Universidad
Rey Juan Carlos

10. SERIES TEMPORALES

Estadística Superior

Dr. Francisco Rabadán Pérez
Dpto. Economía Aplicada I

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

1. Introducción.

- **Serie temporal, cronológica, histórica o de tiempo:** observaciones cuantitativas de un fenómeno ordenadas en el tiempo (Martín-Pliego, 2011;pág. 449)
- El tiempo nos permite describir el comportamiento de la variable e intentar predecir sus valores futuros.
- La consideraremos una distribución bidimensional (X,t,n_{it})

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

1.2. Tipos de Magnitudes

Tipos:

(siempre a intervalos de t constantes)

- Se diferencian:
  - Flujo depende del intervalo de t entre dos observaciones.
  - Stock: el intervalo no la afecta en principio

Magnitudes stock:

- valores concretos en momentos concretos del tiempo, siendo X continua en el tiempo

Magnitudes Flujo:

- *Aquellas que representan el total acumulado de una variable desde la observación anterior*

Notación:

- y_t: serie de observaciones ordenadas correspondientes a t=(1,2,...,T)
- y_{ik}: el año i (i=1,2,...N) y la época del año k (k=1,2,...N)
  - (meses: k=1,2,...12); (trimestres k=1,2,..4)
- T=Nm

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2. Componentes de la serie temporal

Tendencia (T_{ik})

- movimiento general a largo plazo de la serie.

Variaciones estacionales (e_{ik})

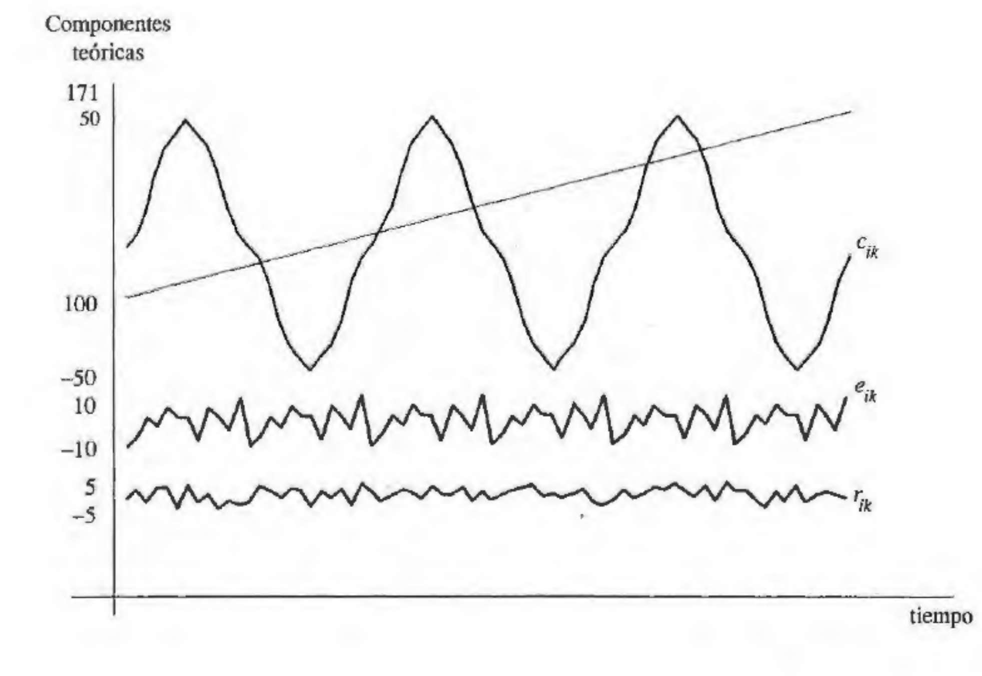
- movimientos en periodos inferiores o iguales al año y que se reproducen en los diferentes años (estaciones del año, clima, semanas o días…)

Variaciones Cíclicas (c_{ik})

- movimientos en periodos superiores al año y que se deben a etapas de prosperidad y depresión en la actividad económica (necesitamos muchas observaciones para detectarlas).

Variaciones Residuales (r_{ik}) o residuos

- variaciones irregulares o erráticas que no muestran un carácter periódico reconocible



Fte: (Martín-Pliego, 2011; pág. 454)

Componente extraestacional (E_{ik}): trata de forma conjunta la tendencia y el ciclo cuando en la práctica son difíciles de distinguir

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.1.Formas de combinar las componentes

Esquema	Modelo
Aditivo	$y_{ik} = T_{ik} + c_{ik} + e_{ik} + r_{ik}$
Multiplicativo I	$y_{ik} = T_{ik} \cdot c_{ik} \cdot e_{ik} \cdot r_{ik}$
Multiplicativo II o Mixto	$y_{ik} = T_{ik} \cdot c_{ik} \cdot e_{ik} + r_{ik}$

Supuesto básico: independencia de las variables residuales respecto a las demás componentes.

Independencia de errores se traduce en sumar los errores al modelo ($+r_{ik}$), lo que sucede en el esquema aditivo y el multiplicativo II.

- Si el modelo se ajusta mas al multiplicativo I → Hay que solucionar este problema antes de continuar.

(Martín-Pliego,2011; 455-458)

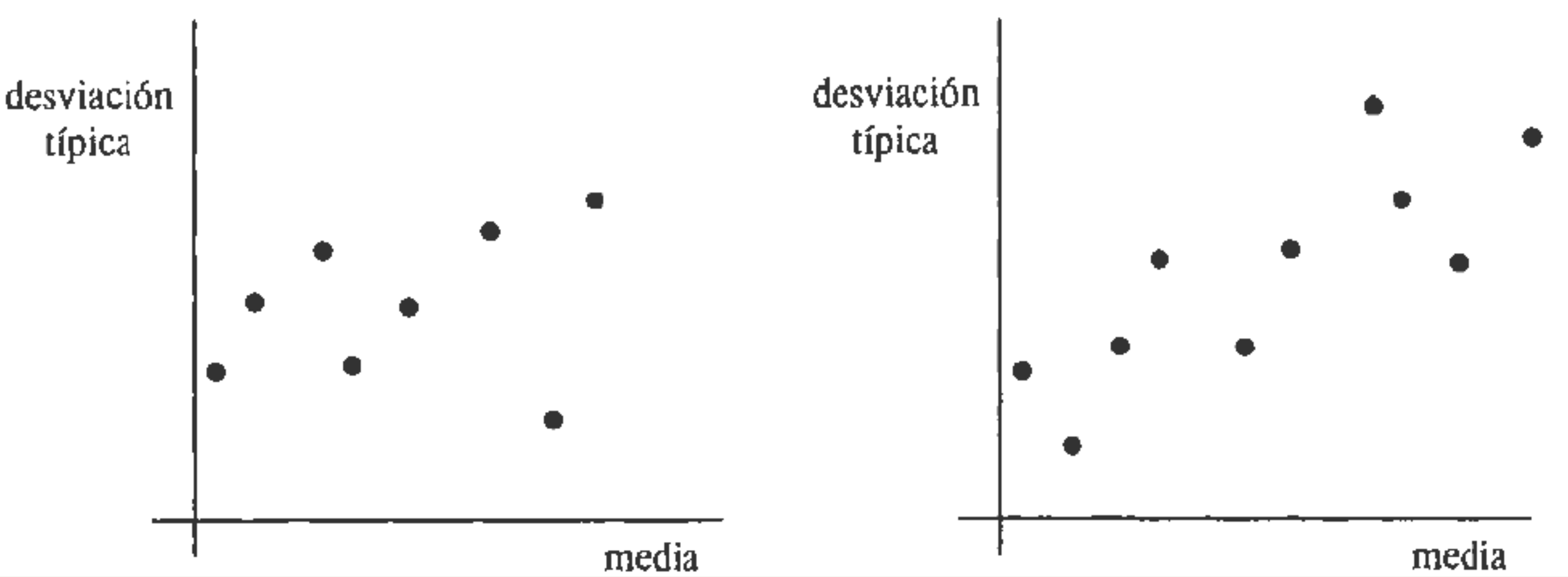
Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.1.Formas de combinar las componentes

- La independencia de los residuos va asociada a la relación que existe entre la media y la varianza de diferentes grupos de la serie.
  - Si dividimos la serie original en grupos de q observaciones puede ocurrir :

la desviación típica parece no depender de la media

la desviación típica parece depender de la media



(Martín-Pliego,2011; 455-458)

2.1.Formas de combinar las componentes

- ¿En la práctica como elegir "q"? Debemos elegir un valor que elimine la componente cíclica y estacional
  - *Para datos mensuales con un ciclo de un año y medio: q=36 (3 años de datos)*
- Hay que tener en cuenta la longitud de la serie:
  - *que puede hacer inviable ese valor teórico deseable de q.*
  - *Para conocer la componente cíclica (necesitamos muchos periodos)*
- Si concluimos que el modelo sigue un esquema Multiplicativo tipo I, la solución es tomar logaritmos y pasamos a un esquema aditivo.

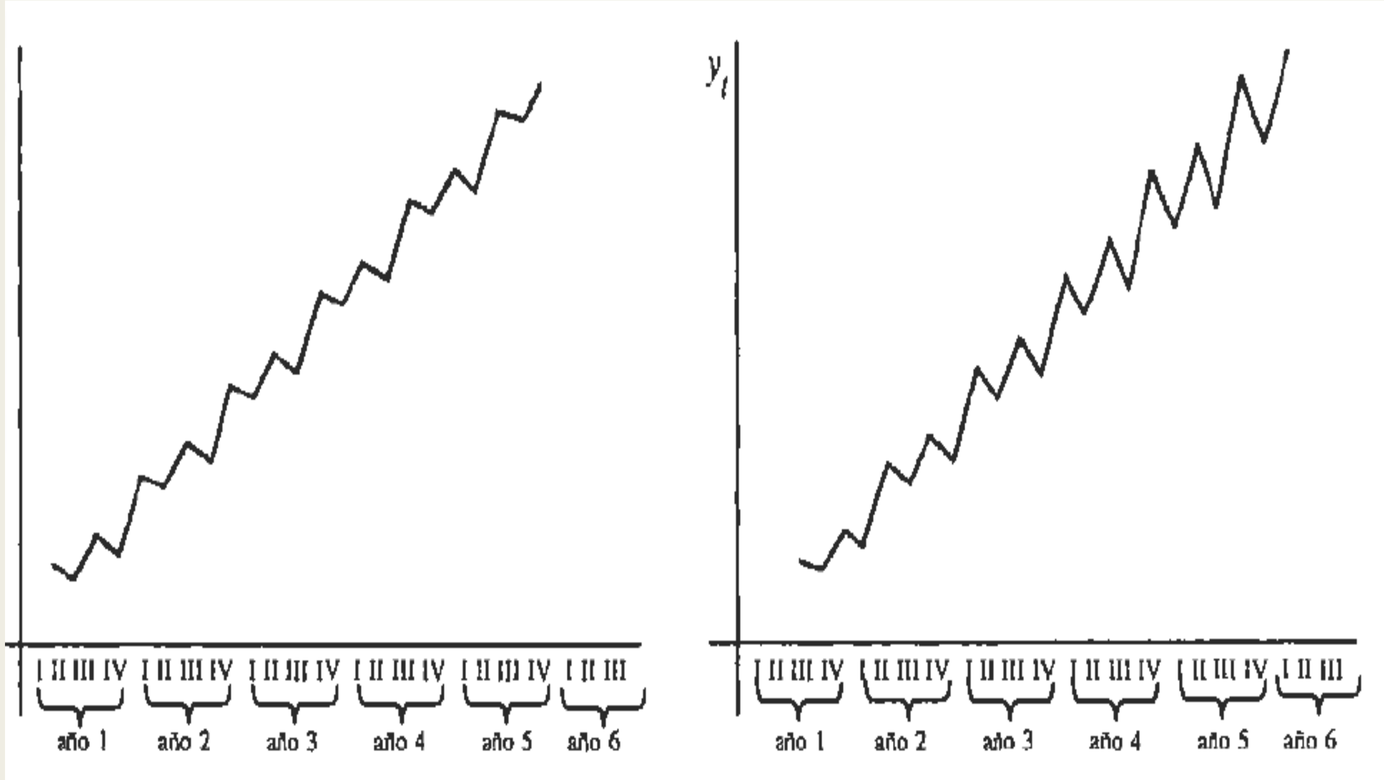
Modelo original	Transformación logarítmica
Multiplicativo I $y_{ik} = T_{ik} \cdot c_{ik} \cdot e_{ik} \cdot r_{ik}$	Aditivo $Y_{ik}^{*} = T_{ik}^{*} + c_{ik}^{*} + e_{ik}^{*} + r_{ik}^{*}$

(Martín-Pliego,2011; 457)

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

2.1.Formas de combinar las componentes

- La componente estacional se combina con la extraestacional de forma distinta:
  - *Aditivo: la oscilación estacional tiende a mantenerse con amplitud constante cada año*
  - *Multiplicativo II: la oscilación estacional tienden a crecer con mayores valores de la variable.*



(Martín-Pliego,2011; 457)

- Tenemos el mismo problema con la componente cíclica:
  - *Pero es difícil tener información sobre los ciclos en esta etapa del análisis.*
  - *En la práctica: lo relevante es la forma en que la componente estacional se sobrepone a la extraestacional en general*

2.1.Formas de combinar las componentes

- Resumen:
- En el esquema Aditivo y Multiplicativo II la componente residual es independiente de las demás → podemos aplicar técnicas estadísticas apropiadas.
- Es necesario realizar una transformación logarítmica en el esquema multiplicativo I para pasar a un esquema aditivo y poder aplicar estas técnicas.
- Cuando se habla de esquema multiplicativo, nos referimos al esquema multiplicativo II

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

2.2. Tendencia

Tendencia (T_{ik})

- Movimiento a largo plazo (no analizamos movimientos inferiores al año)
- Designaremos la serie temporal como Z_t ; t: periodo de tiempo parcial
- Metodología:
  - *Método de ajuste analítico*
  - *Método de Media móviles*
  - *Método de Diferencias*

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.2.1. Método de Ajuste Analítico

Tendencia (T_{ik})

- Ajuste por regresión de los valores de la serie a la variable tiempo a través de una función matemática que explique el comportamiento de la serie temporal
- Tipos de Ajuste:

Tendencia	Ajuste	Valido para
Lineal	$Z(t) = a + bt$	Comportamiento constante
Logarítmica	$Z(t) = a + \log(bt)$	Aumenta o disminuye rápidamente y luego se estabiliza
Semilogarítmica	$Z(t) = \log(a + bt)$	
Polinómica	$Z(t) = a + bt + ct^2 + \cdots + c^* t^n$	El orden del polinomio se determina por el número de fluctuaciones de la serie. • n=2 → un máximo o mínimo • n=3 → dos máximos o mínimos • N=4 → más de 3
Potencial	$Z(t) = at^b$	Para comparar medidas que aumentan a un ritmo concreto. Problema: cero o valores negativos.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.2.1. Método de Ajuste Analítico

Tendencia (T_{ik})

Tipos de Ajuste:

Tendencia	Ajuste	Valido para
Exponencial	$Z(t) = aEXP(bt)$	Los valores aumentan o disminuyen a intervalos cada vez mayores
Crecimiento	$Z(t) = EXP(a + bt)$	
Compuesta	$Z(t) = ab^t$	Tendencia exponencial simple con base a una constante b
Inversa o hiperbólica	$Z(t) = \mathrm{a} + \dfrac{b}{t}$	
Curva S	$Z(t) = EXP(a + b/t)$	Combinación de la exponencial e inversa
Logística	$Z(t) = 1/(a + bc^t)$	El parámetro a suele ser la inversa del límite superior logístico
Media Móvil	Atenúa las fluctuaciones en los datos Utiliza un periodo (número concreto de datos) sobre el que realiza un promedio como punto de línea	

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.2.2. Método de las media móviles

Tendencia (T_{ik})

- MA(p) analiza la tendencia de una serie temporal mediante medias en la siguiente forma:
- P es impar: forma medias relativas a los instantes (p+1)/2, (p+3)/2, (p+5)/2

$$\bar{y}_{\frac{p+1}{2}} = \frac{y_1 + y_2 + \cdots + y_p}{p} \qquad \bar{y}_{\frac{p+3}{2}} = \frac{y_2 + y_3 + \cdots + y_{p+1}}{p} \qquad \bar{y}_{\frac{p+5}{2}} = \frac{y_3 + y_4 + \cdots + y_{p+2}}{p}$$

- P es par: forma medias relativas a los instantes (p+1)/2, (p+3)/2, (p+5)/2... (que no serán valores enteros porque p es par). A continuación se hallan nuevas media móviles originales consecutivas, que ahora serán relativas a los instantes (p+2)/2, (p+3)/2, (p+6)/2, ... (que ya serán valores enteros porque p es par).

$$\bar{y}_{\frac{p+2}{2}} = \frac{\bar{y}_{\frac{p+1}{2}} + \bar{y}_{\frac{p+3}{2}}}{p} \qquad \bar{y}_{\frac{p+4}{2}} = \frac{\bar{y}_{\frac{p+3}{2}} + \bar{y}_{\frac{p+5}{2}}}{p}$$

- Una vez obtenida la serie de media móviles, la tendencia será las líneas que las unen.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.2.3. Método de las diferencias

Tendencia (T_{ik})

- $z_t = y_t - y_{t-1} = \nabla y_t$
- Si y_t no tiene tendencia: se mantiene entorno a un mismo valor. $E(z_t) = k$
- Si y_t tiene tendencia: $z_t$crece o decrece a lo largo del tiempo → tendríamos que seguir tomando diferencias.
  - $w_t = z_t - z_{t-1} = \nabla z_t = \nabla\nabla y_t = \nabla^2 y_t$ ... *y así sucesivamente hasta encontrar una serie aleatoria sin tendencia*
- en cada diferenciación de orden d perdemos d observaciones de y_t

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.3. Variaciones estacionales

Variaciones estacionales (e_{ik})

- Movimientos en periodos inferiores al año que se reproducen a lo largo de los años.
- Importancia: provoca una distorsión en el auténtico movimiento de las series económicas.
- Desestacionalización: tarea difícil que ha dado lugar a multitud de estudios y algoritmos complejos: X11 y X12 ARIMA del *Bureau of the Census* de EE.UU.
- Los métodos más sencillos (pero no tan completos):
  - *Medias móviles*
  - *Diferencias estacionales*
- Desestacionalización: tarea difícil que ha dado lugar a multitud de estudios y algoritmos complejos: X11 y X12 ARIMA del Bureau of the Census de EE.UU.
- Notación:
  - $y_t \rightarrow$ *orden natural de la serie (según aparecen los datos); t=1,2,… T*
  - $y_{ik} \rightarrow i$ *representa al año, y k el periodo (k=12→ año ; k=4→trimestre;…; k=m)*

Universidad Rey Juan Carlos

Variaciones estacionales (e_{ik})

2.3.1. Método de desestacionalización de la tendencia o de las relaciones medias mensuales respecto a la tendencia

- Etapas:
- Ajustar $\bar{y}_{i.} = a - bi$ por MCO a las medias anuales de los datos observados $\bar{y}_{i.} = \frac{\sum_{k=1}^{m} y_{ik}}{m}$
- Calcular las medias mensuales en los diferentes años $\bar{y}_{.k} = \frac{\sum_{i=1}^{N} y_{ik}}{m}$ $\forall k = 1,2 \ldots m$
- Aislar la comp. estacional con la serie de media mensuales corregidas $\overline{y'}_{.k} = \bar{y}_{.k} - \frac{b(k-1)}{m}$
- Calcular la media global corregida $\overline{y'} = \frac{\overline{y'}_{.1} + \overline{y'}_{.2} + \cdots + \overline{y'}_{.m}}{m}$
- Con esquema multiplicativo se calculan los índices de variación estacional $I_k = \frac{\overline{y'}_{.k}}{\overline{y'}} 100$ y se desestacionaliza la serie dividiendo sus valores por los índices de variación estacional. La componente estacional será $E_{ik} = \frac{I_k}{100}$
- Con esquema aditivo: la componente estacional del mes k es $E_{ik} = \overline{y'}_{.k} - \overline{y'}$

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.3.2. Método de desestacionalización del índice estacional

Variaciones estacionales (e_{ik})

- Existen varios métodos. Sólo vamos a ver el **método general del índice estacional** que consta de las siguientes etapas:
- Hallar tendencia por medias móviles tomando un año de periodo sobre la serie cronológica.
- Se centran los valores de las medias móviles en los instantes de tiempo originales.
- Se elimina la tendencia y la variación cíclica incluida en ella dividiendo toda la serie por los valores observados en cada periodo de repetición anual.
- Sobre estos últimos se calculan los índices de variación estacional en forma de porcentajes.
- Se divide la serie cronológica original por los índices de variación estacional correspondientes, y así obtendríamos la serie desestacionalizada.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.3.3. Método de desestacionalización de las medias móviles.

Variaciones estacionales (e_{ik})

- Obtener la componente extraestacional por MA(m) sobre la serie original para eliminar las variaciones estacionales.
- Un procedimiento simple podría ser el siguiente:
- Sea $X_t (t = 1,2,\dots,n)$ una serie temporal de periodo s (s=4 →trimestre; s=12 →meses). Una serie de s puntos, $X^*{}_t$, se obtiene a través de los siguientes pasos pasa s par:
  - *Para las media móviles de s puntos* $X^*_{t+0,5} = \frac{\sum_{j=-\left(\frac{s}{2}\right)+1}^{s/2} x_{t+j}}{s}; \left(t = \frac{s}{2}, \frac{s}{2}+1, \dots, n-\frac{s}{2}\right)$
  - *Para las media móviles centradas de s puntos* $X^*_t = \frac{X^*_{t-0,5}+X^*_{t+0,5}}{2}; \left(t = \frac{s}{2}, \frac{s}{2}+1, \dots, n-\frac{s}{2}\right)$

2.3.4. Método de las diferencias estacionales

- Elimina la mayor parte del efecto estacional de una serie
- Tomar la serie en diferencias de orden m (periodo estacional) $z_t = y_t - y_{t-m.} = \nabla_{m.} y_t$
- Importante: en cada diferenciación de orden m perdemos m observaciones de y_t

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.4. Variaciones cíclicas

Variaciones Cíclicas (c_{ik})

- La más difícil e detectar:
  - *Movimiento a largo plazo con periodo no fácilmente identificable*
  - *En muchos casos incluso variable*
  - *Es frecuente que varios ciclos se superpongan.*
- Identificación en la práctica:
  - *Se elimina la tendencia y la estacionalidad*
  - *La parte restante se analiza como* $k_{ik} = c_{ik} + r_{ik}$
  - *Podríamos incluso prescindir del doble subíndice porque no hay componente estacional.*
  - *Detectamos* c_{ik} *con métodos específicos como el análisis armónico.*

2.4.1. Variaciones cíclicas: análisis armónico.

- Onda armónica:

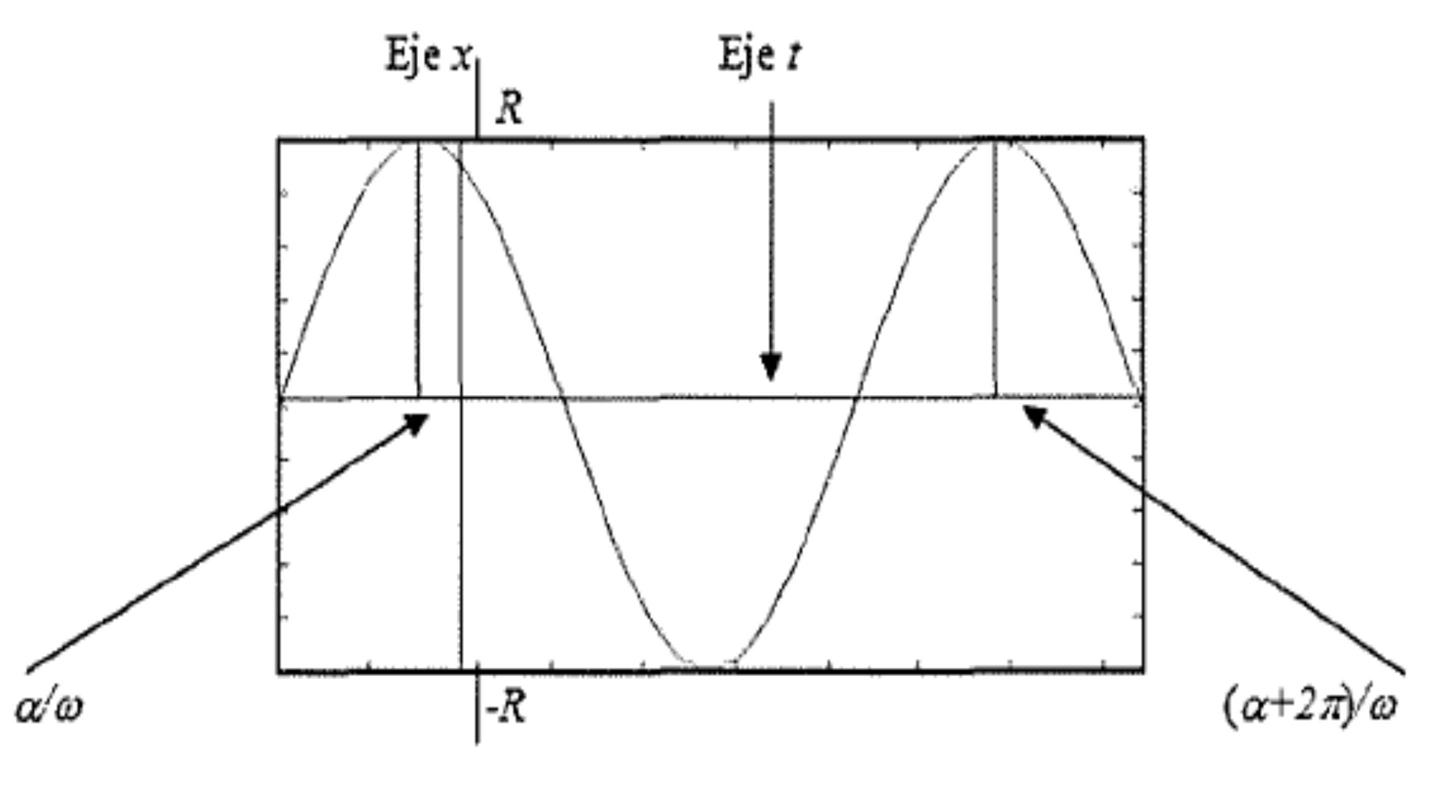
$$X_j = A\cos w_j + B\sin w_j$$

$$X_j = R\cos(w_j - \alpha)$$

- R se denomina amplitud, proporciona el valor máximo de X_j
- $\frac{2\pi}{w}$ es el periodo o intervalo de tiempo necesario para que se produzca una oscilación completa
- $\frac{w}{2\pi}$ frecuencia o número de oscilaciones que se producen entre dos momentos consecutivos del tiempo
- w es el ángulo en radianes
- α es la fase que marca X_j en el origen

- Ambas expresiones son equivalentes

$$R = (A^2 + B^2);\ \alpha = arctan\left({}^{B}/_{A}\right)$$



Fuente: TEP,CP;pág. 181

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.4.1. Variaciones cíclicas: análisis armónico.

Variaciones Cíclicas (c_{ik})

- Para detectar la existencia de un ciclo de orden p se forma el siguiente cuadro:

Oscilación	x_1	x_2		x_p
1ª	x_1	x_2	…	x_p
2ª	x_{p+1}	x_{p+2}	…	x_{2p}
…	…	…	…	…
qª	$X_{(q-1)p+1}$	$X_{(q-1)p+2}$	…	X_{qp}
Medias	$\bar{x}_1$	$\bar{x}_2$	…	$\bar{x}_p$

Fuente: TEP,CP;pág. 181

- Ajustamos los datos medios a:

$$\bar{x}_j = A_0 + A\cos\frac{2\pi}{p}j + B\cos\frac{2\pi}{p}j\ ;\ \forall j = \{1,2 \dots p\}$$

- La fila j-ésima recoge los p valores (no sig) que forman la j-ésima oscilación.
- q o número de oscilaciones, se obtiene dividiendo el número de oscilaciones de x_t por el periodo p_j.
- La última fila presenta las medias de los primeros elementos de cada oscilación (el valor medio de los segundos, etc…)
- A_0 constante para mejorar el ajuste
- $\frac{2\pi}{p}$ se incluye porque si el periodo es $\frac{2\pi}{w}$ entonces $w = \frac{2\pi}{p}$

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

2.4.1. Variaciones cíclicas: análisis armónico.

Variaciones Cíclicas (c_{ik})

- Recordemos: Ajustamos los datos medios a:

$$\bar{x}_j = A_0 + A\cos\frac{2\pi}{p}j + B\cos\frac{2\pi}{p}j\ ;\ \forall j = \{1,2\ldots p\}$$

- El ajuste por MCO resulta:

$$A_0 = \sum_{j=1}^{p}\frac{\bar{x}_j}{p}\ ;A = \frac{2}{p}\sum_{j=1}^{p}\bar{x}_j\cos\left(\frac{2\pi j}{p}\right);B = \frac{2}{p}\sum_{j=1}^{p}\bar{x}_j\,\mathrm{sen}\left(\frac{2\pi j}{p}\right)$$

- Para distintos valores de p se obtienen distintas amplitudes $\mathrm{R}(p)$ correspondientes a cada periodo.
- Los puntos [p,R(p)] forman el periodograma.

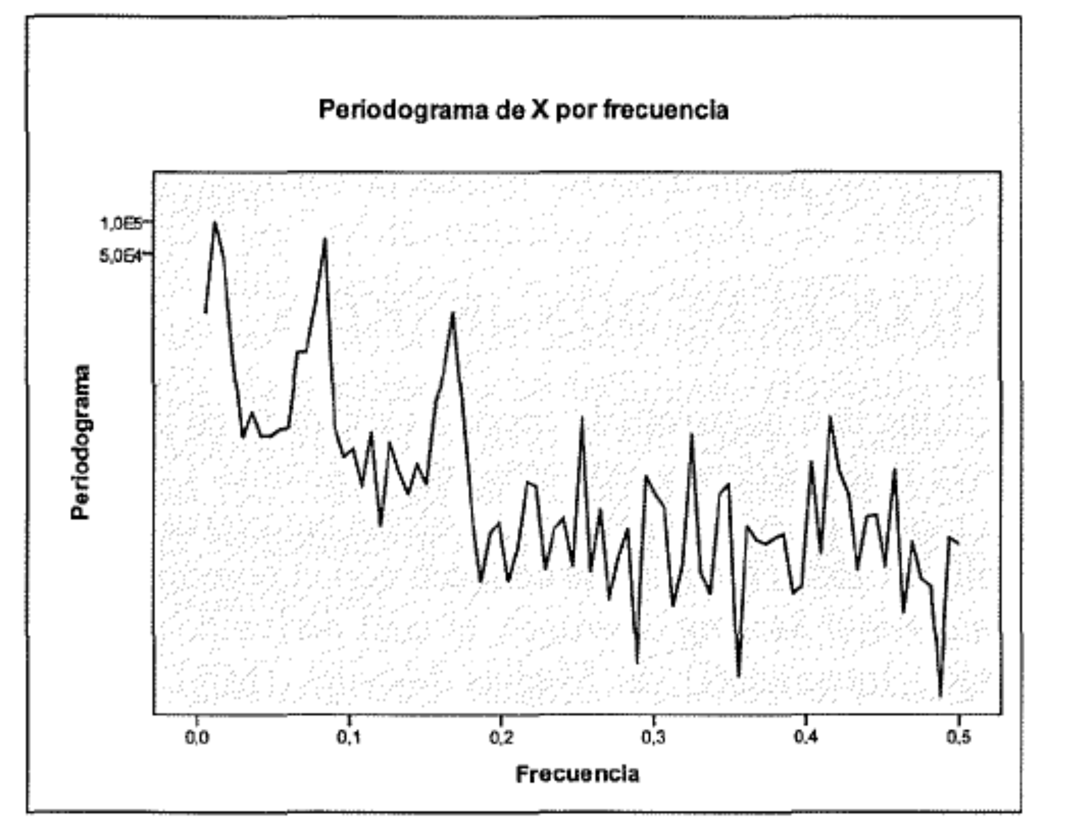
Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

2.4.1. Variaciones cíclicas: análisis armónico.

Variaciones Cíclicas (c_{ik})

- Para distintos valores de p se obtienen distintas amplitudes $\mathrm{R}(p)$ correspondientes a cada periodo.

- Los puntos [p,R(p)] forman el periodograma.
- El periodograma traslada la serie temporal al ámbito de las frecuencias (transformadas de Fourier)
- Picos:
  - No Destacables: no hay estacionalidad
  - Destacables: cada uno de ellos puede ser un ciclo.
- A cada amplitud destacable le corresponde una frecuencia cuya inversa es el periodo estacional o ciclo (CP).
- El periodograma identifica la longitud del periodo estacional, y de existir, la del ciclo.
- Las amplitudes:
  - Más fuertes: corresponden a ciclos o valores mas bajos de las frecuencias
  - Más débiles: suelen corresponder a estaciones o valores no tan bajos de las frecuencias



Fuente: TEP,CP;pág. 183

2.4.1. Variaciones cíclicas: análisis armónico.

- El periodograma acumulativo.
- Representa en el eje de abscisas las frecuencias y en el eje de ordenadas las amplitudes acumuladas.
- Si la serie es aleatoria, coincide con la diagonal del primer cuadrante.
- Desvíos bruscos de la diagonal provocan presencia de ciclos o estaciones para las respectivas frecuencias, que serán ciclos cuando las frecuencias sean bajas.

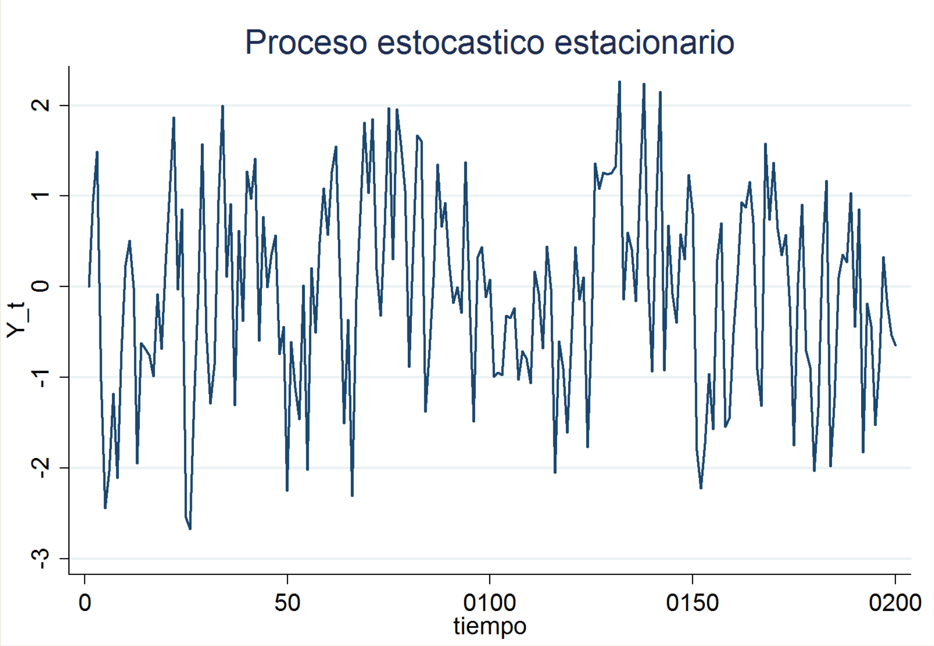
Fuente: TEP,CP;pág. 183

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3. Metodología Box-Jenkins

- Metodología ARIMA: la desarrollaron Box y Jenkins para series temporales a partir de la dependencia existente entre los datos.
- Cada observación es explicada por los valores anteriores.
- ARIMA = AR (*AutoRegresive*) + I (*Integrated*) + MA (*Moving Avarage*)
- Describe un valor como función lineal de datos anteriores + errores aleatorios.
- Puede incluir, además, un componente cíclico o estacional.
- Principio de parsimonia: contener los elementos mínimos necesarios para describir el modelo.
- Box-Jenkins recomienda un mínimo de 50 observaciones.
- Modelar una serie = inferir un modelo ARIMA que se ajuste al conjunto de datos.

Proceso estacionario: no existe cambio sistemático en media o en varianza a lo largo del tiempo.



Fuente: https://gdromero.wordpress.com/2013/10/19/predecir-el-futuro-series-de-tiempo/

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.1. Fases Box-Jenkins

	Fase	Consiste	Notas
1	Recogida de datos	> 50 datos	Series mensuales: entre 6 y 10 años
2	Representación gráfica	Juzgar la estacionariedad	Alternativa: medias y desviaciones típicas por subperiodo
3	Transformación de ST	Logaritmos si ST no es estacionaria en varianza	Práctica: incluso si la heteroscedasticidad es mínima. Se compara la serie original y logarítmica.
4	Eliminación de la Tendencia	Tomando diferencias	Primeras diferencias es lo mas habitual. Una tendencia no lineal, en la práctica lleva a dos diferencias como mucho.
5	Identificación del modelo	Determinar el orden de los AR,MA de las componentes regular y estacional	Decisión: Función de autocorrelación (FAC) y función de autocorrelación parcial (FAP). Práctica: se elige entre los modelos más simples AR(1), AR(2),MA(1),MA(2) y ARMA(1,1)

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.1. Fases Box-Jenkins

	Fase	Consiste	Notas
6	Estimación de coeficientes	Proceso Iterativo en base al modelo	Pueden sugerirse valores iniciales (debido a la iteración)
7	Contraste de validez conjunta	Contrastes de significación de los parámetros,	Covarianzas entre estimadores, Correlaciones, Suma cuadrática de los errores,…
8	Análisis detallado de errores	Diferencias históricas entre valores estimados y observados	Objetivo: ausencia de relación sistemática Observar: errores especialmente significativos.
9	Selección del modelo	Decisión basada en las etapas anteriores	
10	Predicción	Formula inicial de predicción	

3.2. Tipología ARIMA(p,d,q)

- Principio básico: el efecto de un conjunto de variables explicativas x sobre y debe ser estimado después de que se haya controlado el comportamiento pasado de la variable endógena (dependiente).
- Objetivo: examinar en qué medida las variables independientes explican la variación de la variable endógena que no viene explicada por la variación en el pasado de dicha variable endógena.
- En necesario precisar cual es el proceso que autogenera (y explica) la variación de la variable endógena a través del tiempo.
- Exigencia : todas las variables que participan en el modelo de regresión deben ser estacionarias (el tiempo no afecta a la media ni a la varianza de la ST).

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2. Tipología ARIMA(p,d,q)

- Función de transferencia: Análisis de regresión multivariante donde las variables son estacionarias con media y varianza constantes (se elimina la tendencia en el tiempo antes de incluir las variables en el análisis).
- Tomar diferencias para eliminar la tendencia:
  - *Tendencia lineal: 1 diferencia* $\Delta y_t = y_t - y_{t-1}$
  - *Tendencia cuadrática: 2 diferencias*

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$$

- Tomar logaritmos (o incluso el cambio porcentual: tasa) si la varianza no es constante.
  - *El gráfico cronológico de la serie nos puede ayudar a determinar si la serie es estacionaria o no.*
  - *Si hay dudas analizamos el gráfico de las primeras y segundas diferencias.*
- $I(d)$ Variable integrada de orden d: cuando la variable cronológica resulta ser estacionaria después de diferenciarla d veces.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2. Tipología ARIMA(p,d,q)

$$y_t = \delta + \boldsymbol{\phi_1 y_{t-1}} + \boldsymbol{\phi_2 y_{t-2}} + \cdots + \boldsymbol{\phi_p y_{t-p}} + \varepsilon_t \boldsymbol{- \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}}$$

- ¿cuál es el mejor modelo ARIMA?
  - *Análisis de la FAC y la FAP*
  - *Tipos de procesos:*
    - **AR(p)**
    - **MA(q)**
    - ARMA(p,q)
    - ARIMA (p,d,q)

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2. Tipología ARIMA(p,d,q)

$$y_t = \delta + \boldsymbol{\phi_1 y_{t-1}} + \boldsymbol{\phi_2 y_{t-2}} + \cdots + \boldsymbol{\phi_p y_{t-p}} + \varepsilon_t - \boldsymbol{\theta_1 \varepsilon_{t-1}} - \boldsymbol{\theta_2 \varepsilon_{t-2}} - \cdots - \boldsymbol{\theta_q \varepsilon_{t-q}}$$

- Cuatro parámetros desconocidos que deben ser estimados por máxima verosimilitud:
  - δ: *constante*
  - $\boldsymbol{\phi}$ *coeficientes AR;* $|\boldsymbol{\phi}| < \mathbf{1}$
  - $\boldsymbol{\theta}$ *coeficientes MA;* $|\boldsymbol{\theta}| < \mathbf{1}$
  - *Var(*$\boldsymbol{\varepsilon}$*): varianza de los residuos que deben ser sendero aleatorio.*
- Posteriormente:
  - *Contraste paramétrico de los coeficientes*
  - *Contraste no paramétrico: bondad del ajuste del modelo*

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2. Tipología ARIMA(p,d,q)

$$y_t = \delta + \boldsymbol{\phi_1 y_{t-1}} + \boldsymbol{\phi_2 y_{t-2}} + \cdots + \boldsymbol{\phi_p y_{t-p}} + \varepsilon_t \boldsymbol{- \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}}$$

- Los errores deben ser **ruido blanco**($\boldsymbol{\varepsilon_{it*}}$):
  - $\boldsymbol{\varepsilon_{it*} \rightarrow N(0, \sigma)}$
  - *Todos los errores deben estar incorrelados entre sí: linealmente independientes.*
- ¿Cómo determinamos que los residuos son ruido blanco?
  - *Análisis de la FAC y la FAP*
  - *Contraste de Box-Lung. H_0: el residuo es ruido blanco.*
- Si el error no es ruido blanco hay que re-espicificar el modelo ARIMA.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2. Tipología ARIMA(p,d,q)

$$y_t = \delta + \boldsymbol{\phi_1 y_{t-1}} + \boldsymbol{\phi_2 y_{t-2}} + \cdots + \boldsymbol{\phi_p y_{t-p}} + \varepsilon_t - \boldsymbol{\theta_1 \varepsilon_{t-1}} - \boldsymbol{\theta_2 \varepsilon_{t-2}} - \cdots - \boldsymbol{\theta_q \varepsilon_{t-q}}$$

- Identificar, estimar y diagnosticar el modelo ARIMA → se está explicando la variable endógena por sus efectos pasados.
- Ahora podemos introducir variables explicativas (x_{tk}: en el instante t con el retardo k). Si x_{tk}tiene un coeficiente asociado con alto nivel de significación entonces podemos deducir que tiene una inflencia significatica sobre la variable endógena y_t

3.2.1. AR(1)

$$(1 - \phi_1 B) W_t = \varepsilon_t$$

$$W_t - \phi_1 W_{t-1} = \varepsilon_t$$

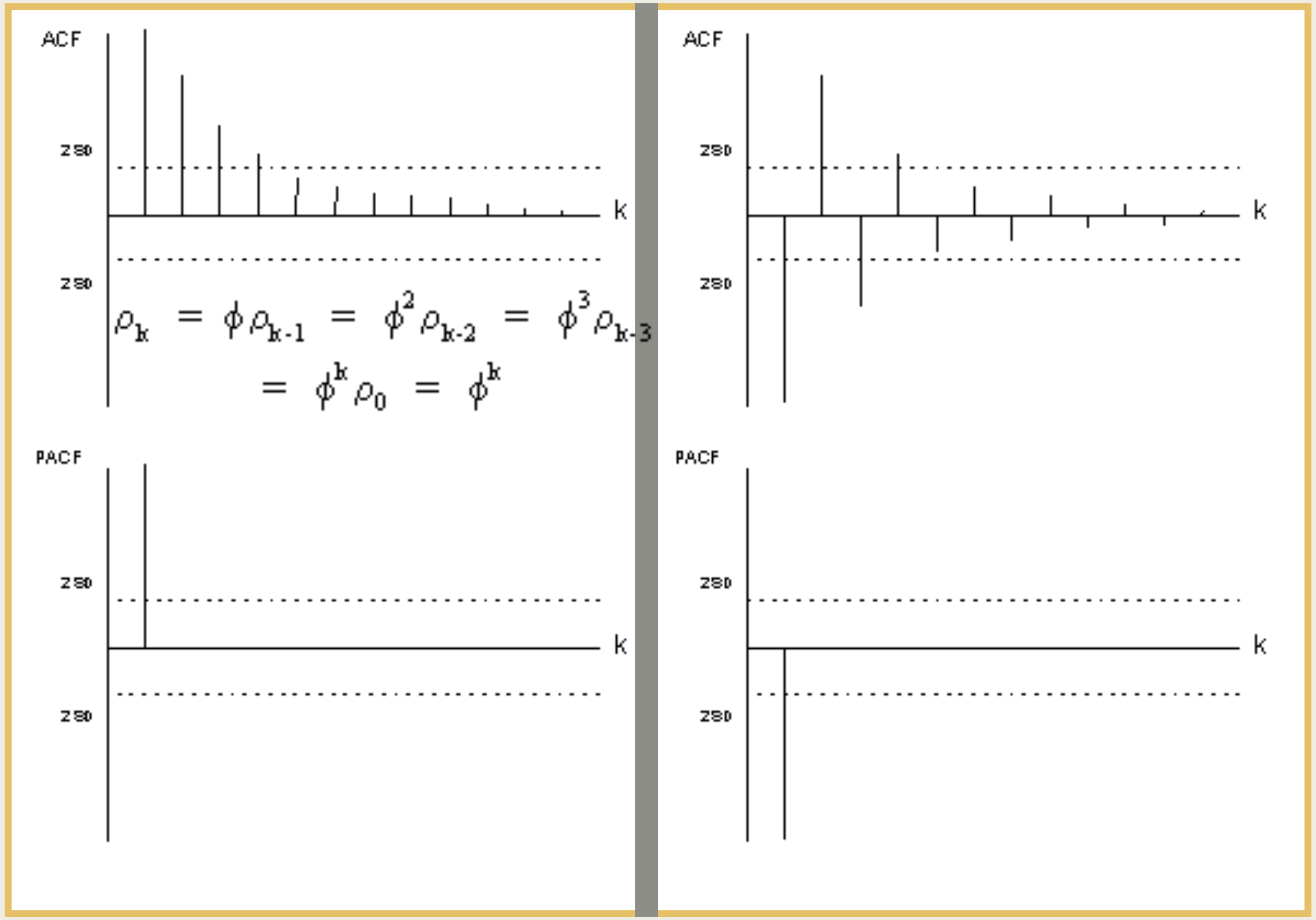
$$W_t - \varepsilon_t = \phi_1 W_{t-1}$$

$$F_t = \phi_1 W_{t-1}$$

$$\mathbb{CN}: |\phi_i| < 1$$

Donde:

- W_t es serie temporal estacionaria
- ε_t es ruido blanco
- F_t es la función de predicción



http://www.xycoon.com/ar1_process.htm

3.2.2. AR(2)

$$(1-\phi_1 B-\phi_1 B^2)W_t=\varepsilon_t$$

$$W_t-\phi_1 W_{t-2}-\phi_2 W_{t-2}=\varepsilon_t$$

$$W_t-\varepsilon_t=\phi_1 W_{t-1}+\phi_2 W_{t-2}$$

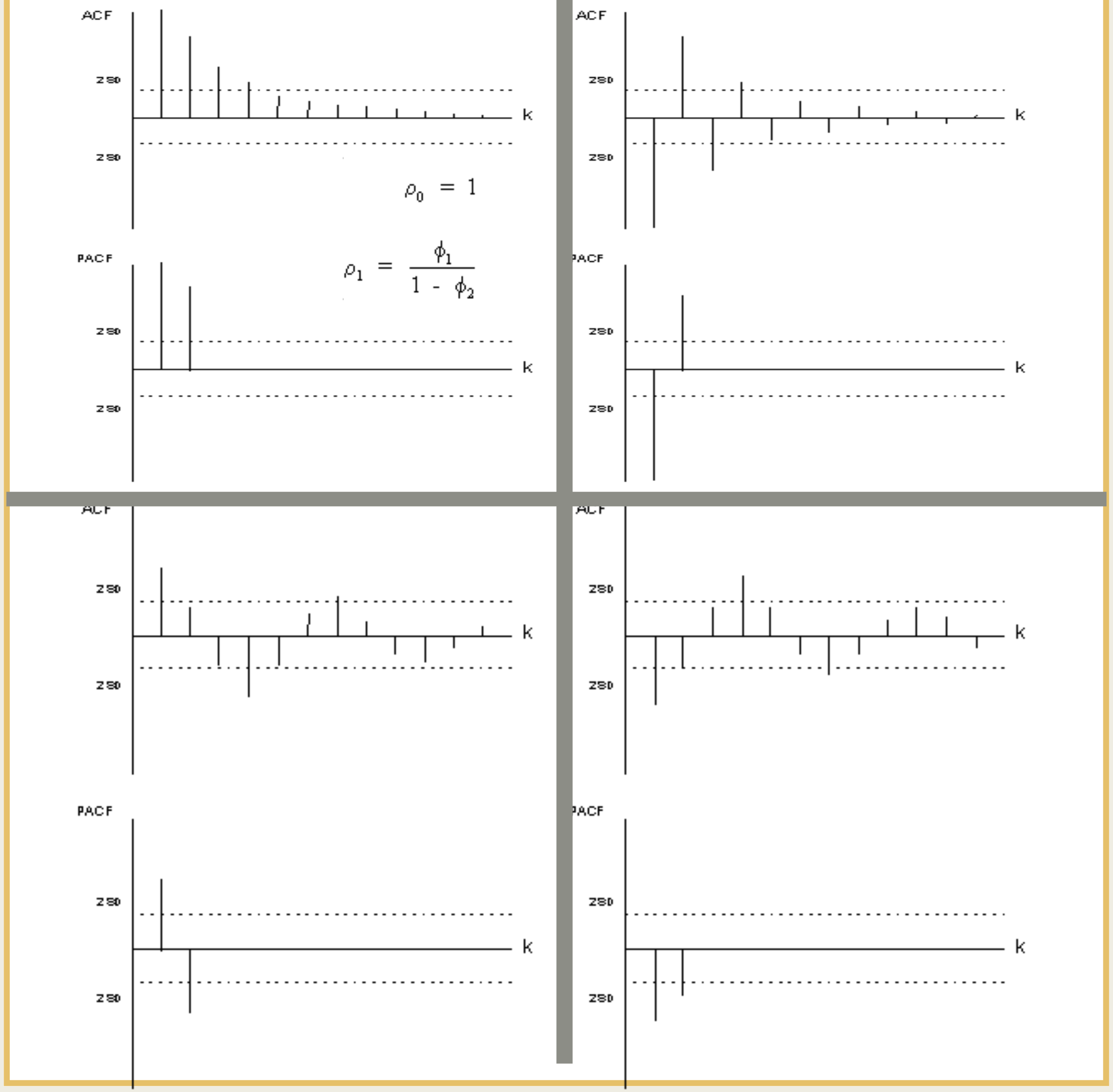
$$F_t=\phi_1 W_{t-1}+\phi_2 W_{t-2}$$

$$\mathbb{CN}:\begin{array}{c}|\phi_2|>-1\\ \phi_1^2+4\phi_2<0\end{array}$$

Donde:

- W_t es serie temporal estacionaria
- ε_t es ruido blanco
- F_t es la función de predicción

http://www.xycoon.com/ar2_process.htm



Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2.3. MA(1)

$$W_t = (1 - \theta_1 B)\varepsilon_t$$

$$W_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$W_t - \varepsilon_t = -\theta_1 (W_{t-1} - F_{t-1})$$

$$F_t = -\theta_1 W_{t-1} + \theta_1 \left(-\theta_1 (W_{t-2} - F_{t-2})\right)$$

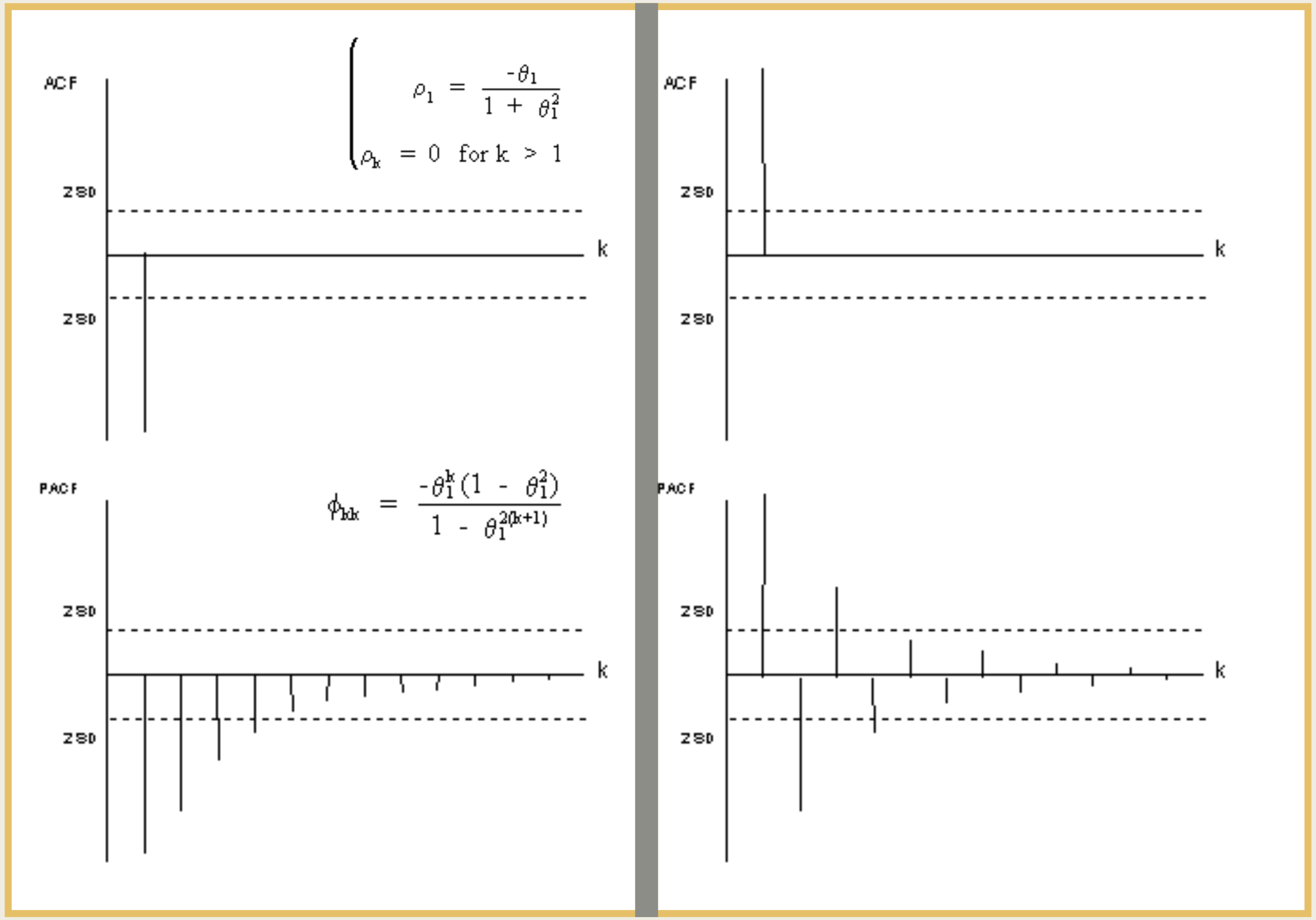
$$F_t = -\theta_1 W_{t-1} - \theta_1^2 W_{t-2} + \theta_1^2 F_{t-2}$$

$$\boxed{F_t = \sum_{i=1}^{\infty} \theta_1^i W_{t-i}}$$

$$\mathbb{CN}: |\theta_i| < 1$$

Donde:

- W_t es serie temporal estacionaria
- ε_t es ruido blanco
- F_t es la función de predicción



http://www.xycoon.com/ma1_process.htm

3.2.4. MA(2)

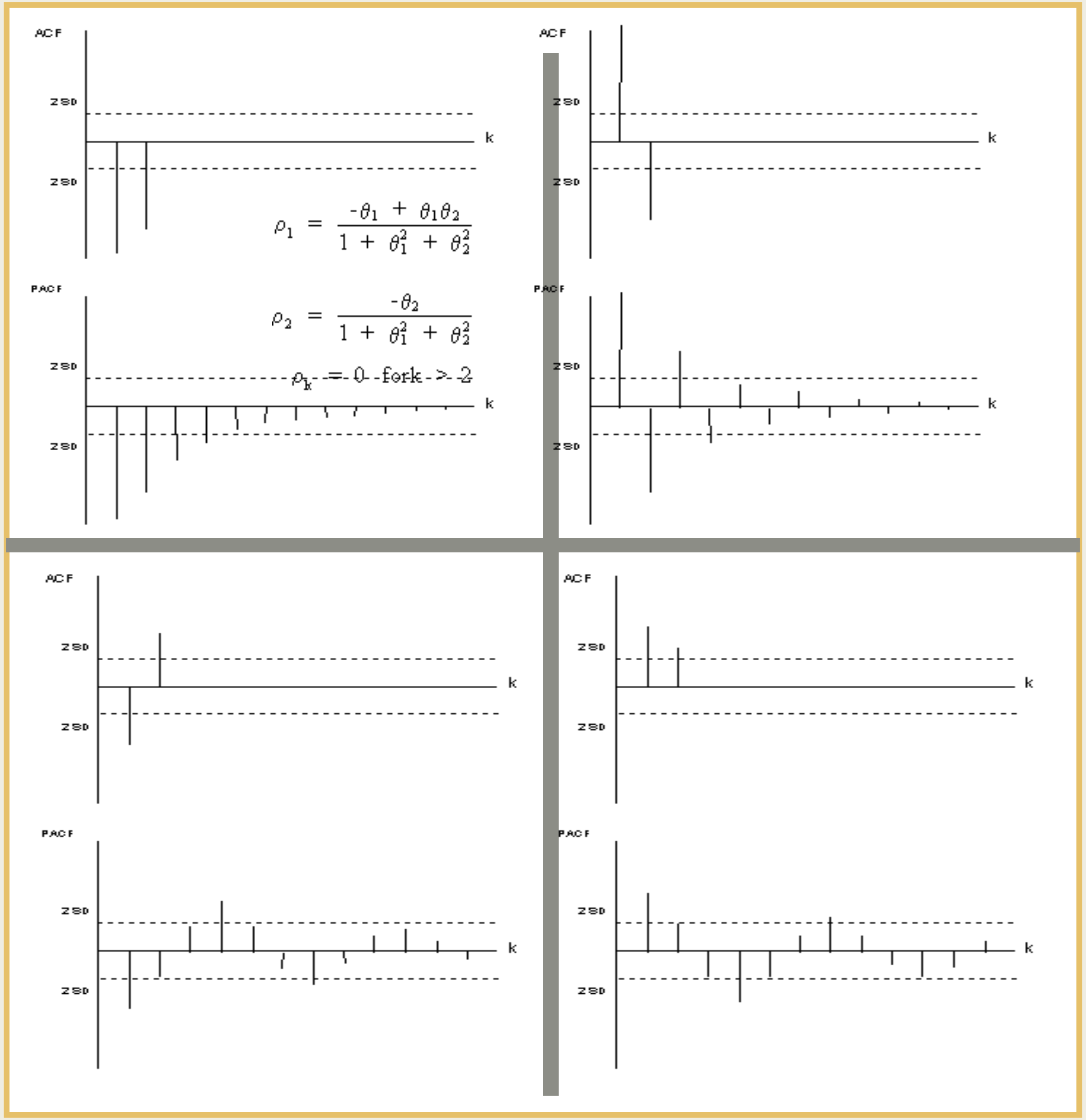
$$W_t = (1 - \theta_1 B - \theta_2 B^2)\varepsilon_t$$

$$W_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

$$F_t = \sum_{i=1}^{\infty} -\theta_1^i W_{t-1} + \sum_{i=1}^{\infty} -\theta_2^i W_{t-1-i}$$

Condiciones de invertibilidad:

$$\theta_1 + \theta_2 < 1$$
$$\theta_2 - \theta_1 < 1$$
$$-1 < \theta_2 < 1$$



Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2.5. ARMA(1,1)

$$(1-\phi_1 B)W_t=(1-\theta_1 B)\varepsilon_t$$

$$W_t-\phi_1 W_{t-1}=\varepsilon_t-\theta_1\varepsilon_{t-1}$$

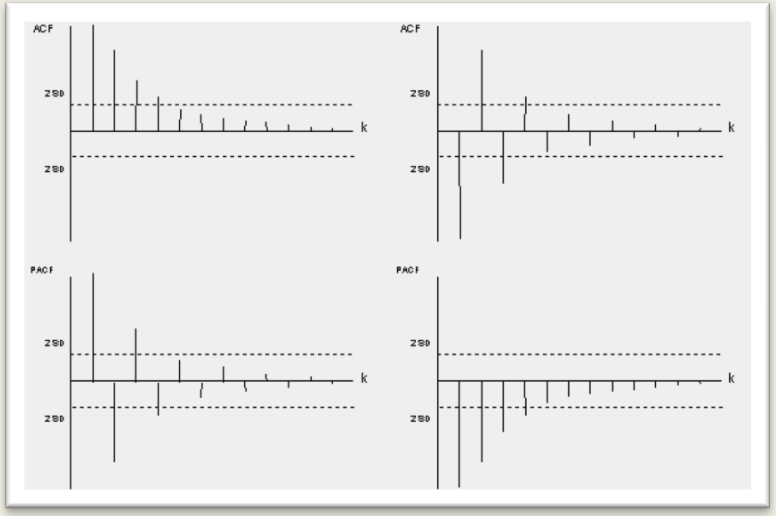
$$F_t=\phi_1 W_{t-1}-\theta_1\varepsilon_t-1$$

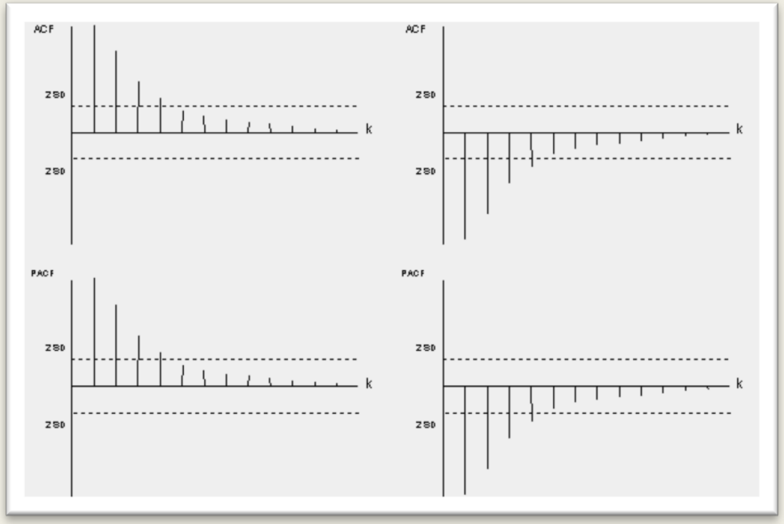
Función de autocorrelación

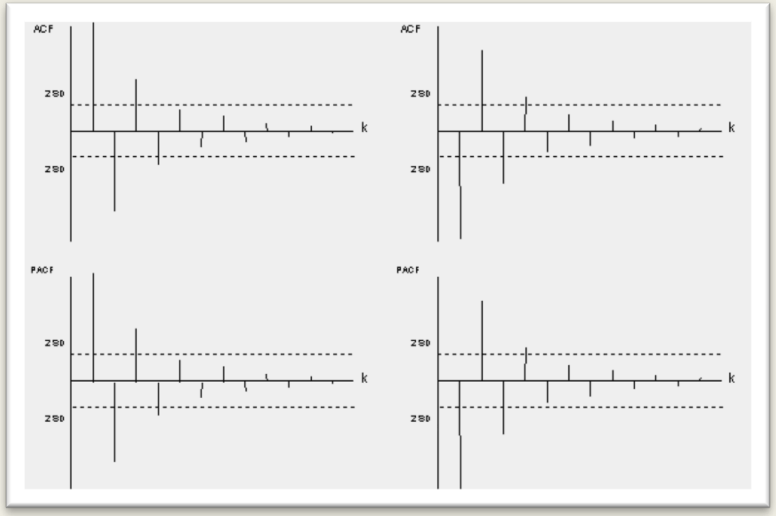
$$\rho_1=\frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$$

$$\rho_k=\phi_1\rho_{k-1}\quad \forall k>1$$

Función de autocorrelación parcial: se caracteriza por un decrecimiento exponencial.







http://www.xycoon.com/arma11_process.htm

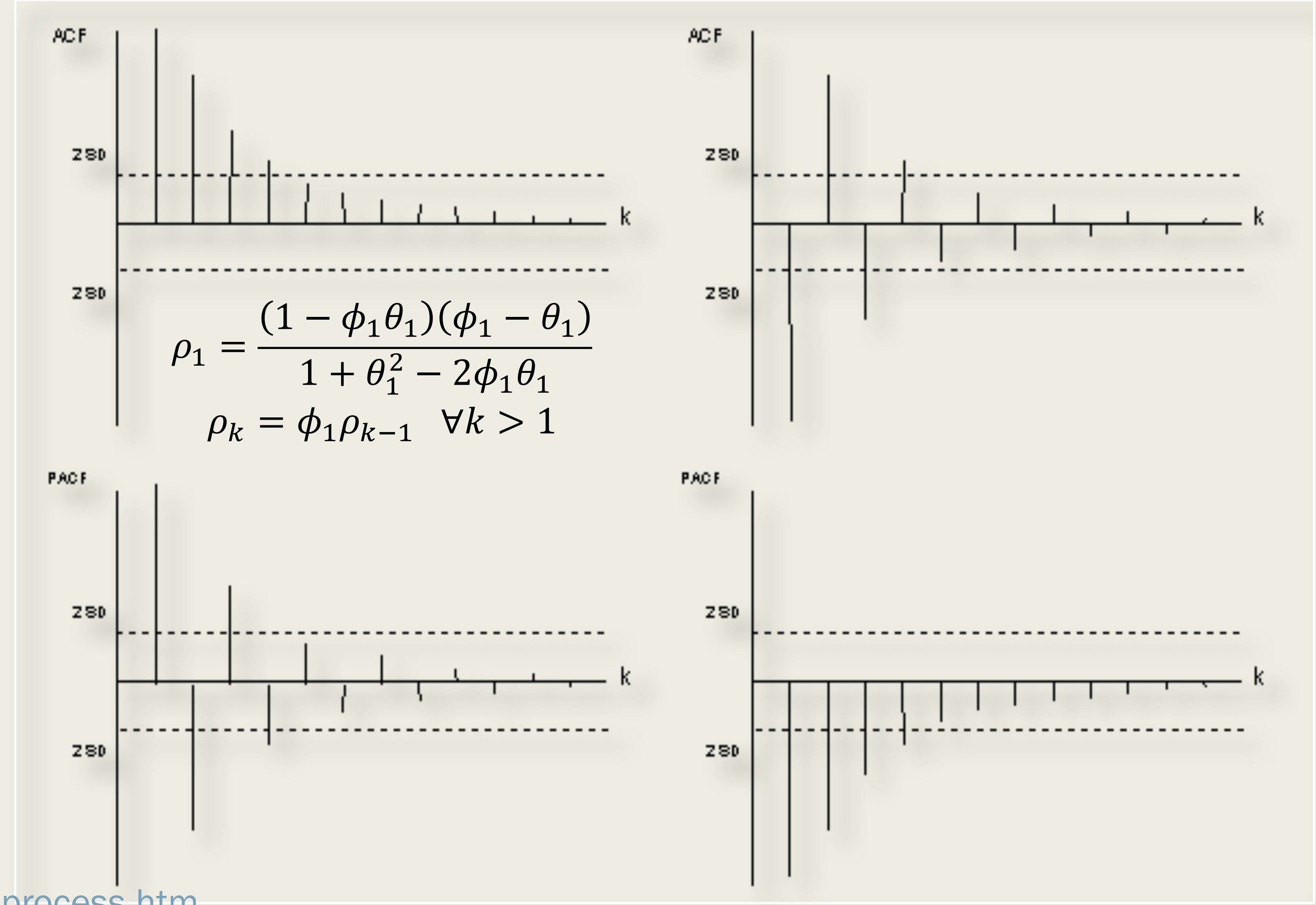
Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.2.5. ARMA(1,1)

$(1-\phi_1 B)W_t=(1-\theta_1 B)\varepsilon_t$

$W_t-\phi_1 W_{t-1}=\varepsilon_t-\theta_1\varepsilon_{t-1}$

$F_t=\phi_1 W_{t-1}-\theta_1\varepsilon_t-1$



$$\rho_1=\frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$$

$$\rho_k=\phi_1\rho_{k-1}\quad\forall k>1$$

http://www.xycoon.com/arma11_process.htm

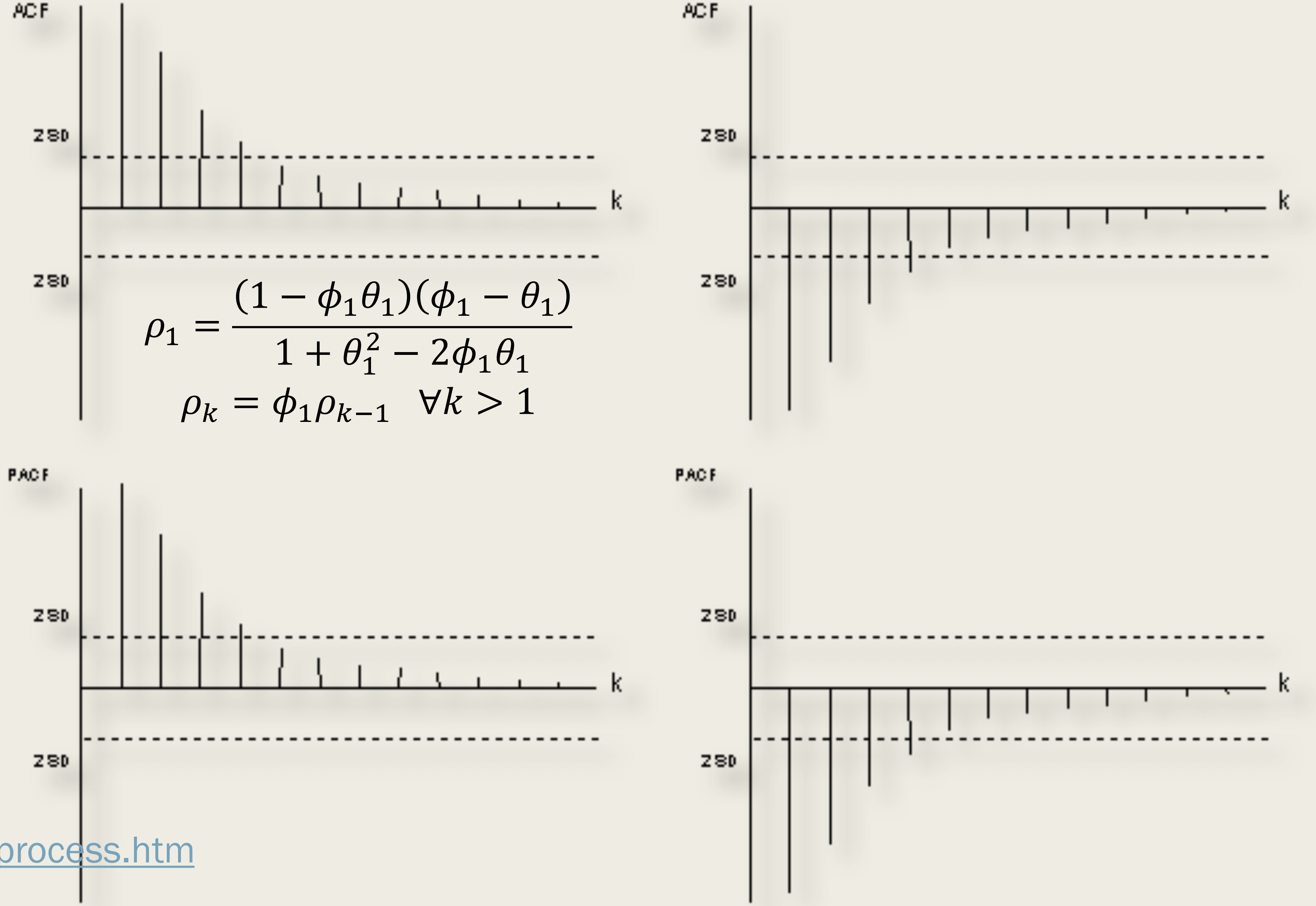
Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

3.2.5. ARMA(1,1)

$(1-\phi_1 B)W_t = (1-\theta_1 B)\varepsilon_t$

$W_t - \phi_1 W_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

$F_t = \phi_1 W_{t-1} - \theta_1 \varepsilon_t - 1$



$$\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$$

$$\rho_k = \phi_1 \rho_{k-1} \quad \forall k > 1$$

http://www.xycoon.com/arma11_process.htm

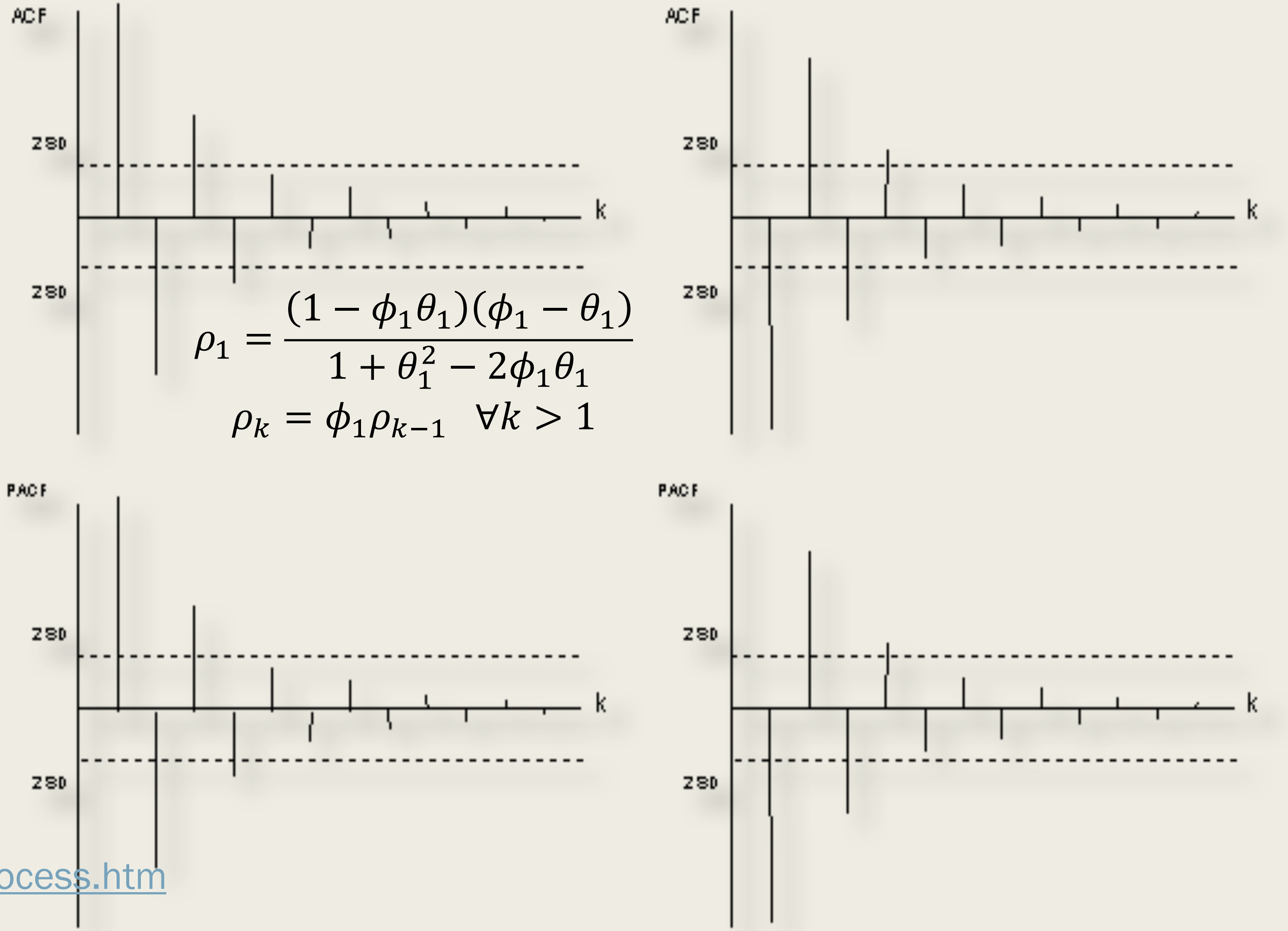
Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

3.2.5. ARMA(1,1)

$(1-\phi_1 B)W_t=(1-\theta_1 B)\varepsilon_t$

$W_t-\phi_1 W_{t-1}=\varepsilon_t-\theta_1\varepsilon_{t-1}$

$F_t=\phi_1 W_{t-1}-\theta_1\varepsilon_t-1$



$$\rho_1=\frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$$

$$\rho_k=\phi_1\rho_{k-1}\quad \forall k>1$$

http://www.xycoon.com/arma11_process.htm

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.3.Modelos de la función de transferencia.

$$\boldsymbol{y_t = \beta_0 + \sum_{i=1}^{k} \beta_k x_{kt} + ARIMA(p,d,q)}$$

- Parsimonia: Box-Jenkins aconsejan utilizar el ARIMA de menor orden posible.
- Se determina el ARIMA a través de las gráficas FAC y FAP
- Tras estimar el ARIMA por máxima-verosimilitud, hay que volver a estimar de forma conjunta la función de transferencia y evaluar el nivel de significación de los parámetros estimados.
- Criterio de validación de la función de transferencia:
  - *El residuo medio del modelo final debe ser menor que el del modelo ARIMA*
  - *El valor del criterio de información Akaike (AIC) debe ser el menor posible.*
  - *El término de error debe ser ruido blanco sin que los errores estén autocorrelacionados.*

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

3.3.Modelos de la función de transferencia.

$$\boldsymbol{y_t = \beta_0 + \sum_{i=1}^{k} \beta_k x_{kt} + ARIMA(p, d, q)}$$

- Box-Jenkins proporciona modelos poderosos respecto del poder predictivo de la variable endógena temporal, sin embargo…
- Aún no está suficientemente desarrollado para el caso de datos temporales de sección cruzada (datos de panel ó *data-panel*).
- No son módelos fáciles de estimar: en ocasiones, modelos similares se pueden identificar con modelos que incluyen retardos de la variable endógena y_t.

3.4.Identificación del modelo.

$$\boldsymbol{y_t = ARIMA(p,d,q)}$$

- Consiste en encontrar los valores p,d,q mas apropiados.
- Tendencia: diferencia $\boldsymbol{y_t}$ d veces.
- Después: análisis de los correlogramas FAC (ACF) y FAP (PACF).
  - *Autoregresivos AR(p): los primeros p términos de la PACF son distintos de cero y el resto son nulos.*
    - En la práctica: se considera que una muestra proviene de AR(p) si los términos de la FAP son casi cero a partir del que ocupa en lugar p. Un valor se considera casi cero si su módulo es inferior a $a/\sqrt{T}$
  - *Media Móvil MA(q): los primeros q términos de la FAC son dintintos de cero y el resto son nulos*

3.4.Identificación del modelo.

$$y_t = ARIMA(p, d, q)$$

1. Decidir si X_t necesita ser transformada para eliminar la estacionariedad: tomar logaritmos o transformación de Box-Cox.
2. Determinar el grado de diferenciación adecuado $\boldsymbol{d}$.
   – *La no estacionariedad →los coeficientes de la FAC tienden a decrecer muy lentamente. La pregunta es: ¿ cuán lentamente ha de ser el decrecimiento de los coeficientes de la FAC para que el proceso sea estacionario?*
   – *En general, sólo ocasionalmente los datos económicos del correlograma dejarán de decrecer tras las primeras diferencias, y en este caso sería necesario tomar segundas diferencias.*
   – *Una diferenciación innecesaria altera el esquema de autocorrelación evidente de una serie estacionaria y lo complica innecesariamente.*
3. Decidir p y q y si existe una componente estacional, decidir los órdenes de los operadores estacionales P y Q (observar el siguiente cuadro)

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.4.Identificación del modelo.

$$y_t = ARIMA(p, d, q)$$

Proceso	Función de Autocorrelación (FAC)	Función de autocorrelación parcial (FAP)
MA(q)	q primeros coeficientes significativos	Exponenciales atenuados u ondas sinusoidales
AR(p)	Exponenciales atenuados u ondas sinusoidales	Los p primeros términos significativos
ARIMA (p,d,q)	Comportamiento irregular en los retardos (1,...q) con q picos. Decrecimiento para retardos posteriores a q	Decrece (aproximadamente con exponenciales atenuados y ondas sinusoidales). No cero.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.5.ARIMA(p,d,q)(P,D,Q) estacionales

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

Recordemos:

- Las variaciones estacionales se producen en un periodo inferior al año, y se reproducen de manera reconocible en los diferentes años.
  - *Producen una distorsión en el autentico movimiento de* $\boldsymbol{y_t}$*.*
  - *Desestacionalización:*
    - tarea no trivial.. X11 y X12 del *Bureau of the Census de EE.UU*.
    - Métodos triviales: método de la tendencia, método de las medias móviles y el método de las diferencias estacionales

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.5.ARIMA(p,d,q)(P,D,Q) estacionales

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

Detectar la estacionalidad:

- Gráfico de la serie cronológica $\boldsymbol{y_t}$
- Gráfico de las subseries estacionales: identifica gráficamente los periodos estacionales.
- Gráfico de las subseries anuales, valida gráficamente los periodos estacionales.
- La FAC y la FAP estimadas también validan los periodos estacionales de acuerdo a:
  - *Los coeficientes de la ACF para retardos de múltiplos del periodo estacional de la serie deben ser significativamente distintos de cero.*
  - *Para una cantidad grande de retardos la ACF se configura en forma de abanico que completa su ciclo girando sobre el eje de abcisas para una cantidad de periodos igual al periodo estacional. La PACF debe presentar estructura de coeficientes significativos para retardos periódicos (largos).*

3.5.ARIMA(p,d,q)(P,D,Q) estacionales

$$y_t = ARIMA(p,d,q)(P,D,Q)$$

Detectar la estacionalidad:

- Periodograma: traslada la serie temporal del tiempo a las frecuencias.
  - *Se representan las frecuencias en el eje X y las amplitudes en el eje Y.*
  - *Si no presenta picos destacables no hay estacionalidad.*
  - A cada amplitud destacable le corresponde una frecuencia cuya inversa es el periodo estacional o el ciclo, con lo que *el periodograma identifica la longitud del periodo estacional y en su caso el ciclo.*
  - *Las amplitudes mas fuertes (correspondientes a valores no tan bajos de las frecuencias) suelen corresponder a estaciones.*
  - *Si hay dudas entre ciclos y estaciones podemos apoyarnos en las funciones de autocorrelación para discriminar.*

3.5.ARIMA(p,d,q)(P,D,Q) estacionales

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

Detectar la estacionalidad:

- Periodograma Acumulativo:.
  - *Se representan las frecuencias en el eje de abcisas (X) y las amplitudes ordenadas en el eje de ordenadas(Y).*
  - $\boldsymbol{y_t}$ *aleatoria → coincide con la diagonal del primer cuadrante*
  - *Desviaciones bruscas de la diagonal muestran presencia de ciclos o estaciones para las respectivas frecuencias, que serán ciclos cuando las frecuencias sean bajas.*
- Densidad Espectral: se interpreta igual que el periodograma.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

3.5.ARIMA(p,d,q)(P,D,Q) estacionales

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- ARIMA (p,d,q)(P,D,Q): modelo para series con tendencia secular y variaciones cíclicas o estacionales.
  - *(p,d,q) tendencia secular o parte regular de la serie*
  - *(P,D,Q) variaciones estacionales o parte estacional de la serie temporal.*
  - *P,D y Q se estiman igual que p,d y q, pero considerando la serie de valores periódicos con periodo* π $\{x_t, x_{t+\pi}, x_{t+2\pi}, \dots\}$ *en vez de la serie original* x_t*.*

3.6.Estimación ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- Método habitual: MCO.
- Si representamos $ARIMA(p,d,q)$ en la forma $\Phi(\boldsymbol{B})\boldsymbol{X_t} = \boldsymbol{v}(\boldsymbol{B})\boldsymbol{a_t}$, el objetivo es encontrar el vector de parámetros $\Phi = \left(\Phi_{\boldsymbol{1}}, \Phi_{\boldsymbol{2}}, \ldots \Phi_{\boldsymbol{p}}\right)$ y $\boldsymbol{v} = \left(\boldsymbol{v_1}, \ldots \boldsymbol{v_p}\right)$ que minimice la suma de cuadradados de los errores $\sum_i a_t^2 = S(\Phi, \boldsymbol{v})$
- La estimación es complicada porque la ecuación es no lineal en los parámetros. Debemos utilizar un método iterativo no lineal como por ejemplo del Marquardt.
- Las estimaciones previas a las iteraciones se obtienen por el método de los momentos.

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$y_t = \boldsymbol{ARIMA(p,d,q)(P,D,Q)}$$

- Contraste de sobreparametrización de Box-Jenkins: ajustar el modelo de orden superior al elegido y comprobar si los parámetros son significativamente distintos de cero.
- Las funciones de autocorrelación de los residuos (ACF y PACF) deben ser nulas en todo su recorrido excepto en cero.
- Si los residuos se comportan como un ruido blanco autocorrelado (problema análogo al encontrado, en los modelos econométricos con perturbaciones autocorreladas)) debemos emplear contrastes como el de **Durbin-Watson** (autocorrelación de primer orden) o el de **Wallis** (para la de cuarto orden)
- Otros contrastes: encaminados a comprobar si los residuos pueden ser considerados ruido blanco.

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- Box y Pierce

$$Q = \sum_{k=1}^{m} r_k^2 \quad ; \quad r_k^2 = \frac{\sum_{t=k+1}^{n} a_t a_{t-k}}{\sum_{t=k+1}^{n} a_t^2} \quad ; Q \sim \chi^2_{m-p-q}$$

- a_t: errores estimados ; n: número de observaciones; m es un valor arbitrario pero conviene que sea lo más elevado posible.
- H0: los a_t son ruido blanco
  - *Rechazamos para valores de Q muy altos.*
  - $RC_\alpha: P(Q > I) = \alpha$

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- Ljung y Box

$$Q' = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k} \quad ; \quad r_k^2 = \frac{\sum_{t=k+1}^{n} a_t a_{t-k}}{\sum_{t=k+1}^{n} a_t^2} \quad ; Q \sim \chi^2_{m-p-q}$$

- Idem anterior pero para **valores de m no muy grandes**
- a_t: errores estimados ; n: número de observaciones; m es un valor arbitrario pero conviene que sea lo más elevado posible.
- H0: los a_t son ruido blanco
  - *Rechazamos para valores de Q muy altos.*
  - $RC_\alpha : P(Q' > I) = \alpha$

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$y_t = \boldsymbol{ARIMA(p,d,q)(P,D,Q)}$$

- Inspección del gráfico de los residuos:
  - *Si los residuos proceden de un proceso ruido blanco → incorrelados entre sí →alternancia en signo sin ningún criterio obvio.*
  - *Por el contrario, rachas de residuos consecutivos de un mismo signo son debidos a: mala especificación del modelo (autocorrelación en los residuos o residuos no estacionarios).*
  - *Grafo* (t, a_t) *: si los residuos representados contra el índice de tiempo t presentan una tendencia conocida puede haber heteroscedasticidad en los residuos.*
- Resto de contrastes: aleatoriedad, autocorrelación, falta de linealidad y ausencia de normalidad en los residuos.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$y_t = ARIMA(p,d,q)(P,D,Q)$$

- **Periodograma de los residuos:** debe presentar amplitudes destacables en casi toda la gama de frecuencias.
- **Periodograma acumulativo de los residuos:** debe producir una curva de amplitudes sobre la recta de reposo sin presentar patrones de oscilación en ninguna zona de frecuencias.
- Contraste del modelo univariante residual: Conviene estimar el modelo excluyendo algunas observaciones al final de la muestra
  - *Si provoca una variación sensible en los parámetros estimados podría indicar una variación reciente en la estructura estocástica subyacente, lo que desaconsejaría el modelo para fines predictivos.*

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.7. Diagnostico, validación y contraste de ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- **Los ARMA deben ser estacionarios e invertibles.**
  - ***problema de raíces inferiores a la unidad :Si representamos el ARMA(p,q) en la forma*** $\Phi(\boldsymbol{B})\boldsymbol{X_t} = \boldsymbol{\upsilon}(\boldsymbol{B})\boldsymbol{a_t}$ *y alguna de las raíces de las ecuaciones* $\Phi(\boldsymbol{B}) = \boldsymbol{0}$ *y* $\boldsymbol{\upsilon}(\boldsymbol{B}) = \boldsymbol{0}$ *es menor que uno en módulo, el modelo es rechazable.*
  - ***problema de raíces unitarias:***
    - Si alguna de las raíces de $\Phi(\boldsymbol{B}) = \boldsymbol{0}$ es muy próxima a la unidad → la $\boldsymbol{y_t}$ puede estar **subdiferenciada** (precisa de alguna diferencia más).
    - Si alguna de las raices de $\boldsymbol{\upsilon}(\boldsymbol{B}) = \boldsymbol{0}$ es muy próxima a la unidad → la $\boldsymbol{y_t}$ puede estar **sobrediferenciada** (precisa de alguna diferencia más).
    - Si coincide una raíz en ambas ecuaciones, se puede cancelar un orden en el proceso, pasando a un ARMA(p-1,q-1)

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

3.8. Predicción en modelos ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- **Los ARIMA proporcionan: Predicción puntual** y **distribución de probabilidad** completa para los valores futuros de la serie.
- Consideramos predicción óptima aquella con mínimo error cuadrático medio.
- Elegimos nuestra predicción a horizonte $l, Z_t(l)$ tal que $E\left(e_t^2(l)\right) = E\left(X_{t+1} - Z_t(l)\right)^2$ fuese mínimo.
- Se puede demostrar que esta predicción coincide con la esperanza condicionada de X_{t+1} (regresión tipo I), es decir:

$$Z_t(l) = E\left({}^{X_{t+1}}/_{X_t, X_{t-1}, \dots X_1}\right)$$

3.8. Predicción en modelos ARIMA(p,d,q)(P,D,Q)

$$\boldsymbol{y_t = ARIMA(p,d,q)(P,D,Q)}$$

- **El cálculo real de la predicción** $Z_t(l)$puede hacerse de forma recursiva en la forma:

$$d_t = \Phi_1 d_{t-1} + \cdots + \Phi_p d_{t-p} + a_t - v_1 a_{t-1} - \cdots - v_q a_{t-q}$$

- d_t es la diferencia de orden d de X_t (supuesto X_t no estacionaria y convertible en estacionaria mediante un proceso de d diferencias consecutivas).
- Para calcular la predicción $Z_t(l)$ comenzamos calculando la estimación de $d_t(1)$ como la esperanza condicionada de d_{t+1}, posteriormente $d_t(2)$… y así sucesivamente hasta calcular la estimación de $d_t(I)$.
- Ahora podemos obtener la predicción de X_t sumando d_t d veces

$$Z_t(l) = \Phi_1 d_t + \Phi_{l+1} d_{t-1} + \Phi_{l+2} d_{t-2} + \cdots = Z_{t+1}$$

4. Métodos autoproyectivos deterministas de predicción

- Tipos de predicción temporal:
  - *Predicciones condicionales: modelos causales. Ej: regresión.*
  - *Predicciones incondicionales: métodos autoproyectivos.* y_t *sólo incluye valores pasados, presentes y futuros.*
    - Dos enfoques alternativos:
      - *El determinista o clásico.*
      - *El estocástico o moderno (Box-Jenkins)*
- Los Métodos autoproyectivos deterministas se utilizan para suavizar las irregularidades y fluctuaciones de una serie temporal a fin de obtener la línea de suavizado libre de variaciones estacionales y óptima para la predicción.

4. Métodos autoproyectivos deterministas de predicción

- El enfoque determinista es
  - *más adecuado cuando se dispone de un número limitado de observaciones.*
  - *Basado en modelos no paramétricos que no ponen atención a la estructura estocástica de la población → análisis descriptivo alejado de la inferencia*
- El enfoque estocástico:
  - *Modelización econométrica formal → probabilístico, inferencial.*
  - *Basado en modelos paramétricos construidos en fases: identificación, estimación, diagnosis formal y predicción.*
  - *Box-Jenkins (ARIMA)*

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad
Rey Juan Carlos

4.1. Suavizado por media móviles

- Se utiliza cuando no hay tendencia clara ni estacionalidad en la serie original.
- Objetivo: obtener la línea de tendencia
- Dada la serie temporal $X_t; t = 1,2 \ldots, T$ se define como serie media móvil de orden n

$$T_s = \frac{\sum_{i=s-\frac{n}{2}}^{s+\frac{n}{2}} x_i}{n} \; ; x = \left\{s - \frac{n}{2}, s + 1 - \frac{n}{2}, \ldots, T - \frac{n}{2}\right\}; n \; impar$$

- Ejemplo: n = 5

$$T_s = \frac{(X_{s-2} + X_{s-1} + X_s + X_{s+1} + X_{s+2})}{5} \; : s = \{3,4, \ldots, T-2\}$$

- La serie temporal T_s es una versión suavizada de X_t.

4.1. Suavizado por media móviles

- Si se elige bien el orden n de la media móvil y sería una representación correcta de las componentes a medio y largo plazo (ciclo-tendencia).
- El método es equivalente a ajustar una tendencia lineal. En el caso de n=5 , ajustar una tendencia lineal a cada cinco puntos consecutivos de la serie inicial y tomar en cada ajuste solamente el punto central de la recta ajustada para alisar la serie original.
- Habitualmente, n=4 o n=5 suelen alisar la serie lo suficiente como para intuir su tendencia y poder hacer predicciones.

Dr. Francisco Rabadán Pérez
Economía Aplicada I
Universidad Rey Juan Carlos

4.2. Suavizado exponencial de Brown

- El **modelo simple de alisado exponencial de Brown**
  - *obtiene predicciones de una serie temporal en función de las observaciones pasadas.*
- Sea $\mathbf{X_t}$ el valor observado de la serie temporal en el instante t. $\mathbf{S_t(l)}$ la predicción de $\mathbf{X_t}$ a horizonte $\boldsymbol{l}$. $\mathbf{S_t}$ va a ser un suavizado de la serie $\mathbf{X_t}$.
  - *La predicción se obtiene como:* $\mathbf{S_t(l) = aX_t + (1-a)S_{t-1}(l)\,; a \in (0,1)}$
    - $\mathbf{a \simeq 0} \rightarrow$ alisado más potente : para datos con fuertes fluctuaciones o gran aleatoriedad.
- En todos los modelos basados en el alisado exponencial se presenta el problema de fijar los valores iniciales ($\mathbf{S_0}$): habitualmente $\boldsymbol{S_1 = X_1}$
  - $\mathbf{a \simeq 0} \rightarrow \mathbf{S_0}$ *Influirá en el resultado durante muchos periodos.*
  - $\mathbf{a \simeq 1} \rightarrow \mathbf{S_0}$ *desaparece su influencia rápidamente, y es muy posible que los datos presenten tendencias o estacionalidad → la predicción no será muy acertada.*

4.2. Suavizado exponencial de Brown

- $x_{t+i} = d + ei + u_{t+i} \quad (i = 0, 1, 2, 3, \ldots)$
- El **modelo exponencial de Brown con tendencia lineal**

$$S'_t(l) = aX_t + (1-a)S'_{t-1}(l)\,;\, a \in (0,1)$$

$$S''_t(l) = aS'_t + (1-a)S'_{t-1}(l)$$

- **Predicción** $S_t(l) = a_t + b_t l$; donde a_t y b_t son estimadores de d y e.
- Fijado a, habitualmente entre 0,1 y 0,3
  - *damos un valor a* $S'_{t=1}$; *criterios:*
    - $S''_1 = X_1$
    - S'_1 = promedio de los primeros valores de X.
  - *damos un valor a* $S''_{t=1}$; *criterios*
- Estimadores:
  - $a_t = 2S'_t - S''_t\,;\, b_t = \frac{S'_t - S''_t}{1-a}$

4.2. Suavizado exponencial de Brown

- $x_{t+i} = d + ei + fi^2 + u_{t+i} \quad (i = 0, 1, 2, 3, \dots)$
- El modelo exponencial de Brown con tendencia cuadrática: $a \in (0, 1)$

$$S'_t(l) = aX_t + (1-a)S'_{t-1}(l)$$
$$S''_t(l) = aS'_t + (1-a)S'_{t-1}(l)$$
$$S'''_t(l) = aS''_t + (1-a)S''_{t-1}(l)$$

- Predicción $S_t(l) = p_t + q_t(l) + \frac{1}{2}(r_t)l^2$:
  - $p_t = 3S'_t - 3S''_t + S'''_t$
  - $q_t = \frac{a[(1-5a)S'_t - (10-8a)S''_t]}{2(1-a)^2}$
  - $r_t = \frac{a^2[S'_t - 2S''_t + S'''_t]}{(1-a)^2}$
- Valores iniciales: $S'_1 = S''_1 = S'''_1 = X_1$
- Sirve para predecir series con puntos de cambio de tendencia (turning points). Los métodos con tendencia lineal y simple no son válidos para este fin.

4.3. Suavizado lineal de Holt

- Al igual que Brown, realiza predicciones bajo el supuesto de tendencia lineal.
- Utiliza dos parámetros $\mathbf{a} \in (\mathbf{0}, \mathbf{1})$ y $\mathbf{b} \in (\mathbf{0}, \mathbf{1})$
- Predicción:
  - $F_t(l) = S_{t-1} + (b_{t-1})l \quad t > 2$
  - $S_t = aX_t + (1-a)[S_{t-1} - b_{t-1}]$
  - $b_t = b[S_t - S_{t-1}] + (1-b)b_{t-1}$
- Valores iniciales: $S_1 = x_1; b_1 = x_6 x_1$

- $\mathbf{N}$: número de observaciones
- x_t: observación t de la serie de tiempo
- S_t: observación t de la serie alisada
- $F_t(l)$: Predicción en el instante t a horizonte l
- b_t parámetro estimado en el instante t
- a primera constante de alisado (relacionado con la componente aleatoria)
- b segunda constante de alisado (relacionado con la tendencia)

4.4. Suavizado estacional de Winters

- Generalización del método de Holt para datos con variaciones estacionales.
- *Predicción:* $F_t(l) = S_t + l(b_t)I_{t+l-L}$;
  - L: nº observaciones anuales
  - I_{t+l-L} es el factor estacionalidad (modelo multiplicativo)
- Fórmulas de actualización:
  - $S_t = \frac{aX_t}{I_{t-L}} + (1-a)[S_{t-1} + b_{t-1}]$
  - $b_t = b(S_t - S_{t-1}) + (1-b)b_{t-1}$
  - $I_t = \frac{cX_t}{S_t} + (1-c)I_{t-L}$
- Tendencia:

$$b_1 = \frac{\left[\frac{(X_{L+1} - X_1)}{L} + \frac{(X_{L+6} - X_1)}{L} + \cdots + (X_{l+L} - X_L)\right]}{L}$$

- Tres parámetros:
  - a: componente aleatoria
  - b: componente tendencia
  - c: componente estacional.
- Necesitamos
  - al menos L periodos para establecer los índices estacionales iniciales
  - Al menos 2L periodos (dos años consecutivos) para establecer la tendencia