionaridad y especialmente para series cortas. Los tres métodos son equivalentes para muestras grandes.

Cuando el modelo tiene componentes MA, el proceso de estimación es no lineal y generalmente usa el método de Gauss-Marquardt para obtener las estimaciones de mínimos cuadrados no lineales. La solución generalmente debe ser verificada usando diferentes conjuntos de valores iniciales para los parámetros.

La manera más eficiente de usar el método de estimación exacto es primero estimar el modelo usando el método de estimación condicional. A continuación se reestima el modelo usando el método de estimación exacto, empleando como valores iniciales las estimaciones obtenidas en el primer método. La ventaja de hacerlo así es que el método condicional arroja buenos valores iniciales para el método exacto, logrando generalmente el óptimo global.

Estimación de la varianza σ_a^2 , del proceso de ruido blanco: Una vez los parámetros del modelo han sido estimados, la estimación máximo verosímil de σ_a^2 puede ser obtenida como:

$$\hat{\sigma}_a^2 = \sum_{t=d+p+1}^n \hat{a}_t^2 / n$$

donde \hat{a}_t es el t-ésimo residual del modelo.

Ejemplo. Estimación del modelo para la serie simulada ARIMA(1,1,1)

Modelo a estimar: $(1-\phi_1 B)(1-B)^d Z_t = \theta_0 + (1-\theta_1 B) a_t$

Resultados:

Estimación condicional: por defecto toma como valores iniciales 0.1 para los parámetros AR y MA y de 1 para la constante

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 500
NONLINEAR ESTIMATION TERMINATED DUE TO:
RELATIVE CHANGE IN (OBJECTIVE FUNCTION)**0.5 LESS THAN 0.1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- MODELO1

VARIABLE		TYPE OF VARIABLE		ORIGINAL OR CENTERED		DIFFERENCING		
Z		RANDOM		ORIGINAL		1 (1-B)		
PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRAINT	VALUE	STD ERROR	T VALUE
1 CONST		CNST	1	0	NONE	.0037	.1075	.03
2 THETA1	Z	MA	1	1	NONE	.3043	.0808	3.77
3 PHI1	Z	D-AR	1	1	NONE	.7297	.0587	12.43

EFFECTIVE NUMBER OF OBSERVATIONS . . . 498
R-SQUARE 0.994
RESIDUAL STANDARD ERROR. 0.967875E+00

Estimación exacta: toma como valores iniciales las estimaciones obtenidas por el método condicional.

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 500
 NONLINEAR ESTIMATION TERMINATED DUE TO:
 RELATIVE CHANGE IN THE STANDARD ERROR LESS THAN 0.1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- MODELO1

VARIABLE		TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING					
Z		RANDOM	ORIGINAL	1 (1-B)					

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRRAINT	VALUE	STD ERROR	T VALUE	
1	CONST		CNST	1	0	NONE	.0038	.1117	.03
2	THETA1	Z	MA	1	1	NONE	.3011	.0783	3.84
3	PHI1	Z	D-AR	1	1	NONE	.7277	.0563	12.93
EFFECTIVE NUMBER OF OBSERVATIONS				. .		498			
R-SQUARE					0.994			
RESIDUAL STANDARD ERROR.				0.967867E+00				
--									

Los resultados son muy similares (efecto del tamaño muestral grande). Se observa que la constante no es significativa, como lo habíamos detectado usando otro procedimiento en la fase de identificación.

- **Diagnósticos o verificación del modelo**

La etapa final en la construcción del modelo consiste en validar el modelo estimado. Se trata de responder al pregunta: Es el modelo estadísticamente consistente con los supuestos teóricos sobre los que se basa?

Algunos chequeos importantes:

i) Verificación de la estacionaridad e invertibilidad del modelo estimado: Obtención de las raíces de los polinomios AR y MA y verificación que sus módulos sean mayores que 1. La sobrediferenciación puede conducir a encontrar raíces unitarias (o cercanas a la unidad) en el polinomio MA del modelo estimado. La subdiferenciación puede conducir a encontrar raíces unitarias (o cercanas a la unidad) en el polinomio AR.

Para el ejemplo anterior, este chequeo produce los siguientes resultados.

ALL ROOTS OF THE ARMA POLYNOMIALS LIE OUTSIDE THE UNIT CIRCLE
ROOT(S) FOR THE FOLLOWING AR OR MA FACTOR:

1
1 - (0.73)B

PAIR	REAL	IMAGINARY	DISTANCE
1	1.374279	0.000000	1.374279

ROOT(S) FOR THE FOLLOWING AR OR MA FACTOR:

1
1 - (0.30)B

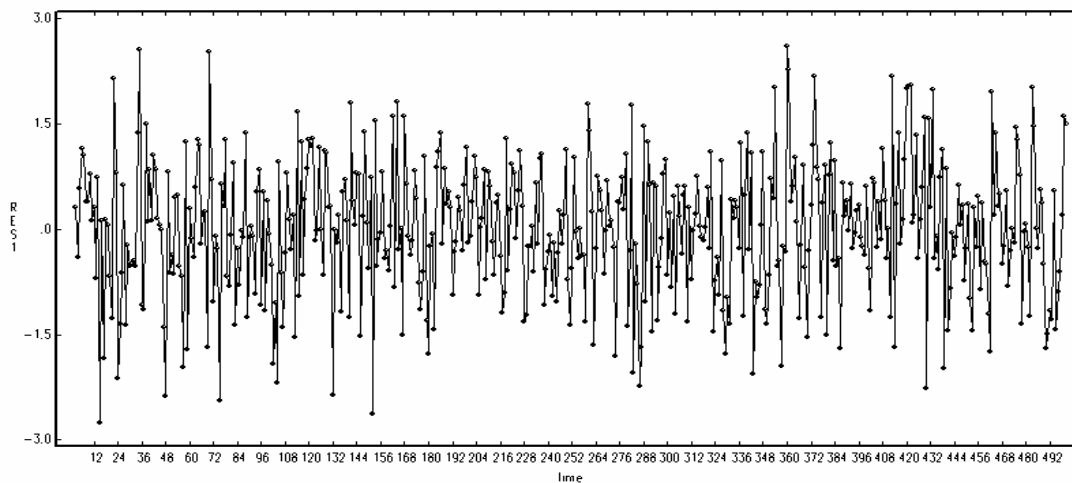
PAIR	REAL	IMAGINARY	DISTANCE
1	3.321306	0.000000	3.321306

Los resultados muestran que el modelo estimado es estacionario e invertible.

ii) Verificación de que los residuales del modelo son ruido blanco, es decir, son no correlacionados y su varianza es constante.

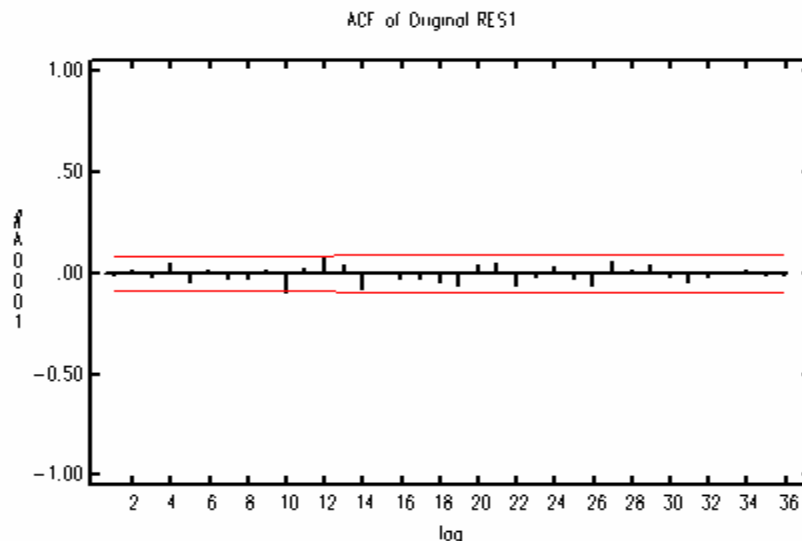
➤ Grafique los residuales contra el tiempo

Resultados



Los residuales parecen provenir de un proceso con media y varianza constante, pues el gráfico no presenta patrones sistemáticos a abrirse o a comprimirse.

➤ Obtenga la ACF de los residuales:



- Calcule el estadístico de Ljung-Box, $Q = n(n+2) \sum_{j=1}^K (n-j)^{-1} \hat{\rho}_j^2(\hat{a})$, que sirve para contrastar la hipótesis nula $H_0: \rho_1(a) = \rho_2(a) = \dots = \rho_K(a) = 0$, donde n es el número de residuales del modelo, los $\hat{\rho}_j^2(\hat{a})$ son los coeficientes de autocorrelación estimados usando la serie de residuales. Rechace H_0 si $Q > \chi(\alpha, K-m)$, donde m es el número de parámetros ARMA estimados en el modelo y $\chi(\alpha, K-m)$ es el percentil α superior de una distribución Chi-cuadrado con $K-m$ grados de libertad.

Resultados

```
NAME OF THE SERIES . . . . . RES1
TIME PERIOD ANALYZED . . . . . 3 TO 500
MEAN OF THE (DIFFERENCED) SERIES . . . 0.0002
STANDARD DEVIATION OF THE SERIES . . . 0.9679
T-VALUE OF MEAN (AGAINST ZERO) . . . . 0.0041
```

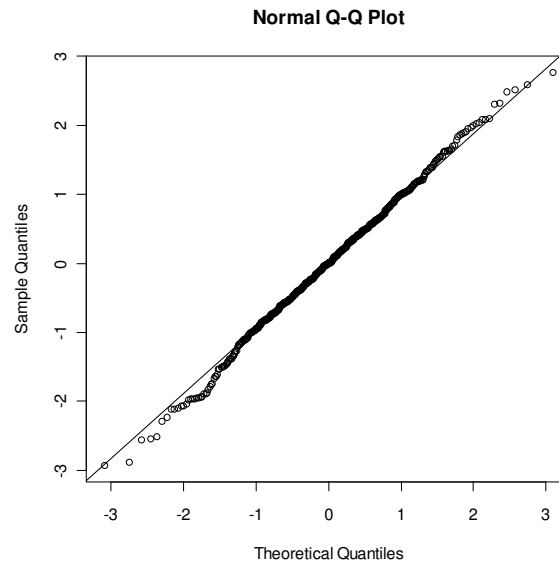
AUTOCORRELATIONS

1- 12	-.00	.01	-.02	.06	-.04	.02	-.02	-.02	.01	-.10	.02	.08
ST.E.	.04	.04	.04	.04	.04	.05	.05	.05	.05	.05	.05	.05
Q	.0	.1	.3	1.8	2.8	3.0	3.3	3.5	3.6	8.4	8.6	12.2
13- 24	.04	-.08	.01	-.03	-.03	-.04	-.06	.04	.05	-.06	-.02	.03
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
Q	13.2	16.2	16.2	16.6	16.9	17.9	19.7	20.6	21.9	23.6	23.7	24.2
25- 36	-.03	-.06	.06	.01	.04	-.02	-.05	-.02	-.00	.02	-.01	-.01
ST.E.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
Q	24.7	26.7	28.5	28.6	29.7	29.9	31.0	31.2	31.2	31.4	31.5	31.5

Por ejemplo, para $K=23$, $Q=23.7$ y para un nivel de significación aproximado de $\alpha=0.05$, $\chi(\alpha, K-m) = \chi(0.05, 21)=32.67$. Como $Q<32.67$, no rechazamos H_0 , es decir el modelo parece ser adecuado.

iii) Verificación de la normalidad del ruido blanco a_t

Es necesaria para la validez de los contrastes de hipótesis, tales como la significancia de los parámetros. La normalidad puede verificarse sobre los residuales del modelo a través del gráfico cuantil-cuantil normal. Para el ejemplo,



Si hay normalidad, el gráfico mostrará una línea recta de 45° , si los residuales están estandarizados. Desviaciones sistemáticas de este patrón indican la violación del supuesto de normalidad. Adicionalmente se pueden emplear otros contrastes de normalidad, tal como el contraste de Jarque-Bera.

En general, si x_1, \dots, x_T una muestra aleatoria de tamaño T de una variable aleatoria X . Entonces, recuerde que:

La media muestral es
$$\hat{\mu}_X = \frac{1}{T} \sum_{t=1}^T x_t$$

La varianza muestral es $\sigma_x^2 = \frac{1}{T-1} \sum_{t=1}^T (x_t - \mu_x)^2$

El coeficiente de asimetría muestral es $\hat{S}(x) = \frac{1}{(T-1)\sigma_x^3} \sum_{t=1}^T (x_t - \mu_x)^3$

El coeficiente de curtosis muestral es $\hat{K}(x) = \frac{1}{(T-1)\sigma_x^4} \sum_{t=1}^T (x_t - \mu_x)^4$

Si la muestra procede de una distribución normal, entonces cuando $T \rightarrow \infty$

$$\hat{S}(x) \xrightarrow{d} N(0, 6/T)$$

$$\hat{K}(x) \xrightarrow{d} N(0, 24/T)$$

Bajo el supuesto de normalidad, el verdadero coeficiente de asimetría, $S(x)$, es 0 y el de curtosis, $K(x)$, es 3. Las propiedades asintóticas anteriores pueden ser usadas para probar normalidad en los residuales del modelo ajustado:

Para probar asimetría se considera la hipótesis $H_0 : S(x) = 0$ contra la alternativa $H_1 : S(x) \neq 0$. El estadístico de la prueba es

$$t = \frac{\hat{S}(x)}{\sqrt{6/T}}$$

Se rechaza H_0 a un nivel de significación α si $|t| > z(\alpha/2)$, donde $z(\alpha/2)$ es el cuantil $\alpha/2$ – superior de la distribución $N(0,1)$.

Para probar exceso de curtosis se considera la hipótesis $H_0 : K(x) - 3 = 0$ contra la alternativa $H_1 : K(x) - 3 \neq 0$. El estadístico de la prueba es

$$t = \frac{\hat{K}(x) - 3}{\sqrt{24/T}}$$

Se rechaza H_0 a un nivel de significación α si $|t| > z(\alpha/2)$.

Para probar asimetría y curtosis simultáneamente, Jarque y Bera (1987) combinan las dos pruebas anteriores y usan el estadístico

$$JB = \frac{\hat{S}^2(x)}{6/T} + \frac{(\hat{K}(x) - 3)^2}{24/T} \xrightarrow{d} \chi^2_2$$

Se rechaza el supuesto de normalidad si $JB > \chi^2_2(\alpha)$, donde $\chi^2_2(\alpha)$ es el cuantil α -superior de la distribución χ^2_2 .

Para los residuales del ejemplo, $JB = 0.87$. Para un nivel de significancia de 0.05, $\chi^2_2(0.05) = 5.99$. No se rechaza el supuesto de normalidad.

iv) Verificación de la significancia individual de los parámetros estimados:

Para contrastar la significancia individual de cada parámetro ϕ_j , $j=1,2,\dots,p$, y θ_i , $i=1,2,\dots,q$ del modelo, se construye el estadístico t definido como:

$t = (\text{estimación del parámetro}) / (\text{error estándar de la estimación del parámetro})$

Con el valor de t rechaza H_0 si $|t| > z_{\alpha/2}$, donde $z_{\alpha/2}$ es el percentil $\alpha/2$ superior de la distribución normal(0,1).

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 500

NONLINEAR ESTIMATION TERMINATED DUE TO:
RELATIVE CHANGE IN THE STANDARD ERROR LESS THAN 0.1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- MODELO1

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING					
Z	RANDOM	ORIGINAL	1 (1-B)					
PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS-TRAI NT	VALUE	STD ERROR	T VALUE
1 CONST		CNST	1	0	NONE	.0038	.1117	.03
2 THETA1	Z	MA	1	1	NONE	.3011	.0783	3.84
3 PHI1	Z	D-AR	1	1	NONE	.7277	.0563	12.93

Observamos que el valor t correspondiente a la constante es muy pequeño (menor que 2 como regla general) y por tanto la constante (que corresponde a θ_0 no es significativa. Con este resultado deberíamos ajustar de nuevo el modelo y realizar de nuevo todos los diagnósticos.

Excluyendo la constante, el nuevo modelo estimado es:

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 500

NONLINEAR ESTIMATION TERMINATED DUE TO:
RELATIVE CHANGE IN THE STANDARD ERROR LESS THAN 0.1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- MODELO2

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING					
Z	RANDOM	ORIGINAL	1 (1-B)					
PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS-TRAI NT	VALUE	STD ERROR	T VALUE
1 THETA1	Z	MA	1	1	NONE	.3011	.0784	3.84
2 PHI1	Z	D-AR	1	1	NONE	.7276	.0564	12.90

EFFECTIVE NUMBER OF OBSERVATIONS 498
R-SQUARE 0.994
RESIDUAL STANDARD ERROR. 0.967868E+00

Recuerde que la serie fue simulada de un ARIMA(1,1,1), para $\theta_1=0.3$, $\phi_1=0.7$ y el término de error tiene una distribución $N(0,1)$.

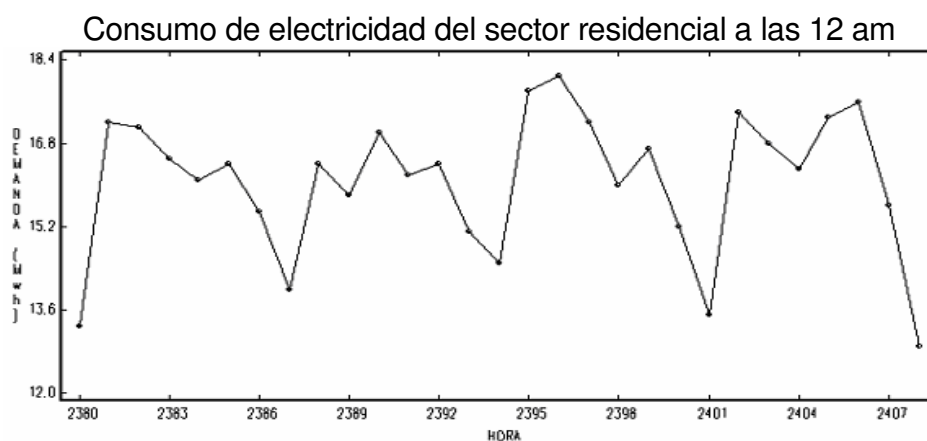
Ejemplos empíricos. Estimación para las series W1-W7 de Wei. Ver resultados de la estimación en la macro Wei_data2.r y diagnósticos en Wei_data3.r

5. Análisis de Series Estacionales.

Muchas series de tiempo económicas y financieras contienen una componente estacional que se repite después de un período regular de tiempo.

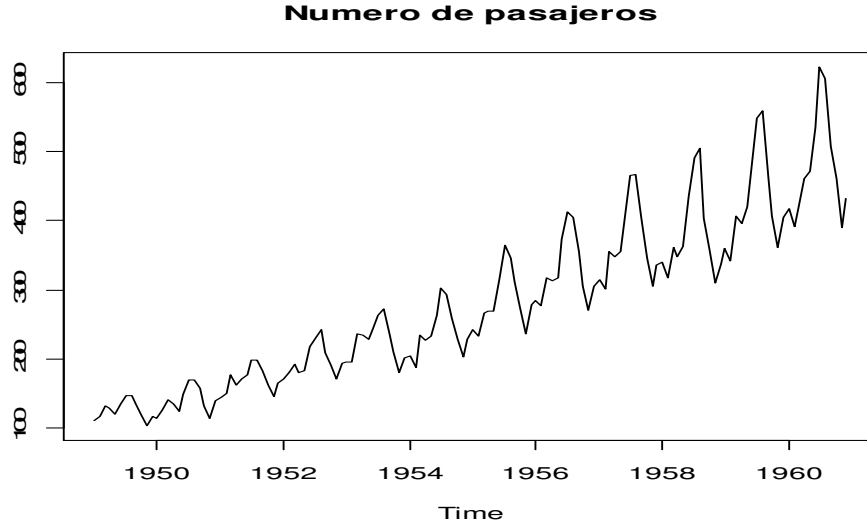
Ejemplos:

a) Considere la serie del consumo de electricidad del sector residencial a las 12 am, cuyo gráfico para 4 semanas es el siguiente:



La gráfica muestra que existen fluctuaciones que se repiten semanalmente, indicando que el consumo a esta hora no solamente está relacionado día a día, sino que también está relacionado semana a semana.

b) Considere el siguiente gráfico con las cifras del número de pasajeros en vuelos internacionales (en miles) para el período enero de 1949 a diciembre de 1960. En este caso el fenómeno estacional se repite año a año.



En general, suponga que una serie Z_t es una serie estacional con período de estacionalidad S . El período estacional está definido como la extensión de cada onda o fluctuación. En el caso del ejemplo del consumo de electricidad a las 12am, $S=7$. Para el ejemplo del número de pasajeros, $S=12$.

Para analizar esta clase de series es útil arreglar los datos en una tabla de la siguiente forma (denominada tabla de Buys-Ballot):

	1	2	3	...	S	Total	Promedio
	Z_1	Z_2	Z_3	...	Z_S	$T_{\cdot 1}$	$\bar{Z}_{\cdot 1}$
	Z_{S+1}	Z_{S+2}	Z_{S+3}	...	Z_{2S}	$T_{\cdot 2}$	$\bar{Z}_{\cdot 2}$
	\vdots	\vdots	\vdots	...	\vdots	\vdots	\vdots
	$Z_{(n-1)S+1}$	$Z_{(n-1)S+2}$	$Z_{(n-1)S+3}$...	Z_{nS}	$T_{\cdot n}$	$\bar{Z}_{\cdot n}$
Total	$T_{1\cdot}$	$T_{2\cdot}$	$T_{3\cdot}$...	$T_{S\cdot}$	T	T/S
Promedio	$\bar{Z}_{1\cdot}$	$\bar{Z}_{2\cdot}$	$\bar{Z}_{3\cdot}$...	$\bar{Z}_{S\cdot}$	T/n	T/nS

Donde:

$T_{j\cdot}$ es el total para la j-ésima estación

$\bar{Z}_{j\cdot}$ es el promedio para la j-ésima estación

$T_{.j}$ es el total para el j-ésimo período

$\bar{Z}_{.j}$ es el promedio para el j-ésimo período

T es el gran total de la serie.

- **Métodos tradicionales**

Convencionalmente, se considera que las series de tiempo están compuestas de una mezcla de componentes de tendencia-ciclo (denotada como P_t), estacional (denotada como S_t) e irregular (denotada como ε_t). Si se puede considerar que estas componentes son aditivas, entonces la serie de tiempo puede ser escrita como

$$Z_t = P_t + S_t + \varepsilon_t$$

Para estimar las componentes se han propuesto varios métodos.

El método de Regresión

En este método, una serie de tiempo **estacional aditiva** puede ser escrita como el modelo de regresión

$$\begin{aligned} Z_t &= P_t + S_t + \varepsilon_t \\ &= \alpha_0 + \sum_{i=1}^m \alpha_i U_{it} + \sum_{j=1}^k \beta_j V_{jt} + \varepsilon_t \end{aligned}$$

donde $P_t = \alpha_0 + \sum_{i=1}^m \alpha_i U_{it}$ y las U_{it} son variables de ciclo-tendencia y

$S_t = \sum_{j=1}^k \beta_j V_{jt}$ donde las V_{jt} son variables estacionales.

Por ejemplo, una componente de tendencia-ciclo lineal P_t puede ser escrita como

$$P_t = \alpha_0 + \alpha_1 t$$

En forma general, una componente de tendencia- ciclo puede ser escrita como un polinomio de orden m en el tiempo

$$P_t = \alpha_0 + \sum_{i=1}^m \alpha_i t^i$$

La componente estacional S_t puede ser escrita como:

a) Una combinación lineal de variables indicadoras (dummy),

$$S_t = \sum_{j=1}^{S-1} \beta_j D_{jt}$$

donde $D_{jt} = 1$ si t corresponde a la j -ésima estación y $D_{jt} = 0$ en otro caso. Observe que cuando el período estacional es S , solamente se necesitan $S-1$ variables indicadoras. En otras palabras, a β_S se le asigna el valor de 0 de modo que los coeficientes β_j , $j \neq S$, representan el efecto estacional de la j -ésima estación comparado con la estación S .

b) Alternativamente la componente estacional S_t puede ser escrita como

$$S_t = \sum_{j=1}^{[S/2]} \left[\beta_j \sin\left(\frac{2\pi jt}{S}\right) + \gamma_j \cos\left(\frac{2\pi jt}{S}\right) \right]$$

donde $[S/2]$ indica la parte entera de $S/2$.

Con esta especificación, el modelo $Z_t = \alpha_0 + \sum_{i=1}^m \alpha_i U_{it} + \sum_{j=1}^k \beta_j V_{jt} + \varepsilon_t$ es

$$Z_t = \alpha_0 + \sum_{i=1}^m \alpha_i t^i + \sum_{j=1}^{[S/2]} \left[\beta_j \sin\left(\frac{2\pi jt}{S}\right) + \gamma_j \cos\left(\frac{2\pi jt}{S}\right) \right] + \varepsilon_t$$

Para un conjunto de datos Z_1, Z_2, \dots, Z_n , y valores dados de m y S , puede usar el método de los mínimos cuadrados para obtener las estimaciones $\hat{\alpha}_i$, $\hat{\beta}_j$ y $\hat{\gamma}_j$ de los parámetros α_i , β_j y γ_j .

Las estimaciones de las componentes P_t , S_t y ε_t son, para el caso de variables indicadoras estacionales

$$\hat{P}_t = \hat{\alpha}_0 + \sum_{i=1}^m \hat{\alpha}_i t^i$$

$$\hat{S}_t = \sum_{j=1}^{S-1} \hat{\beta}_j D_{jt}$$

$$\hat{\varepsilon}_t = Z_t - \hat{P}_t - \hat{S}_t$$

o, para el caso de variables trigonométricas estacionales,

$$\hat{P}_t = \hat{\alpha}_0 + \sum_{i=1}^m \hat{\alpha}_i t^i$$

$$\hat{S}_t = \sum_{j=1}^{\lfloor S/2 \rfloor} \left[\hat{\beta}_j \sin\left(\frac{2\pi jt}{S}\right) + \hat{\gamma}_j \cos\left(\frac{2\pi jt}{S}\right) \right]$$

$$\hat{\varepsilon}_t = Z_t - \hat{P}_t - \hat{S}_t$$

- **Modelos ARIMA estacionales**

Los métodos tradicionales anteriores se basan en el supuesto de que la componente estacional es determinística e independiente de las otras componentes no estacionales. Sin embargo, en muchas series de tiempo estos supuestos no se cumplen. Lo más probable es que la componente estacional sea aleatoria y correlacionada con las otras componentes no estacionales.

Considere el ejemplo del número de pasajeros y la tabla Buys-Ballot,

Año	Enero	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total	Prom
1949	112	118	132	129	121	135	148	148	136	119	104	118	1520	126.67
1950	115	126	141	135	125	149	170	170	158	133	114	140	1676	139.67
1951	145	150	178	163	172	178	199	199	184	162	146	166	2042	170.17
1952	171	180	193	181	183	218	230	242	209	191	172	194	2364	197.00
1953	196	196	236	235	229	243	264	272	237	211	180	201	2700	225.00
1954	204	188	235	227	234	264	302	293	259	229	203	229	2867	238.92
1955	242	233	267	269	270	315	364	347	312	274	237	278	3408	284.00
1956	284	277	317	313	318	374	413	405	355	306	271	306	3939	328.25
1957	315	301	356	348	355	422	465	467	404	347	305	336	4421	368.42
1958	340	318	362	348	363	435	491	505	404	359	310	337	4572	381.00
1959	360	342	406	396	420	472	548	559	463	407	362	405	5140	428.33
1960	417	391	419	461	472	535	622	606	508	461	390	432	5714	476.17
Total	2901	2820	3242	3205	3262	3740	4216	4213	3629	3199	2794	3142	40363	3363.58
Prom	241.75	235.00	270.17	267.08	271.83	311.67	351.33	351.08	302.42	266.58	232.83	261.83	280.30	517.47

La tabla indica que las cifras del número de pasajeros no solamente se relacionan mes a mes sino que también están relacionadas año a año. Esto implica que para pronosticar el número de pasajeros para el mes de diciembre, debemos examinar no solamente el número de pasajeros en los meses vecinos tales como noviembre y enero, sino también el número de pasajeros en el mes de diciembre de los años anteriores.

En general, la tabla de Buys-Ballot indica que la serie Z_t posee relaciones dentro y entre períodos. Las relaciones dentro de los períodos representan las correlaciones entre $\dots, Z_{t-2}, Z_{t-1}, Z_t, Z_{t+1}, Z_{t+2}, \dots$, y las relaciones entre los períodos representan las correlaciones entre $\dots, Z_{t-2S}, Z_{t-S}, Z_t, Z_{t+S}, Z_{t+2S}, \dots$.

Derivación del modelo ARIMA estacional

Suponga que no sabemos que la serie Z_t contiene una componente estacional y ajustamos un modelo ARIMA a la serie, es decir,

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)b_t \quad (\text{A})$$

Entonces, obviamente la serie b_t no será un proceso de ruido blanco, puesto que contiene las correlaciones entre períodos (estacionales) que no se tuvieron en cuenta. Suponga que b_t es un proceso estacionario y sea

$$\rho_{j(s)} = \frac{E(b_{t-js} - \mu_b)(b_t - \mu_b)}{\sigma_b^2}, \quad j = 1, 2, 3, \dots$$

la función de autocorrelación de b_t , la cual representa las relaciones entre períodos que no se tuvieron en cuenta. Entonces, para un proceso no estacionario b_t estas relaciones también se pueden representar por un modelo ARIMA de la forma

$$\Phi_p(B^s)(1-B^s)^D b_t = \Theta_q(B^s)a_t \quad (B)$$

donde

$$\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$$

y

$$\Theta_q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs}$$

son polinomios en B^s sin factores comunes. Los ceros de estos polinomios caen fuera del círculo unidad y $\{a_t\}$ es un proceso de ruido blanco.

Combinando las ecuaciones (A) y (B), se obtiene el modelo estacional multiplicativo de Box y Jenkins,

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D \dot{Z}_t = \theta_q(B)\Theta_q(B^s)a_t \quad (C)$$

donde,

$$\dot{Z}_t = \begin{cases} Z_t - \mu & \text{si } d = D = 0 \\ Z_t & \text{en otro caso} \end{cases}$$

A $\phi_p(B)$ y $\theta_q(B)$ se les llama los polinomios (o factores) autorregresivos y de medias móviles *regulares u ordinarios*, respectivamente, y a $\Phi_p(B^S)$ y $\Theta_q(B^S)$ se les llama los polinomios (o factores) autorregresivos y de medias móviles *estacionales*, respectivamente.

$(1-B)$ es la primera diferencia regular u ordinaria y $(1-B^S)$ es la primera diferencia estacional.

Notación: El modelo en (C) se acostumbra denotar como un modelo $ARIMA(p,d,q)X(P,D,Q)_S$

- **Comportamiento de la ACF y PACF de modelos estacionales**

Son más complicadas que para los modelos ordinarios.

Ejemplo. El proceso $ARIMA(0,0,0)X(1,0,0)_S$ con $S=12$.

Corresponde a un proceso estacional puro autorregresivo de orden 1. Su especificación es,

$$(1 - \Phi B^{12})Z_t = a_t$$

Su función de autocorrelación es

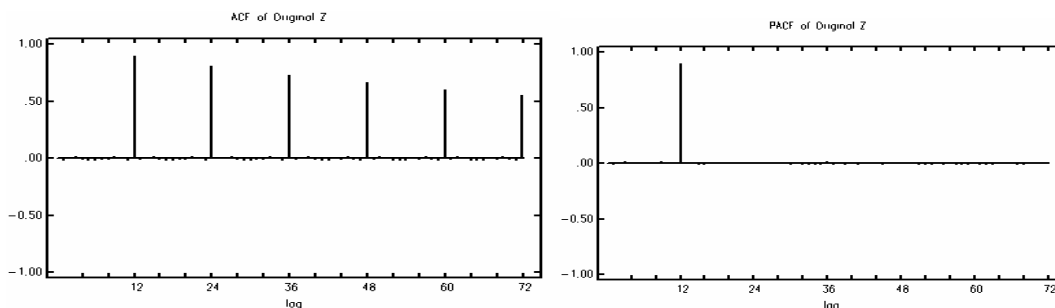
$$\rho_{j(S)} = \begin{cases} \Phi^j & \text{si } j = 1, 2, 3, \dots \\ 0 & \text{en otro caso} \end{cases}$$

y toma el valor de cero en otro caso.

Su función de autocorrelación parcial es:

$$\phi_{kS,kS} = \begin{cases} \Phi & \text{si } k = 1 \\ 0 & \text{en otro caso} \end{cases}$$

Para $\Phi = 0.9$, los correlogramas son los siguientes.



En general, para los modelos estacionales puros autorregresivos (sin componente ordinaria) de orden P , la ACF decae exponencialmente y/o en forma de ondas sinusoidales amortiguadas en los rezagos estacionales y la PACF presenta coeficientes diferentes de cero en los rezagos estacionales hasta el rezago PS . A partir del rezago $SP+1$ la PACF se anula (tiene un corte)

Ejemplo. El proceso $ARIMA(0,0,0)X(0,0,1)_S$ con $S=12$.

Corresponde a un proceso estacional puro de medias móviles de orden 1. Su especificación es,

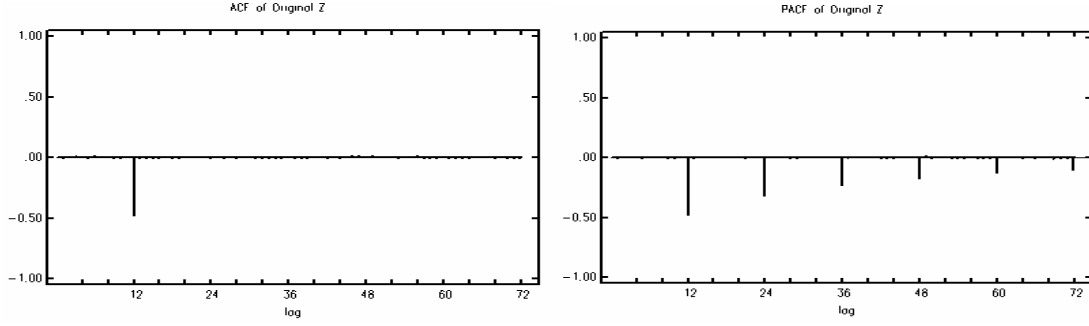
$$Z_t = (1 - \Theta B^{12})a_t$$

Su función de autocorrelación es

$$\rho_{j(S)} = \begin{cases} \frac{-\Theta}{1 + \Theta^2} & \text{si } j = 1 \\ 0 & \text{en otro caso} \end{cases}$$

Su función de autocorrelación parcial decae exponencialmente en los rezagos estacionales.

Para $\Theta = 0.8$, los correlogramas son los siguientes.



En general, para los modelos estacionales puros de medias móviles (sin componente ordinaria) de orden Q , la ACF presenta coeficientes diferentes de cero en los rezagos estacionales hasta el rezago QS . A partir del rezago $SQ+1$ la ACF se anula (tiene un corte). La PACF decae exponencialmente y/o en forma de ondas sinusoidales amortiguadas en los rezagos estacionales.

Ejemplo. El modelo $ARIMA(0,1,1)X(0,1,1)_S$, con $S=12$. Su especificación es

$$(1-B)(1-B^{12})Z_t = (1-\theta B)(1-\Theta B^{12})a_t$$

Este modelo ha sido muy útil para representar una gran variedad de series de tiempo estacionales. Este modelo fue utilizado por Box y Jenkins para modelar la serie del número pasajeros en vuelos internacionales y por esto ha sido llamado el modelo de línea aérea (airline model). Las autocovarianzas del proceso $W_t = (1-B)(1-B^{12})Z_t$ son:

$$\gamma_0 = (1+\theta^2)(1+\Theta^2)\sigma_a^2$$

$$\gamma_1 = -\theta(1+\Theta^2)\sigma_a^2$$

$$\gamma_{11} = \theta\Theta\sigma_a^2$$

$$\gamma_{12} = -\Theta(1+\theta^2)\sigma_a^2$$

$$\gamma_{13} = \theta\Theta\sigma_a^2$$

Por tanto, la ACF es

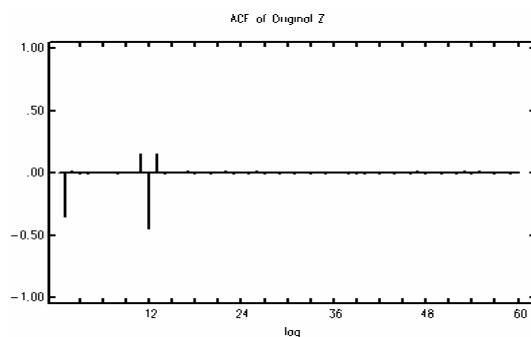
$$\rho_1 = \frac{-\theta}{(1+\theta^2)}$$

$$\rho_{11} = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)} = \rho_{13}$$

$$\rho_{12} = \frac{-\Theta}{(1+\Theta^2)}$$

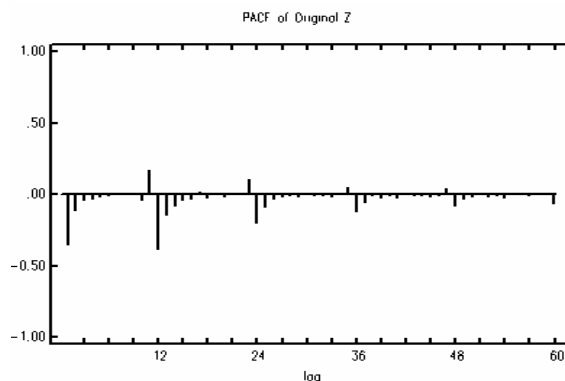
$$\rho_j = 0 \text{ en otro caso}$$

El correlograma es el siguiente



Se observa que la ACF tiene un corte después del primer rezago, indicando que el proceso ordinario es MA(1). En el rezago 12 (primer rezago estacional) aparece el patrón estacional de medias móviles de orden 1, acompañado por los coeficientes “satélites” en el rezago 11 y 13. A partir de $k=14$ la FAC se anula.

En general, la PACF de los modelos ARIMA estacionales son complicadas. Para el modelo anterior el correlograma parcial es



La PACF decae exponencialmente en los rezagos no estacionales y estacionales.

En general, los modelos estacionales puramente autorregresivos en su parte ordinaria y estacional, la PACF tiene cortes en los rezagos no estacional y estacional. Su ACF presenta decaimiento exponencial y/o ondas seno amortiguadas en los rezagos estacionales y no estacionales.

Los modelos estacionales puramente de medias móviles en su parte ordinaria y estacional, la ACF tiene cortes en los rezagos no estacional y estacional. Su PACF presenta decaimiento exponencial y/o ondas seno amortiguadas en los rezagos estacionales y no estacionales.

Para modelos estacionales mixtos autorregresivos y de medias móviles, tanto su ACF como su PACF presentan decaimientos exponenciales y/o ondas seno amortiguadas en los rezagos estacionales y no estacionales. En este caso la utilidad de la EACF es muy limitada.

Nota. Todo modelo ARIMA estacional se puede escribir como un modelo ARIMA con restricciones sobre los parámetros.

Por ejemplo, considere el modelo $ARIMA(0,0,1)X(0,0,1)_S$, con $S=12$, dado por

$$Z_t = (1 - \theta B)(1 - \Theta B^{12})a_t$$

Este modelo se puede escribir como

$$Z_t = (1 - \theta B)(1 - \Theta B^{12})a_t = (1 - \theta B - \Theta B^{12} + \theta \Theta B^{13})a_t$$

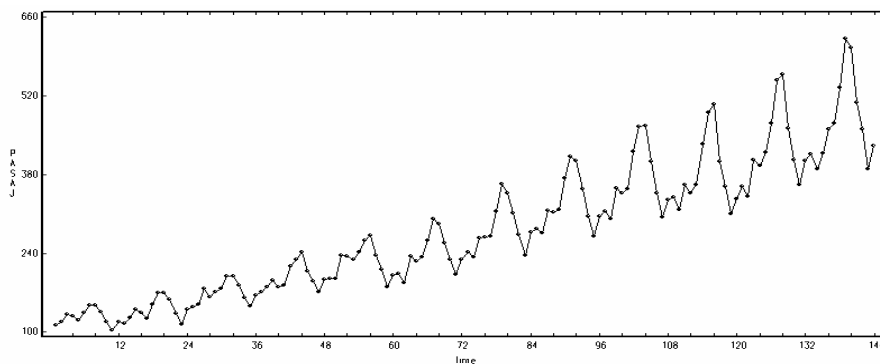
El cual corresponde a un modelo ARIMA(0,0,13), es decir a un MA(13) con restricciones en los parámetros dadas por $\theta_2 = \theta_3 = \dots = \theta_{11} = 0$ y $\theta_{13} = -\theta\Theta$. Esto implica que la estimación de los modelos ARIMA estacionales se puede realizar usando, en general, estimación no lineal bajo restricciones en los parámetros.

Para el estudio más detallado de patrones estacionales ver Guerrero (2003).

Ejemplo: Series de tiempo del número de pasajeros en líneas aéreas, Box-Jenkins (1976). Ver macro pasajer.sca

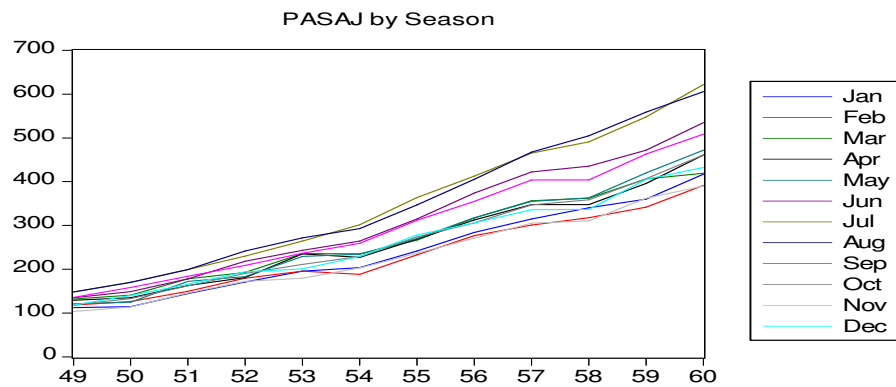
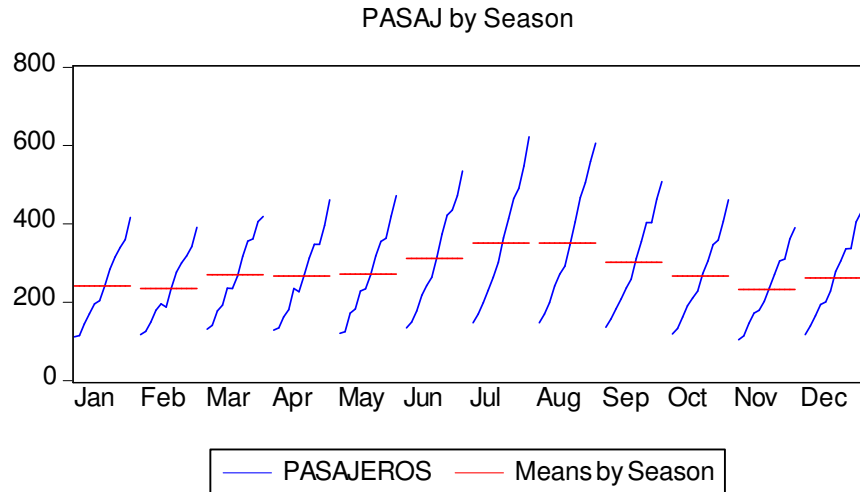
Identificación del proceso Z_t :

PASO 1: Gráfico de la serie $Z_t = \text{PASAJ}_t$



Se observa que la serie no es estacionaria en media y varianza.

Análisis Gráfico Estacional



Identificación de la transformación estabilizadora de varianza, λ .

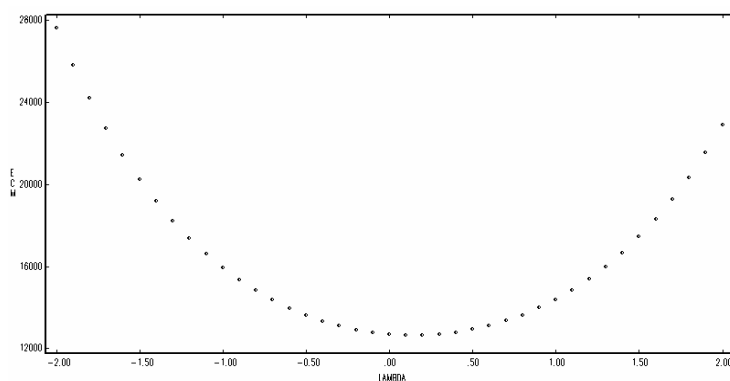
VARIABLE	LAMBDA	ECM
COLUMN-->	1	1
ROW		
1	-2.000	27636.617
2	-1.900	25828.211
3	-1.800	24202.049
4	-1.700	22739.928
5	-1.600	21425.736
	⋮	
17	-.400	13342.986
18	-.300	13108.209
19	-.200	12923.093
20	-.100	12785.930
21	.000	12695.463
22	.100	12650.876
23	.200	12651.775
24	.300	12698.190
25	.400	12790.573
26	.500	12929.808
27	.600	13117.219


```

      :
      :
38      1.700 19273.805
39      1.800 20352.150
40      1.900 21555.568
41      2.000 22897.506

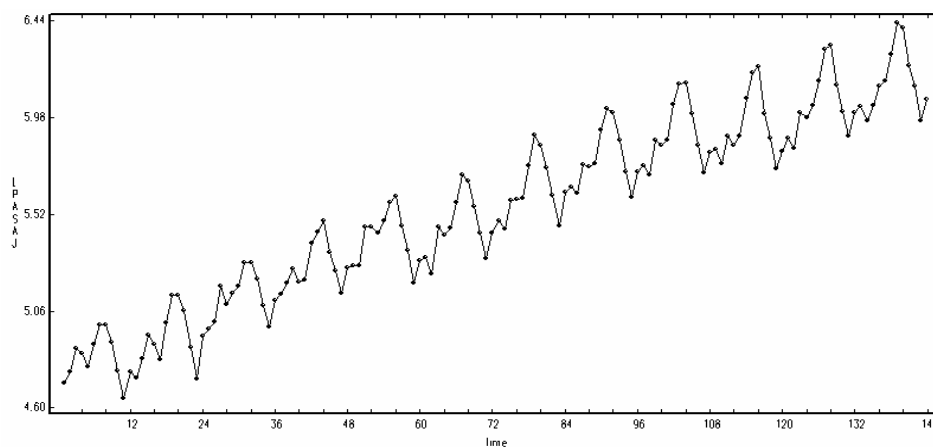
```

Gráficamente,

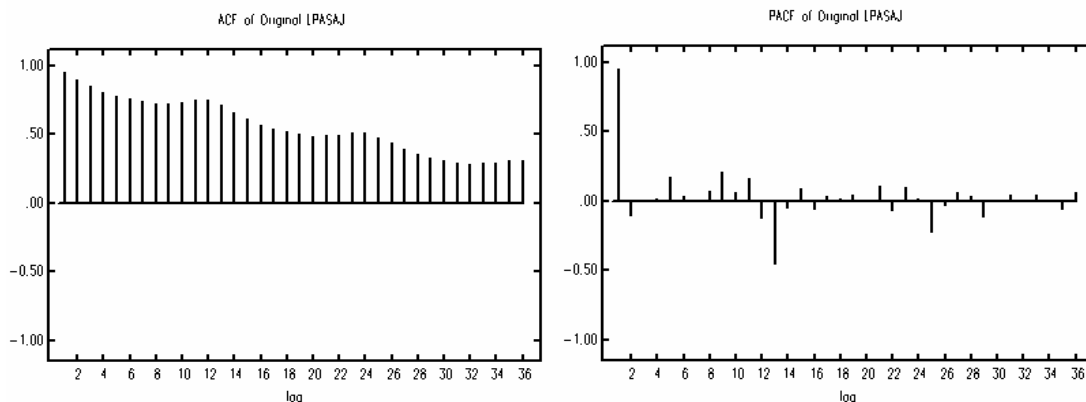


Los resultados señalan que para $\lambda=0$, es decir el logaritmo natural de la serie parece ser la transformación adecuada. El siguiente gráfico presenta la serie $Y_t = \ln(Z_t) = LPASAJ_t$ con varianza estabilizada.

Gráfico de la serie $Y_t = \ln(Z_t)$

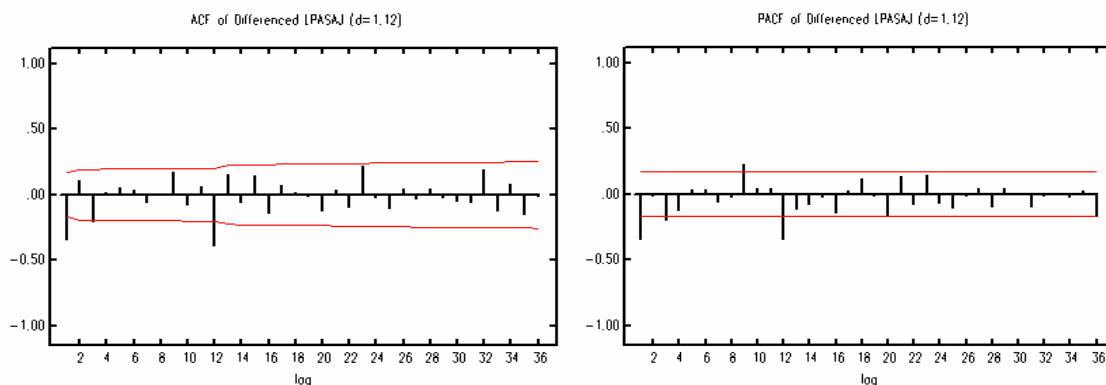


PASO 2: Examen de la ACF y PACF original de $Y_t = LPASAJ_t$



El gráfico de la serie transformada en logaritmo y los resultados de su ACF y PACF muestran que la serie no es estacionaria. Además presentan fluctuaciones en los rezagos correspondientes al período estacional.

Debemos diferenciar la parte ordinaria una vez y la parte estacional una vez. (Algunas veces basta con diferenciar la parte estacional una vez, para lograr estacionaridad en la serie). A continuación presentamos los resultados del ACF y PACF de la serie diferenciada $(1-B)(1-B^{12})Y_t$



El análisis de la PACF muestral parece indicar que hay decaimiento exponencial tanto en los rezagos no estacionales como estacionales; por su parte la ACF muestral señala que hay un corte después del primer rezago de la parte ordinaria y después del primer rezago de la parte estacional.

En conclusión, el logaritmo natural de la serie de datos de pasajeros parece ser consistente con un modelo $ARIMA(0,1,1) \times (0,1,1)_{12}$.

La estimación del modelo usando máxima verosimilitud exacta es la siguiente:

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- MOD								
VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING					
LPASAJ	RANDOM	ORIGINAL		1	12			
				(1-B)	(1-B)			
PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRAIINT	VALUE	STD ERROR	T VALUE
1	LPASAJ	MA	1	1	NONE	.4024	.0801	5.03
2	LPASAJ	MA	2	12	NONE	.5568	.0728	7.65
EFFECTIVE NUMBER OF OBSERVATIONS				131				
R-SQUARE				0.993				
RESIDUAL STANDARD ERROR				0.159459E-01				

Los diagnósticos realizados muestran que el modelo parece ser adecuado.

- **El modelo ARIMA general.** En forma general el modelo ARIMA puede ser escrito como

$$\prod_{j=1}^M \phi_j(B) \prod_{i=1}^K (1-B^{s_i})^{d_i} \dot{Z}_t = \theta_0 + \prod_{k=1}^N \theta_k(B) a_t$$

El modelo contiene K factores de diferenciación, M factores autorregresivos y N factores de medias móviles. Esta extensión es útil para representar series no estándar que, por ejemplo tengan una mezcla de componentes estacionales de diferentes períodos.

Si $K=0$, entonces $\dot{Z}_t = Z_t - \mu$ y $\dot{Z}_t = Z_t$ en otro caso.

El parámetro θ_0 representa una tendencia determinística y se considera solamente cuando $K \neq 0$.

El j-ésimo factor autorregresivo es

$$\phi_j(B) = 1 - \phi_{j1}B - \phi_{j2}B^2 - \dots - \phi_{jp_j}B^{p_j}$$

El k-ésimo factor de medias móviles es

$$\theta_j(B) = 1 - \theta_{j1}B - \theta_{j2}B^2 - \dots - \theta_{jq_j}B^{q_j}$$

En la mayoría de las aplicaciones K, M y N son menores o iguales que 2. Los parámetros autorregresivos y de medias móviles más allá del primer factor son considerados parámetros estacionales del modelo.

6. Pronósticos para Series de Tiempo Univariadas

- Uno de los objetivos más importantes del análisis de una serie de tiempo es pronosticar sus valores futuros. Los pronósticos son esenciales en la planeación y control de operación en una gran cantidad de áreas tales como manejo de producción, inventario de sistemas, control de calidad, planeamiento financiero y análisis de inversión.

Pronósticos de Error Cuadrático Medio Mínimo para Modelos ARMA

Cuando se pronostica, el objetivo es producir pronósticos que no tengan error o que el error sea tan pequeño como sea posible. Esto conduce a la obtención de pronósticos con error cuadrático medio mínimo.

Considere el modelo ARMA $\phi(B)Z_t = \theta(B)a_t$, con $\mu = 0$, estacionario e invertible.

Z_t puede ser escrito como $Z_t = \psi(B)a_t$, donde $\psi(B) = \phi^{-1}(B)\theta(B) = \sum_{j=0}^{\infty} \psi_j B^j$.

Entonces,

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \dots$$

Suponga que estamos interesados en pronosticar el valor de Z , l períodos adelante basados en la historia hasta el período $t=n$, es decir, usando $Z_n, Z_{n-1}, Z_{n-2}, \dots$. Se quiere pronosticar

$$Z_{n+l} = a_{n+l} + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \psi_3 a_{n+l-3} + \dots$$

- Sea $\hat{Z}_n(l)$ el predictor de Z_{n+l} y sea $e_n(l) = Z_{n+l} - \hat{Z}_n(l)$ el error de predicción.

Entonces el predictor $\hat{Z}_n(l)$ que minimiza $E[e_n(l)]^2$, es

$$\hat{Z}_n(l) = \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots$$

$\hat{Z}_n(l)$ es llamado el predictor l períodos adelante (o de horizonte l) de Z basado en la historia hasta el período n .

En efecto, basados en la historia $Z_n, Z_{n-1}, Z_{n-2}, \dots$, el predictor puede ser escrito como

$$\hat{Z}_n(l) = \psi_l^* a_n + \psi_{l+1}^* a_{n-1} + \psi_{l+2}^* a_{n-2} + \dots = \sum_{j=0}^{\infty} \psi_{l+j}^* a_{n-j} \quad (D)$$

donde los ψ_j^* deben ser determinados.

Ahora bien, Z_{n+l} puede ser escrito como

$$Z_{n+l} = a_{n+l} + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \dots + \psi_{l-1} a_{n+l-(l-1)} + \psi_l a_{n+l-(l)} + \psi_{l+1} a_{n+l-(l+1)} + \dots$$

$$Z_{n+l} = a_{n+l} + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \dots + \psi_{l-1} a_{n+1} + \psi_l a_n + \psi_{l+1} a_{n-1} + \dots$$

$$Z_{n+l} = \sum_{j=0}^{l-1} \psi_j a_{n+l-j} + \sum_{j=0}^{\infty} \psi_{l+j} a_{n-j}$$

Entonces

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j} + \sum_{j=0}^{\infty} \psi_{l+j} a_{n-j} - \sum_{j=0}^{\infty} \psi_{l+j}^* a_{n-j}$$

$$e_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j} + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*) a_{n-j}$$

Elevando al cuadrado y tomando valor esperado,

$$E(e_n(l)^2) = E\left(\sum_{j=0}^{l-1} \psi_j a_{n+l-j} + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*) a_{n-j}\right)^2$$

$$= E\left(\sum_{j=0}^{l-1} \psi_j^2 a_{n+l-j}^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2 a_{n-j}^2\right) + E(\text{suma de productos cruzados de los } a_t)$$

$$= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_a^2 \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2$$

De lo anterior, el pronóstico con error cuadrático medio mínimo se obtiene cuando,

$$\sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2 = 0, \text{ es decir, cuando}$$

$$\psi_{l+j}^* = \psi_{l+j}$$

Por tanto, reemplazando en (D),

$$\hat{Z}_n(l) = \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots$$

Propiedades del pronóstico

$$1) \hat{Z}_n(l) = E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots)$$

En efecto,

$$\begin{aligned}
E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) &= E(a_{n+l} + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \dots + \\
&\quad \psi_{l-1} a_{n+1} + \psi_l a_n + \psi_{l+1} a_{n-1} + \dots | Z_n, Z_{n-1}, Z_{n-2}, \dots) \\
&= E(a_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \psi_1 E(a_{n+l-1} | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \\
&\quad \psi_2 E(a_{n+l-2} | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \dots + \psi_{l-1} E(a_{n+1} | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \\
&\quad \psi_l E(a_n | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \psi_{l+1} E(a_{n-1} | Z_n, Z_{n-1}, Z_{n-2}, \dots) + \dots
\end{aligned}$$

donde

$$E(a_{n+j} | Z_n, Z_{n-1}, Z_{n-2}, \dots) = \begin{cases} 0 & \text{si } j > 0 \\ a_{n+j} & \text{si } j \leq 0 \end{cases}$$

Por tanto,

$$E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) = \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots = \hat{Z}_n(l)$$

2) El error de pronóstico es $e_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j}$

$$3) E(e_n(l)) = E\left(\sum_{j=0}^{l-1} \psi_j a_{n+l-j}\right) = 0$$

$$4) \text{var}(e_n(l)) = \text{var}\left(\sum_{j=0}^{l-1} \psi_j a_{n+l-j}\right) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$$

5) Para un proceso Gaussiano (normal), el intervalo de predicción con $(1-\alpha)\%$ de confianza está dado por:

$$\hat{Z}_n(l) \pm z_{\alpha/2} \sigma_a \left[\sum_{j=0}^{l-1} \psi_j^2 \right]^{1/2}$$

- 6) Si $l=1$, $e_n(1) = Z_{n+1} - \hat{Z}_n(1) = a_{n+1}$, es decir, los errores de pronóstico 1 período adelante no son correlacionados, y bajo normalidad son independientes.
- 7) Para $l>1$, los errores de los pronósticos están correlacionados, cuando se calculan para diferentes orígenes. Es decir, $\text{cov}(e_n(l), e_{n-j}(l)) \neq 0$, para $j < l$
- 8) Para diferentes horizontes, los errores de pronóstico basados en el mismo origen están correlacionados. Por ejemplo,

$$\text{cov}(e_n(2), e_n(1)) = E((a_{n+2} + \psi_1 a_{n+1})a_{n+1}) = E(a_{n+2}a_{n+1}) + \psi_1 E(a_{n+1}^2) = \psi_1 \sigma_a^2$$

- **Pronósticos de Error Cuadrático Medio Mínimo para modelos ARIMA**

Considere el modelo no estacionario ARIMA, $\phi(B)(1-B)^d Z_t = \theta(B)a_t$, donde $\phi(B)$ estacionario y $\theta(B)$ es invertible. El predictor óptimo $\hat{Z}_n(l)$ de Z_{n+l} es también de la forma,

$$\hat{Z}_n(l) = E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots)$$

y su varianza es $\text{var}(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$

Para obtener la varianza del error del pronóstico para el modelo ARIMA general, se debe escribir el modelo en el período $t+l$ en su representación AR, puesto que el modelo es invertible,

$$\pi(B)Z_t = a_t, \text{ donde } \pi(B) = \frac{\phi(B)(1-B)^d}{\theta(B)} = 1 - \sum_{j=1}^{\infty} \pi_j B^j$$

y las ψ_j se calculan recursivamente usando

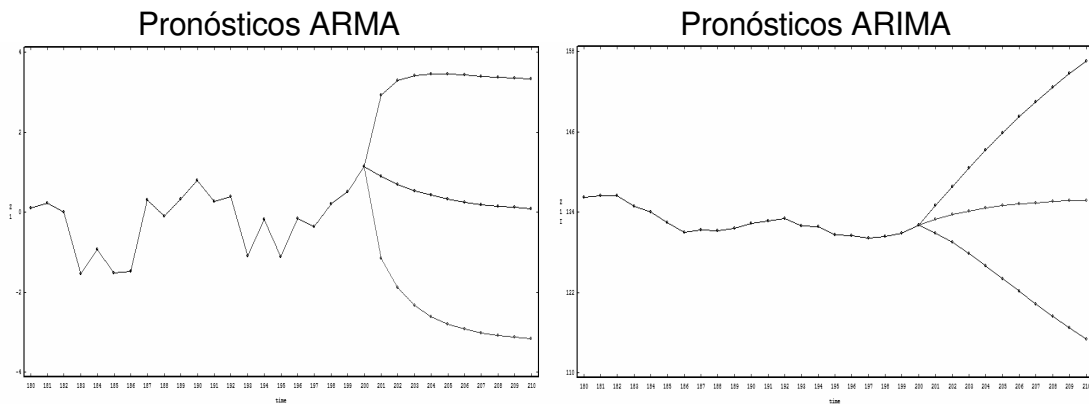
$$\psi_j = \sum_{i=0}^{j-1} \pi_{j-i} \psi_i$$

donde π_i se obtienen al expresar el modelo en la forma autorregresiva anterior.

El intervalo de predicción tiene la misma forma que en el caso del modelo ARMA.

Observación

Aunque los intervalos de predicción se expresan de la misma forma, mientras que en un proceso ARMA $\lim_{l \rightarrow \infty} \sum_{j=0}^{l-1} \psi_j^2$ converge y por tanto los límites del intervalo de pronóstico se van aproximando a dos líneas paralelas, en el caso de un modelo ARIMA, dicho límite no existe y la $\sum_{j=0}^{l-1} \psi_j^2$ se incrementa a medida que $l \rightarrow \infty$. Los límites llegan a ser cada vez más amplios a medida que l crece. Esto significa que a medida que el horizonte de pronóstico crece, también crece indefinidamente la incertidumbre en el pronóstico.



- **Cálculo de los pronósticos:**

Del resultado $\hat{Z}_n(l) = E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots)$, los pronósticos pueden ser obtenidos fácilmente por medio del empleo directo de la forma de ecuación en diferencia del modelo.

El modelo general ARIMA(p,d,q) puede ser escrito como $\Psi(B)Z_t = \theta(B)a_t$

$$(1 - \Psi_1 B - \Psi_2 B^2 - \dots - \Psi_{p+d} B^{p+d}) Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

o, para $t = n + l$,

$$Z_{n+l} = \Psi_1 Z_{n+l-1} + \Psi_2 Z_{n+l-2} + \dots + \Psi_{p+d} Z_{n+l-(p+d)} + a_{n+l} - \theta_1 a_{n+l-1} - \theta_2 a_{n+l-2} - \dots - \theta_q a_{n+l-q}$$

Tomando esperanza condicional, el pronóstico se puede calcular como:

$$\begin{aligned} E(Z_{n+l} | Z_n, Z_{n-1}, \dots) &= \Psi_1 E(Z_{n+l-1} | Z_n, Z_{n-1}, \dots) + \Psi_2 E(Z_{n+l-2} | Z_n, Z_{n-1}, \dots) + \dots + \\ &\quad \Psi_{p+d} E(Z_{n+l-(p+d)} | Z_n, Z_{n-1}, \dots) + E(a_{n+l} | Z_n, Z_{n-1}, \dots) - \theta_1 E(a_{n+l-1} | Z_n, Z_{n-1}, \dots) - \\ &\quad \theta_2 E(a_{n+l-2} | Z_n, Z_{n-1}, \dots) - \dots - \theta_q E(a_{n+l-q} | Z_n, Z_{n-1}, \dots) \end{aligned}$$

o,

$$\begin{aligned} \hat{Z}_n(l) &= \Psi_1 \hat{Z}_n(l-1) + \Psi_2 \hat{Z}_n(l-2) + \dots + \Psi_{p+d} \hat{Z}_n(l-p-d) + \hat{a}_n(l) - \theta_1 \hat{a}_n(l-1) - \theta_2 \hat{a}_n(l-2) - \\ &\quad \dots - \theta_q \hat{a}_n(l-q) \end{aligned}$$

donde

$$\begin{aligned}
\hat{Z}_n(j) &= E(Z_{n+j} | Z_n, Z_{n-1}, \dots) & \text{si } j \geq 1 \\
\hat{Z}_n(j) &= Z_{n+j} & \text{si } j \leq 0 \\
\hat{a}_n(j) &= 0 & \text{si } j \geq 1 \\
\hat{a}_n(j) &= Z_{n+j} - \hat{Z}_{n+j-1}(1) = a_{n+j} & \text{si } j \leq 0
\end{aligned}$$

Ejemplo. Pronósticos de un ARMA(1,1)

$$(1 - \phi_1 B)(Z_t - \mu) = (1 - \theta_1 B)a_t$$

1) Cálculo de $\hat{Z}_n(l)$, como la media condicional de la ecuación en diferencia.

Para $t = n + l$, se puede escribir el modelo como

$$Z_{n+l} = \mu + \phi_1(Z_{n+l-1} - \mu) + a_{n+l} - \theta_1 a_{n+l-1}$$

Tomando esperanza condicional sobre la historia,

$$\begin{aligned}
E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) &= \mu + \phi_1(E(Z_{n+l-1} | Z_n, Z_{n-1}, Z_{n-2}, \dots) - \mu) + \\
&\quad E(a_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) - \theta_1 E(a_{n+l-1} | Z_n, Z_{n-1}, Z_{n-2}, \dots)
\end{aligned}$$

Por tanto,

$$\hat{Z}_n(l) = \mu + \phi_1(\hat{Z}_n(l-1) - \mu) + \hat{a}_n(l) - \theta_1 \hat{a}_n(l-1)$$

Si $l = 1$,

$$\hat{Z}_n(1) = \mu + \phi_1(\hat{Z}_n(0) - \mu) + \hat{a}_n(1) - \theta_1 \hat{a}_n(0)$$

$$\hat{Z}_n(1) = \mu + \phi_1(Z_n - \mu) - \theta_1 a_n$$

Si $l = 2$,

$$\begin{aligned}
\hat{Z}_n(2) &= \mu + \phi_1(\hat{Z}_n(1) - \mu) + \hat{a}_n(2) - \theta_1 \hat{a}_n(1) \\
&= \mu + \phi_1([\mu + \phi_1(Z_n - \mu) - \theta_1 a_n] - \mu) \\
&= \mu + \phi_1^2(Z_n - \mu) - \phi_1 \theta_1 a_n
\end{aligned}$$

Si $l = 3$,

$$\begin{aligned}
\hat{Z}_n(3) &= \mu + \phi_1(\hat{Z}_n(2) - \mu) + \hat{a}_n(3) - \theta_1 \hat{a}_n(2) \\
&= \mu + \phi_1^3(Z_n - \mu) - \phi_1^2 \theta_1 a_n
\end{aligned}$$

En general, para $l > 1$,

$$\hat{Z}_n(l) = \mu + \phi_1^l(Z_n - \mu) - \phi_1^{l-1} \theta_1 a_n$$

2) Cálculo de la varianza del error de pronóstico, $var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$, donde

ψ_j se obtienen de la representación de medias móviles del modelo ARMA(1,1), es decir, se obtienen de

$$(1 - \phi_1 B)(1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots) = (1 - \theta_1 B)$$

La solución es

$$\psi_j = \phi_1^{j-1}(\phi_1 - \theta_1), \text{ si } j \geq 1$$

Reemplazando,

$$var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 = \sigma_a^2 \left(1 + \sum_{j=1}^{l-1} [\phi_1^{j-1}(\phi_1 - \theta_1)]^2 \right)$$

Si $l \rightarrow \infty$,

$$var(e_n(l)) \rightarrow \sigma_a^2 [1 + (\phi_1 - \theta_1)^2 / (1 - \phi_1^2)]$$

Ejemplo. La varianza del error de pronóstico de un IMA(1,1)

Si en el modelo anterior $\phi_1=1$ y $\mu=0$, se obtiene el modelo no estacionario IMA(1,1)

$$(1-B)Z_t = (1-\theta_1 B)a_t$$

Para calcular las ponderaciones ψ_j necesarias para obtener la varianza del error de

pronóstico $var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$, debido a que no existe la representación MA para

Z_t , se debe obtener el modelo escrito en forma AR (la cual existe si $|\theta_1| < 1$),

Para el período $t+l$,

$$\pi(B)Z_{t+l} = a_{t+l}$$

donde

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = \frac{1-B}{1-\theta_1 B}$$

o,

$$1-B = 1 - (\pi_1 + \theta_1)B - (\pi_2 - \pi_1 \theta_1)B^2 - (\pi_3 - \pi_2 \theta_1)B^3 - \dots$$

Igualando los coeficientes de las potencias iguales a ambos lados de la ecuación, se obtiene,

$$\pi_j = (1-\theta_1)\theta_1^{j-1}, \text{ si } j \geq 1$$

Dados los π_j , los ψ_j requeridos para calcular la varianza del error de pronóstico,

se obtienen recursivamente de $\psi_j = \sum_{i=0}^{j-1} \pi_{j-i} \psi_i$. Por tanto,

$$\psi_1 = \pi_1 = (1-\theta_1)$$

$$\psi_2 = \pi_2 - \pi_1 \psi_1 = (1-\theta_1)\theta_1 + (1-\theta_1)^2 = (1-\theta_1)$$

En general, $\psi_j = (1-\theta_1)$, para $1 \leq j \leq l-1$

De esta forma, la varianza del error de pronóstico es

$$\text{var}(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 = \sigma_a^2 [1 + (l-1)(1-\theta_1)^2]$$

Para este caso, si $l \rightarrow \infty$,

$$\text{var}(e_n(l)) \rightarrow \infty$$

Es decir, no converge.

Ejemplo numérico. Pronósticos de un AR (1)

Considere el modelo AR(1)

$$(1 - \phi_1 B)(Z_t - \mu) = a_t$$

Con $\phi_1 = 0.6$, $\mu = 9$ y $\sigma_a^2 = 0.1$. Suponga que se tienen las observaciones $Z_{97} = 9.6$, $Z_{98} = 9.0$, $Z_{99} = 9.0$, $Z_{100} = 8.9$. Se quieren pronosticar $Z_{101}, Z_{102}, Z_{103}$ y Z_{104} , junto sus límites de predicción del 95%.

1) Pronósticos. El modelo AR(1) puede ser escrito como

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + a_t$$

$$Z_t = \mu + \phi_1(Z_{t-1} - \mu) + a_t$$

Entonces, la forma general de los pronósticos es, para $l \geq 1$

$$\hat{Z}_n(l) = \mu + \phi_1(\hat{Z}_n(l-1) - \mu)$$

$$= \mu + \phi_1^l(Z_n - \mu)$$

Por tanto,

$$\hat{Z}_{100}(1) = 9 + 0.6(8.9 - 9) = 8.94$$

$$\hat{Z}_{100}(2) = 9 + (0.6)^2(8.9 - 9) = 8.964$$

$$\hat{Z}_{100}(3) = 9 + (0.6)^3(8.9 - 9) = 8.9784$$

$$\hat{Z}_{100}(4) = 9 + (0.6)^4(8.9 - 9) = 8.98704$$

2) Para obtener los límites de predicción, se deben calcular las ψ_j de la relación

$$(1 - \phi_1 B)(1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots) = 1$$

La solución es

$$\psi_j = \phi_1^j, \text{ si } j \geq 0$$

Por tanto los límites de predicción del 95% son:

$$\text{Para } Z_{101} \text{ son } 8.94 \pm 1.96\sqrt{0.1} \quad \text{o} \quad 8.320 < Z_{101} < 9.560$$

$$\text{Para } Z_{102} \text{ son } 8.964 \pm 1.96\sqrt{0.1(1 + \psi_1^2)}$$

$$\text{Para } Z_{103} \text{ son } 8.9784 \pm 1.96\sqrt{0.1(1 + \psi_1^2 + \psi_2^2)}$$

$$\text{Para } Z_{104} \text{ son } 8.98704 \pm 1.96\sqrt{0.1(1 + \psi_1^2 + \psi_2^2 + \psi_3^2)}$$

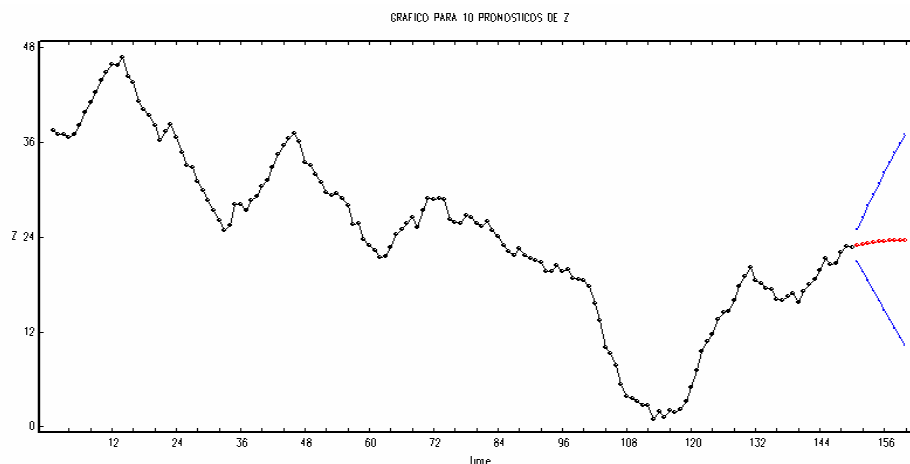
Ejemplo: Pronósticos para el modelo simulado.

Usando 150 datos simulados para una serie generada por un modelo ARIMA(1,1,1), se pronosticaron 10 observaciones a partir del dato 150. Los resultados fueron los siguientes.

10 FORECASTS, BEGINNING AT 150

TIME	FORECAST	STD. ERROR	ACTUAL IF KNOWN
151	22.9835	1.0180	
152	23.1402	1.7331	
153	23.2660	2.4354	
154	23.3669	3.1252	
155	23.4478	3.7973	
156	23.5128	4.4480	
157	23.5649	5.0753	
158	23.6066	5.6783	
159	23.6402	6.2573	
160	23.6671	6.8130	

VARIABLE	PRON	LIMINF	LIMSUP
COLUMN-->	1	1	1
ROW			
1	22.983	20.948	25.020
2	23.140	19.674	26.606
3	23.266	18.395	28.137
4	23.367	17.117	29.617
5	23.448	15.853	31.042
6	23.513	14.617	32.409
7	23.565	13.414	33.716
8	23.607	12.250	34.963
9	23.640	11.126	36.155
10	23.667	10.041	37.293



Pronósticos para la Serie Original cuando se usan Transformaciones (Ver Guerrero, 2003)

- Cuando la serie es transformada usando una transformación potencial, se debe tener presente que los pronósticos de la serie original son **las retransformaciones de los pronósticos de la serie transformada corregidos por un factor de disminución de sesgos**. Guerrero (2003), en la página 253, presenta el procedimiento que debería ser seguido.

En general, Guerrero (1993) mostró que bajo la transformación de Box y Cox $T(Z_t)$ para la serie Z_t , el pronóstico aproximadamente insesgado para Z_{t+h} , basado en la información hasta el período t es de la forma

$$\hat{Z}_t(h) = T^{-1}[\hat{T}(Z_t)(h)] \cdot \hat{c}_{t,\lambda}(h)$$

donde $\hat{c}_{t,\lambda}(h)$ es un factor de corrección de sesgos dado por,

$$\hat{c}_{t,\lambda}(h) = \begin{cases} \left\{ 0.5 + 0.5 \left[1 + 2(\lambda^{-1} - 1) \frac{V\hat{r}_t[e_t(h)]}{[\lambda^{-1} + \hat{T}(Z_t)(h)]^2} \right]^{1/2} \right\}^{1/\lambda} & \text{si } \lambda \neq 0 \\ \exp(V\hat{r}_t[e_t(h)]/2) & \text{si } \lambda = 0 \end{cases}$$

y donde $\hat{T}(Z_t)(h)$ es el pronóstico de la serie transformada h períodos adelante

y $V\hat{r}_t[e_t(h)]$ es la varianza estimada del error del pronóstico h períodos

adelante de la serie transformada, es decir $V\hat{r}_t[e_t(h)] = \hat{\sigma}_a^2 \sum_{j=0}^{h-1} \hat{\psi}_j^2$.

Por ejemplo, si la transformación empleada fue $\lambda=0$, es decir, bajo la transformación de Box y Cox se aplicó la transformación logaritmo natural a la serie Z_t , y se obtuvo $Y_t = \ln(Z_t)$, el pronóstico de Z_t h períodos adelante es

$$\hat{Z}_t(h) = \exp(\hat{Y}_n(h)) \cdot \exp(V\hat{a}_t[e_n(h)]/2)$$

$$\hat{Z}_t(h) = \exp(\hat{Y}_n(h) + V\hat{a}_t[e_n(h)]/2)$$

donde $\hat{Y}_n(h)$ es el pronóstico h pasos delante de la serie transformada.

Los intervalos sesgados para los pronósticos de la serie original son de la forma

$$T^{-1} \left\{ \hat{Z}_t(h) \pm z_{\alpha/2} \sqrt{V\hat{a}_t[e_t(h)]} \right\}$$

El sesgo se puede reducir multiplicando los límites por el factor de corrección $\hat{c}_{t,\lambda}(h)$. Para el caso de la transformación logarítmica, el intervalo sería,

$$\exp \left\{ \hat{Z}_t(h) \pm z_{\alpha/2} \sqrt{V\hat{a}_t[e_t(h)]} \right\} \cdot \exp(V\hat{a}_t[e_n(h)]/2)$$

Ejemplo. Serie mensual de número de pasajeros en vuelos internacionales.

Para la serie mensual del número de pasajeros se obtuvo $\lambda=0$, es decir, para estabilizar la varianza de la serie se usó el logaritmo natural, y se obtuvo $Y_t = \ln(Z_t)$. Usando los resultados anteriores, el pronóstico de Z_t h períodos adelante sesgado es

$$\hat{Z}_t(h) = \exp(\hat{Y}_n(h))$$

y el pronóstico con sesgo reducido es

$$\hat{Z}_t(h) = \exp(\hat{Y}_n(h) + V\hat{a}_t[e_n(h)]/2)$$

donde $\hat{Y}_n(h)$ es el pronóstico h períodos adelante, de la serie transformada. Los intervalos sesgados para los pronósticos de la serie original son de la forma

$$T^{-1} \left\{ \hat{Z}_t(h) \pm z_{\alpha/2} \sqrt{\hat{V}ar_t[e_t(h)]} \right\}$$

Los intervalos con sesgos reducidos son

$$\exp \left\{ \hat{Z}_t(h) \pm z_{\alpha/2} \sqrt{\hat{V}ar_t[e_t(h)]} \right\} \cdot \exp(\hat{V}ar_t[e_n(h)]/2)$$

Usando estos resultados se calcularon 24 pronósticos y los intervalos de predicción del 95% para la serie original a partir de enero de 1961. Las columnas indican:

Sesg: pronósticos sesgados

SesgR: pronósticos con sesgo reducido

LinfS: límite inferior del intervalo sesgado

LsupS: límite superior del intervalo sesgado

LinfR: límite inferior del intervalo con sesgo reducido

LsupR: límite superior del intervalo con sesgo reducido

AmpS: amplitud del intervalo sesgado

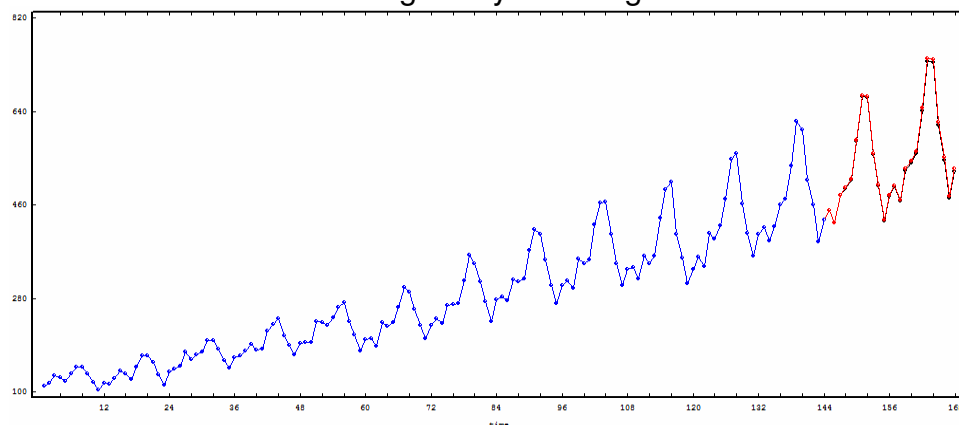
AmpR: amplitud del intervalo con sesgo reducido

Pron	Sesg	SesgR	LinfS	LsupS	LinfR	LsupR	AmpS	AmpR
1	450	451	418	485	419	485	66	66
2	426	426	391	464	391	464	73	73
3	479	480	435	528	435	528	93	93
4	492	493	443	547	443	548	105	105
5	509	510	454	571	454	572	117	118
6	583	584	516	660	517	661	144	144
7	670	671	588	764	589	765	176	176
8	667	669	581	766	582	768	185	185
9	558	560	483	645	484	647	163	163
10	497	499	427	579	428	580	151	152
11	430	431	367	503	368	505	136	137
12	477	479	405	562	406	564	157	158
13	496	498	414	594	415	597	181	181

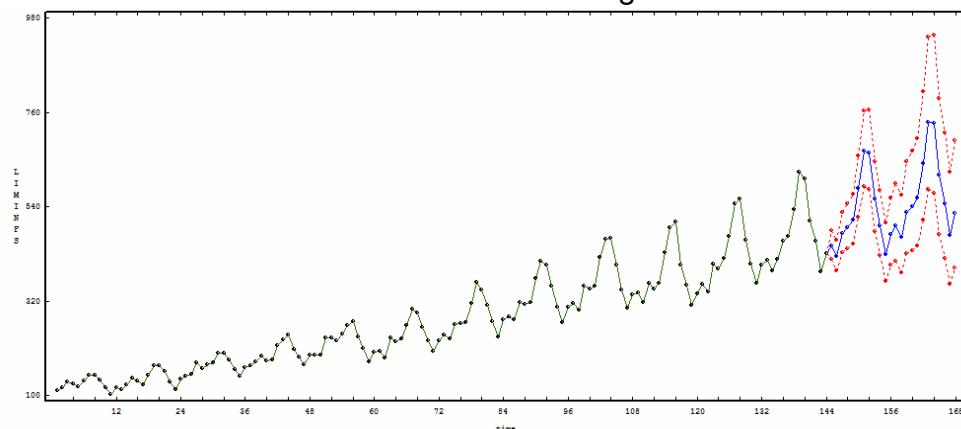
14	469	471	387	568	389	571	181	182
15	527	530	431	646	433	649	215	216
16	542	545	438	670	441	674	232	233
17	560	564	449	699	452	704	250	252
18	642	647	510	809	513	814	299	301
19	738	743	581	937	585	944	356	359
20	734	740	573	941	578	948	367	370
21	615	620	476	794	480	800	318	320
22	547	552	421	712	424	719	292	294
23	473	478	361	621	364	626	260	262
24	525	531	398	694	402	701	296	299

El efecto de la corrección se observa tanto en el pronóstico, como en la amplitud de los intervalos. A continuación se presentan los gráficos con los pronósticos y los intervalos de predicción.

Pronósticos sesgados y con sesgos reducidos



Pronósticos e intervalos con sesgos reducidos



Evaluación de los pronósticos

Es importante evaluar la capacidad predictiva del modelo. En el trabajo aplicado con series de tiempo es común evaluar la capacidad predictiva de un modelo comparando los valores observados con los pronosticados. Para desarrollar estas ideas es usual ajustar el modelo con las “T” primeras observaciones y analizar luego cómo el modelo predice las “h” observaciones siguientes, donde $n=T+h$, siendo n el tamaño de la muestra disponible. Algunas de las medidas más empleadas son:

- La raíz del error cuadrático medio:

$$RMSE = \sqrt{\frac{\sum_{t=T+1}^n (y_t - \hat{y}_t)^2}{h}}$$

- El promedio de los valores absolutos de los errores de pronósticos

$$MAE = \frac{\sum_{t=T+1}^n |y_t - \hat{y}_t|}{h}$$

- La raíz del error cuadrático medio porcentual:

$$RMSPE = \sqrt{\frac{\sum_{t=T+1}^n y_t \left(\frac{y_t - \hat{y}_t}{y_t} \right)^2}{h}}$$

- El promedio del error porcentual absoluto

$$MAPE = \frac{\sum_{t=T+1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|}{h}$$

Descomposición del error cuadrático medio

Se puede probar que el error cuadrático medio se puede descomponer de la siguiente forma:

$$\frac{\sum_{t=T+1}^n (\hat{y}_t - y_t)^2}{h} = (\bar{\hat{y}} - \bar{y})^2 + (s_{\hat{y}} - s_y)^2 + 2(1-r) s_{\hat{y}} s_y$$

donde r es el coeficiente de correlación entre y_t y \hat{y}_t ; \bar{y} , $\bar{\hat{y}}$, s_y , $s_{\hat{y}}$ son las medias y desviaciones estándar de y_t y \hat{y}_t . Con base en esta descomposición se definen:

- La proporción del sesgo:

$$PS = \frac{(\bar{\hat{y}} - \bar{y})^2}{\frac{\sum_{t=T+1}^n (\hat{y}_t - y_t)^2}{h}}$$

- La proporción de la varianza

$$PV = \frac{(s_{\hat{y}} - s_y)^2}{\frac{\sum_{t=T+1}^n (\hat{y}_t - y_t)^2}{h}}$$

- La proporción de la covarianza

$$PC = \frac{2(1-r) s_{\hat{y}} s_y}{\frac{\sum_{t=T+1}^n (\hat{y}_t - y_t)^2}{h}}$$

Observe que $PS + PV + PC = 1$.

PS es una medida de la desviación del promedio del pronóstico con relación al promedio de la serie observada (medida del error sistemático). Valores grandes de PS indican que un sesgo sistemático está presente y por lo tanto el modelo se debería revisar. La proporción de la varianza nos informa la capacidad que tiene el modelo de replicar la variabilidad de la serie. PC mide el error no sistemático. La situación ideal sería $PS = PV = 0$ y $PC = 1$.

Ejemplo: Ver script de R.

Raíces unitarias estacionales

- Como vimos el modelo para la serie de datos de pasajeros es

$$(1-B)(1-B^{12})\ln(Z_t) = (1-\theta_1 B)(1-\Theta_{12} B^{12})a_t$$

Es decir, la serie del logaritmo tiene una raíz unitaria y una raíz unitaria estacional. La decisión fue tomada en forma subjetiva con base en la ACF muestral. Esta decisión puede ser derivada en forma objetiva por medio de pruebas para la existencia de raíces estacionales.

- Suponga que las observaciones trimestrales de una serie son generadas por

$$Z_t = Z_{t-4} + a_t$$

Para este modelo la primera diferencia estacional es estacionaria, puesto que

$$(1-B^4)Z_t = Z_t - Z_{t-4} = a_t$$

- Existen diferentes alternativas para tratar la estacionalidad en una serie no estacionaria.
- ✓ Cuando el patrón estacional es puramente determinístico. Para este caso se definen variables dummy D_{it} que toman el valor de 1 si la observación Z_t pertenece al i -ésimo trimestre y toman el valor de cero en otro caso. Para probar **la existencia de una raíz unitaria no estacional** en esta clase de series, según sea el proceso generador de datos use las siguientes ecuaciones

$$\Delta Z_t = \beta_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \gamma Z_{t-1} + \sum_{i=2}^P \beta_i \Delta Z_{t-i+1} + \varepsilon_t$$

(con intercepto)

$$\Delta Z_t = \beta_0 + \beta_1 t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \gamma Z_{t-1} + \sum_{i=2}^P \beta_i \Delta Z_{t-i+1} + \varepsilon_t$$

(con intercepto y tendencia)

La hipótesis nula de una raíz unitaria, es decir $\gamma = 0$, puede ser probada usando la prueba de Dickey-Fuller con el estadístico τ_μ para el primer caso, o con el estadístico τ_τ , en el segundo caso.

- ✓ El patrón estacional es aleatorio. En este caso se sospecha de la existencia de una raíz unitaria **estacional**. Para explicar la metodología, observe que el polinomio $(1 - \gamma B^4)$ puede ser factorizado como

$$(1 - \gamma B^4) = (1 - \gamma^{1/4} B)(1 + \gamma^{1/4} B)(1 - i\gamma^{1/4} B)(1 + i\gamma^{1/4} B)$$

Si Z_t tiene una raíz unitaria estacional, $\gamma=1$. La ecuación anterior es un poco restrictiva puesto que solamente permite una raíz unitaria en la frecuencia anual.

Hylleberg, Engle, Granger y Yoo (1990), desarrollaron una técnica que permite probar la existencia de raíces unitarias en varias frecuencias: se puede verificar la existencia de una raíz unitaria (se le llama una raíz en la frecuencia cero), una raíz unitaria en la frecuencia semianual, o una raíz unitaria estacional.

Para entender la prueba denominada HEGY, suponga que Z_t es generada por

$$A(B)Z_t = a_t$$

donde $A(B)$ es un polinomio de cuarto orden tal que

$$A(B) = (1 - c_1 B)(1 + c_2 B)(1 - c_3 iB)(1 + c_4 iB)$$

Si $c_1 = c_2 = c_3 = c_4 = 1$, entonces $A(B) = 1 - B^4$, lo que implica que $\gamma = 1$ y en ese caso existe una raíz unitaria estacional.

Los otros casos posibles son:

i) Si $c_1 = 1$, una solución homogénea a la ecuación anterior es $Z_t = Z_{t-1}$

Este es el caso de una raíz unitaria no estacional, en el cual la diferenciación adecuada es $(1-B)Z_t$.

ii) Si $c_2 = 1$, una solución homogénea a la ecuación anterior es $Z_t + Z_{t-1} = 0$.

En este caso, la serie tiende a replicarse a sí misma en intervalos de 6 meses, y por tanto hay una raíz unitaria semianual. Por ejemplo, si $Z_t = 1$ entonces $Z_{t+1} = -1$, $Z_{t+2} = 1$, $Z_{t+3} = -1$, $Z_{t+4} = 1$, y así sucesivamente.

iii) Si $c_3 = 1$ o $c_4 = 1$, la serie tiene un ciclo anual. Por ejemplo, Si $c_3 = 1$, una solución homogénea a la ecuación anterior es $Z_t = iZ_{t-1}$. Por tanto si $Z_t = 1$ entonces $Z_{t+1} = i$, $Z_{t+2} = i^2 = -1$, $Z_{t+3} = -i$, $Z_{t+4} = -i^2 = 1$, de manera que la serie se replica a sí misma cada 4 períodos. En esta caso el grado apropiado de diferenciación es $(1+B^2)Z_t$.

Con lo anterior, para realizar el test se siguen los siguientes pasos:

Paso 1. para datos trimestrales obtenga las siguientes variables:

$$y_{1,t-1} = (1 + B + B^2 + B^3)Z_{t-1} = Z_{t-1} + Z_{t-2} + Z_{t-3} + Z_{t-4}$$

$$y_{2,t-1} = (1 - B + B^2 - B^3)Z_{t-1} = Z_{t-1} - Z_{t-2} + Z_{t-3} - Z_{t-4}$$

$$y_{3,t-1} = (1 - B^2)Z_{t-1} = Z_{t-1} - Z_{t-3}$$

Paso 2. Estime la ecuación

$$(1 - B^4)Z_t = \gamma_1 y_{1,t-1} + \gamma_2 y_{2,t-1} + \gamma_5 y_{3,t-1} + \gamma_6 y_{3,t-2} + \varepsilon_t$$

Esta ecuación puede ser modificada introduciendo un intercepto, dummies estacionales determinísticas, una tendencia lineal determinística y como en la prueba aumentada de Dickey-Fuller, introducir valores rezagados $(1-B^4)Z_{t-i}$. Haga los diagnósticos apropiados para chequear el modelo y asegúrese que los residuales son ruido blanco.

Paso 3.

Obtenga el estadístico t para probar $H_0: \gamma_1 = 0$ contra $H_1: \gamma_1 < 0$. Los valores críticos de la prueba son reportados por Hylleberg et al (1990). Si no rechaza H_0 , concluya que $c_1 = 1$, y por tanto que hay una raíz unitaria no estacional.

✓ Obtenga el estadístico t para probar $H_0: \gamma_2 = 0$ contra $H_1: \gamma_2 < 0$. Si no rechaza H_0 , concluya que $c_2 = 1$, y por tanto que hay una raíz unitaria con frecuencia semianual.

✓ Obtenga el estadístico F para probar $H_0: \gamma_5 = \gamma_6 = 0$. Los valores críticos de la prueba son reportados por Hylleberg et al (1990). Si el valor calculado el estadístico es menor que el valor crítico, no rechace H_0 , y concluya que hay una raíz unitaria estacional.

- Resultados de estudios Monte Carlo (Hylleberg, 1995) indican que el tamaño y la potencia de la prueba HEGY son razonables cuando el proceso generador de datos es AR. Sin embargo, la prueba se comporta pobremente en presencia de errores MA.

Para datos mensuales, Franses (1990) extiende los resultados de la prueba HEGY. Ver También en Maddala y Kim (2000), cap. 12, sec. 12.3, o Guerrero(2003), cap.5, sec. 5.4.2

Lecturas:

Wei (2006) "Time Series análisis". Capítulo 8, 9.

Guerrero (2003) "Análisis Estadístico de Series de tiempo Económicas". Capítulo 5.

Maddala y Kim (2000) "Unit Roots, Cointegration and Structural Changes".

Pascual, Romo y Ruiz (2004) "Bootstrap prediction intervals for power-transformed time series", International Journal of Forecasting, Vol 21, No 2, April-June 2005, pp 219-235.

Liu (2005) "Time Series Analysis and Forecasting", cap. 8.