



mr
manual de referencia

ECONOMÍA | EMPRESA

PREDICCIÓN EN EL DOMINIO DEL TIEMPO

ANÁLISIS DE SERIES TEMPORALES PARA INGENIEROS

Juan Carlos García Díaz

EMPRESA

EDITORIAL
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Predicción en el dominio del tiempo

Análisis de series temporales para ingenieros

Juan Carlos García Díaz

2016

EDITORIAL
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Colección *Manual de Referencia*

Los contenidos de esta publicación han sido evaluados mediante el sistema *doblo ciego*, siguiendo el procedimiento que se recoge en:
<http://www.upv.es/entidades/AEUPV/info/891747normalc.html>

Para referenciar esta publicación utilice la siguiente cita: GARCÍA DÍAZ, J.C. (2016). *Predicción en el dominio del tiempo. Análisis de series temporales para ingenieros*. Valencia: Universitat Politècnica de València

© Juan Carlos García Díaz

© 2016, Editorial Universitat Politècnica de València
distribución: Telf.: 963 877 012 / www.lalibreria.upv.es / Ref.: 0482_04_01_01

Imprime: Byprint Percom, sl

ISBN: 978-84-9048-438-8

Impreso bajo demanda

La Editorial UPV autoriza la reproducción, traducción y difusión parcial de la presente publicación con fines científicos, educativos y de investigación que no sean comerciales ni de lucro, siempre que se identifique y se reconozca debidamente a la Editorial UPV, la publicación y los autores. La autorización para reproducir, difundir o traducir el presente estudio, o compilar o crear obras derivadas del mismo en cualquier forma, con fines comerciales/lucrativos o sin ánimo de lucro, deberá solicitarse por escrito al correo edicion@editorial.upv.es.

Impreso en España

Dedicatoria:

... a Ana, Carlos y Javier

Índice

Introducción	VII
Capítulo 1. Análisis descriptivo de series temporales.....	1
1.1. Introducción.....	1
1.2. Componentes de una serie temporal	2
1.3. Métodos de descomposición de una serie temporal.....	13
1.3.1. Método clásico descomposición	13
1.3.2. Descomposición STL.....	13
1.3.3. Desestacionalización.....	14
1.4. Series temporales con doble o triple estacionalidad	25
1.5. Parámetros y criterios de evaluación de la bondad en el ajuste de modelos.	27
Ejercicios propuestos.....	33

Capítulo 2. Técnicas de suavizado de series temporales	35
2.1. Introducción.....	35
2.2. Medias móviles.....	35
2.2.1. Medias móviles centradas	36
2.2.2. Medias móviles unilaterales	37
2.3. Suavizado exponencial	42
2.3.1. Suavizado exponencial simple	42
2.3.2. Modelo de Holt.....	46
2.3.3. Modelo de Holt-Winters.....	47
Ejercicios propuestos	58
Capítulo 3. Introducción a los modelos ARIMA	61
3.1. Introducción.....	61
3.2. Procesos estocásticos	61
3.3. Identificación de la autocorrelación en series temporales	62
3.4. Propiedades de un proceso estocástico	63
3.5. Modelos lineales	67
3.6. Modelos auto regresivos (AR).....	70
3.7. Modelos de medias móviles (MA).....	76
3.8. Modelos autorregresivos de medias móviles (ARMA)	78
3.9. Modelos ARIMA no estacionarios	82
3.10. Modelos ARIMA estacionales.....	88
3.11. Modelos ARIMA generales.....	91
Ejercicios propuestos	96

Capítulo 4. Modelización ARIMA.....	99
4.1. Metodología Box-Jenkins para el ajuste de modelos ARIMA	99
4.2. Obtención de una serie estacionaria	100
4.2.1. Varianza constante.....	100
4.2.2. Valor medio constante	100
4.2.3. Presencia de estacionalidad en el comportamiento de la serie.....	100
4.3. Identificación de un modelo tentativo	103
Capítulo 5. Estimación de modelos ARIMA.....	119
5.1. Estimación del modelo	119
5.2. Análisis de la significación de los parámetros.....	119
5.3. Análisis de la estacionariedad e invertibilidad de los modelos	120
Ejercicios propuestos.....	127
Capítulo 6. Diagnóstico y validación de modelos ARIMA	129
6.1. Introducción.....	129
6.2. Normalidad de los residuos y media cero	130
6.3. Varianza constante de los residuos	131
6.4. Residuos incorrelacionados e independientes	132
6.4.1. Los contrastes de Box-Pierce y de Ljung-Box.....	132
6.4.2. Contrastes de aleatoriedad de los residuos: tests basados en rachas	133
6.4.3. Contrastes sobre los residuos disponibles en Statgraphics.....	135
6.5. Peridiograma acumulado	138
Ejercicios propuestos.....	165

Capítulo 7. Predicción de series temporales	169
7.1. Introducción.....	169
7.2. Predicción puntual	169
7.3. Predicción por intervalos de confianza.....	170
7.4. Actualización de las predicciones.....	170
Ejercicios propuestos	186
Capítulo 8. Análisis de intervención y detección de valores atípicos.....	187
8.1. Introducción.....	187
8.2. Análisis de intervención	188
8.2.1. Variable impulso	188
8.2.2. Variable escalón.....	190
8.3. Valores atípicos	191
8.3.1. Atípicos aditivos (AO)	192
8.3.2. Atípicos innovativos (IO).....	192
8.3.3. Cambio de nivel (LS).....	193
8.3.4. Cambio transitorio (TC).....	194
8.3.5. Métodos de detección de valores atípicos	194
Capítulo 9. Modelos de regresión dinámica. Función de transferencia-ruido	205
9.1. Introducción.....	205
9.2. Modelo función de transferencia-ruido.....	206
9.3. Concepto de ganancia de la función de transferencia.....	208
9.4. Método del preblanqueo	213
9.5. Identificación de los parámetros (b, r, s)	219

9.6. Método LFT (modelo de la Función de Transferencia Lineal).....	227
9.7. Validación del modelo de función de transferencia-ruido.....	233
Ejercicios propuestos	235
Anexo 1. Metodología Box-Jenkins.....	237
Anexo 2. Tabla valores críticos distribución chi-cuadrado.....	238
Anexo 3. Análisis y modelización con Statgraphics Centurion	240
Bibliografía.....	253

Introducción

Predicciones y decisiones

La predicción constituye una de las herramientas fundamentales para la toma de decisiones en cualquier ámbito, desde el doméstico hasta el empresarial pasando incluso por el ámbito político o institucional. El ser humano decide en función de previsiones realizadas de forma consciente o inconsciente. Detrás de una decisión hay una previsión anterior. Así por ejemplo, la mayoría de decisiones empresariales relativas a la localización de plantas industriales, capacidad productiva instalada o los planes maestros de producción para los próximos meses, están condicionadas a la previsión de la demanda realizada. El operador del sistema eléctrico español (Red Eléctrica de España) realiza predicciones horarias de la demanda de energía eléctrica para el día siguiente con el objetivo de poder planificar el mix de fuentes de energía disponibles que cubran eficazmente dicha demanda de forma económica, fiable y segura.

Una serie temporal es una secuencia cronológica de observaciones sobre una variable en particular. Por lo general, las observaciones se realizan a intervalos regulares (horas, días, meses, años). Ejemplos comunes de series de tiempo son el Producto Interno Bruto, tasa de desempleo, y carga de pasajeros de las aerolíneas. Un análisis de series de tiempo se compone de dos pasos: (1) la construcción de un modelo que representa una serie de tiempo, y (2) utilización del modelo para predecir los valores futuros de dicha variable.

Tal y como argumenta Daniel Peña en su libro *Análisis de series temporales* (Peña, 2005), la metodología actual para analizar series temporales es fruto de la confluencia

de varias líneas de trabajo desarrolladas en distintos campos científicos. Desde el estudio de series temporales astronómicas y climáticas, pasando por el desarrollo de los métodos de suavizado exponencial para prever series de producción, inventarios y ventas en la década 1960-70, por la predicción y control de sistemas lineales, desarrollada en ingeniería de control y automática, y por la teoría de procesos no estacionarios y no lineales, llegando finalmente hasta los modelos de series temporales multivariantes.

Definición de previsión, predicción o pronóstico

Prever, predecir o pronosticar y previsión, predicción o pronóstico, pueden utilizarse como sinónimos y se pueden utilizar indistintamente ya que según la *Real Academia Española* de la lengua dichas entradas corresponden a:

- Prever: Ver con anticipación. Conocer, conjeturar por algunas señales o indicios lo que ha de suceder.
- Previsión: Acción y efecto de prever. Acción de disponer lo conveniente para atender a contingencias o necesidades previsibles.
- Predecir: Anunciar por revelación, ciencia o conjetura algo que ha de suceder.
- Predicción: Acción y efecto de predecir. Palabras que manifiestan aquello que se predice.
- Pronosticar: Conocer por algunos indicios lo futuro.
- Pronóstico: Acción y efecto de pronosticar. Señal por donde se conjetura o adivina algo futuro.

En inglés el término adecuado a utilizar es *forecasting* como predecir o prediciendo y *forecast* como predicción.

De las definiciones anteriores podemos extraer como definición práctica de predecir la siguiente: “Predecir es estimar, en presencia de incertidumbre, los valores futuros de una variable temporal conociendo sus valores pasados”.

Objetivos del manual

El objetivo general de este libro es la presentación de técnicas estadísticas de análisis, modelización y predicción de series temporales orientadas a la toma de decisiones para la Gestión y Dirección de Empresas Industriales. Estas herramientas son presentadas de forma teórica y aplicadas en numerosos ejemplos y ejercicios prácticos realizados con software estadístico comercial estándar como Statgraphics o Minitab. Los objetivos generales se muestran a continuación:

1. Describir la evolución temporal observada de una característica de interés de una unidad observable determinada.
2. Prever la evolución futura (desconocida) a corto plazo de dicha característica, utilizando tan sólo información (conocida) sobre su pasado. Esta información constituye lo que se conoce como una serie temporal univariante.

Este manual pretende que los lectores adquirieran las siguientes competencias:

- Comprender el significado de la realización de previsiones estadísticas y los riesgos asociados a ellas.
- Reconocer la estructura y componentes del comportamiento de una serie temporal.
- Tener una visión de las distintas técnicas a su alcance y de la idoneidad de cada una de ellas.
- Ser capaces de realizar previsiones con las mejores condiciones que la metodología estadística permite.
- Conocer y manejar correctamente el software disponible para la realización de previsiones.
- Ser capaces de evaluar el valor de las previsiones obtenidas, el horizonte temporal de las mismas y su grado de precisión.

Estructura del manual

Este manual está pensado para analizar los ejemplos desarrollados con conocimientos previos en el análisis de series temporales en el dominio del tiempo. En cada capítulo se hace una breve introducción teórica a los métodos utilizados en el mismo sin pretender sustituir a una buena descripción teórica de dichas herramientas que puede seguirse en cualquier manual recogido en la bibliografía.

El manual está estructurado de la forma siguiente. El primer capítulo introduce el análisis descriptivo como primer paso en el análisis de cualquier serie temporal. Seguidamente se presentan los métodos clásicos de suavizado, desde la medias móviles hasta el modelo de Holt-Winter. En el capítulo 3 y hasta el 7 se presentan los modelos ARIMA y la metodología de ajuste de dichos modelos desde la identificación, modelización, estimación y validación de dichos modelos mediante la conocida metodología Box-Jenkins. El capítulo 8 se introduce el Análisis de Intervención y detección de valores atípicos, y finalmente se cierra el manual con el capítulo 9 relativo a los modelos de regresión dinámica o de función de transferencia-ruido. El manual se completa con anexos que permiten la comprensión de los ejercicios desarrollados en el texto.

Capítulo 1

Análisis descriptivo de series temporales

1.1. Introducción

Entendemos por Serie Temporal la secuencia de valores que una variable toma a lo largo del tiempo de forma equiespaciada. El consumo horario de energía eléctrica en MWh en España, el precio horario de la energía eléctrica en euros por MWh en España, las ventas mensuales de un determinado producto en una empresa en miles de euros, la temperatura ambiental media diaria a lo largo del año, serían ejemplos de series temporales cuyos valores futuros sería interesante poder conocer antes de que se produzcan.

El primer paso en el análisis de series temporales siempre ha de ser un análisis descriptivo del comportamiento pasado de la misma. Todo análisis de series temporales ha de iniciarse con una representación gráfica de la misma, utilizando los ejes cartesianos, de forma que en el eje de abscisas representaremos el tiempo y en el de ordenadas, la serie observada Z_t . Obtendremos una serie de puntos (t, Z_t) que, al unirlos, nos dan una visión del patrón de comportamiento de la serie de la que se pueden sacar unas primeras conclusiones de la evolución histórica de la misma [García Díaz, 2011b].

El estudio descriptivo de series temporales se basa en la idea de descomponer la variación de la serie en varias componentes básicas.

1.2. Componentes de una serie temporal

En el estudio clásico de las series temporales se considera que el comportamiento de la serie es fruto de la participación de cuatro componentes, la tendencia, las variaciones cíclicas, las variaciones estacionales y las variaciones aleatorias [González y del Puerto, 2009].

Tendencia (T): es una componente de la serie que refleja su evolución a largo plazo. Esta componente, en el conjunto de toda la serie, puede ser de naturaleza estacionaria o constante, de naturaleza lineal, de naturaleza exponencial, u otras posibilidades. Ejemplos de este comportamiento pueden verse en las Figuras 1.3 y 1.5.

Componente cíclica (C): es una componente de la serie que recoge las oscilaciones periódicas de amplitud superior a un año. Estas oscilaciones no son regulares y se presentan en los fenómenos económicos cuando se dan de forma alternativa etapas de crecimiento o recesión. Solamente es caracterizable cuando se disponga de una larga historia de la serie, y en la práctica es difícil de ajustar lo cual no significa que dicha componente no esté condicionando el comportamiento de la misma.

Componente estacional (S): es una componente de la serie que recoge oscilaciones que se producen en periodos de repetición iguales o inferiores a un año originalmente en series con datos mensuales como se puede observar en el ejemplo 6. Su nombre proviene precisamente de las estaciones climatológicas. Pero podemos encontrarnos con estacionalidad distinta a la mensual como la semanal, diaria u horaria En el ejemplo 1.1 podemos ver un claro ejemplo [García Díaz, 2011a].

Componente aleatorio o irregular (I): es una componente de la serie temporal que recoge las fluctuaciones debidas a la ocurrencia de eventos imprevisibles. Es por definición imprevisible y actúa en cualquier serie temporal en mayor o menor medida.

Ejemplo 1.1. Serie con componente estacional pura sin tendencia. Análisis descriptivo de la serie temporal consumo en MWh de energía eléctrica horaria en España de lunes a viernes (CEEsem).

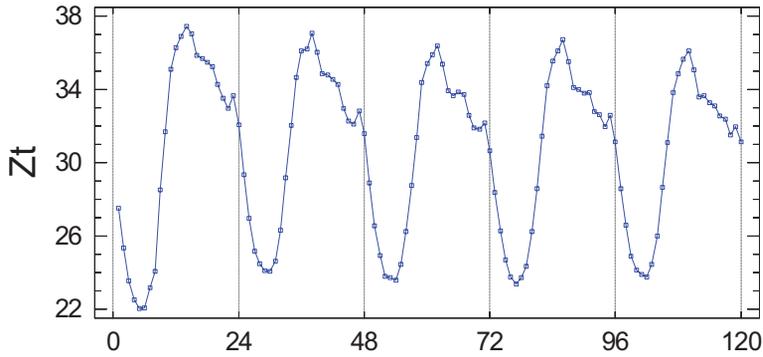


Figura 1.1. Serie $Z_t = \text{CEEsem}$ del consumo en MWh de energía eléctrica horaria en España de una semana laborable de lunes a viernes

Como se puede observar en la Figura 1.1 esta serie presenta una componente estacional de periodo o longitud 24 horas sin apenas tendencia. La Figura 1.2 es el llamado box-plot estacional donde se refleja como el consumo eléctrico es máximo entre las 13 y 14 horas y mínimo sobre las 5 horas.

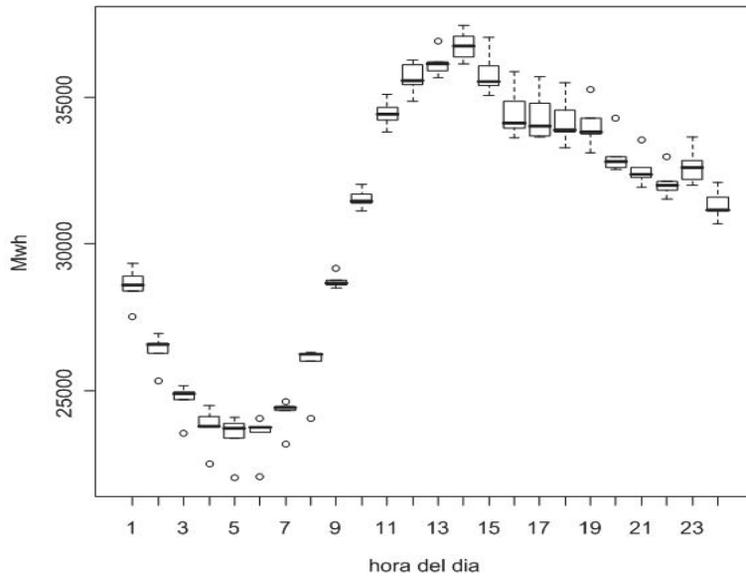


Figura 1.2. Box-plot estacional de CEEsem

Ejemplo 1.2. Serie con tendencia y posible estacionalidad. Análisis descriptivo de la serie temporal producción mensual en toneladas de cemento en España entre 1988 y 2004 ($Z_t = \text{"cemento"}$).

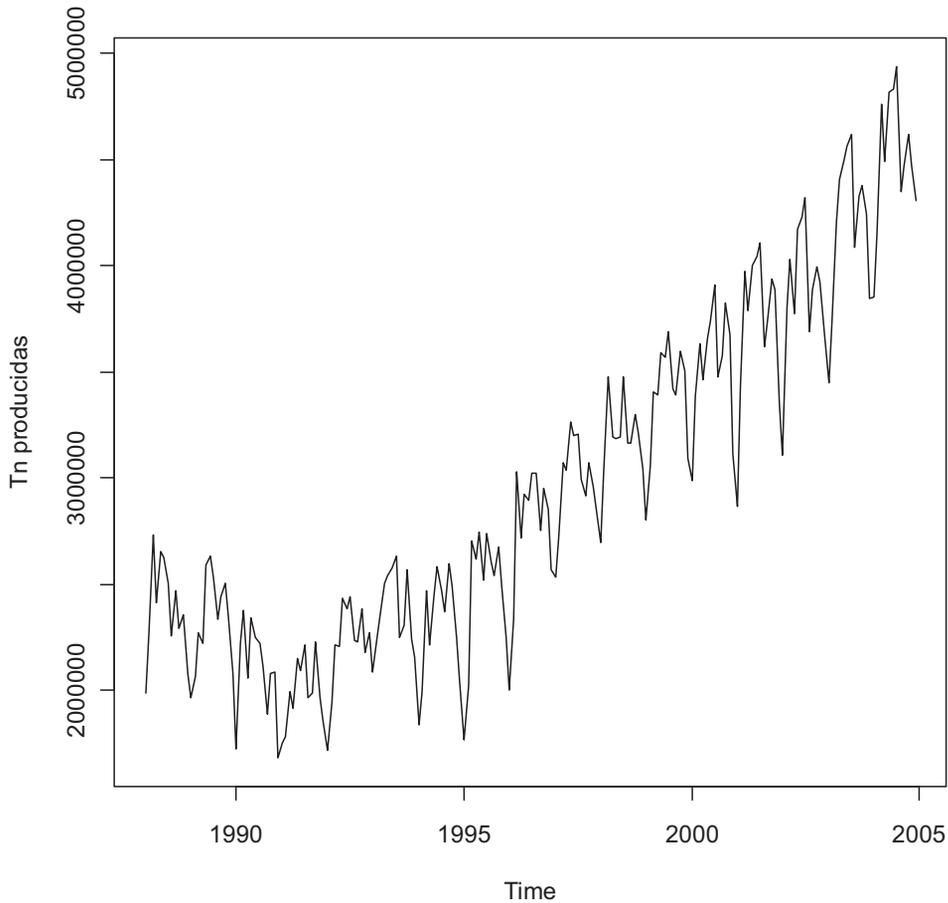


Figura 1.3. Serie $Z_t = \text{"cemento"}$ contiene la producción mensual en toneladas de cemento en España entre 1988 y 2004

En la Figura 1.3, se observa una fuerte tendencia al alza entre 1995 y 2005 correspondiente al crecimiento exponencial del sector inmobiliario en esa década después de una bajada suave entre 1988 y 1992. Se puede apreciar una clara componente estacional.

Se ha realizado la descomposición de la serie en sus componentes tendencia, estacionalidad y aleatoria. En la Figura 1.4 observamos cada componente por separado y podremos determinar que efectivamente el patrón de comportamiento estacional es también muy marcado.

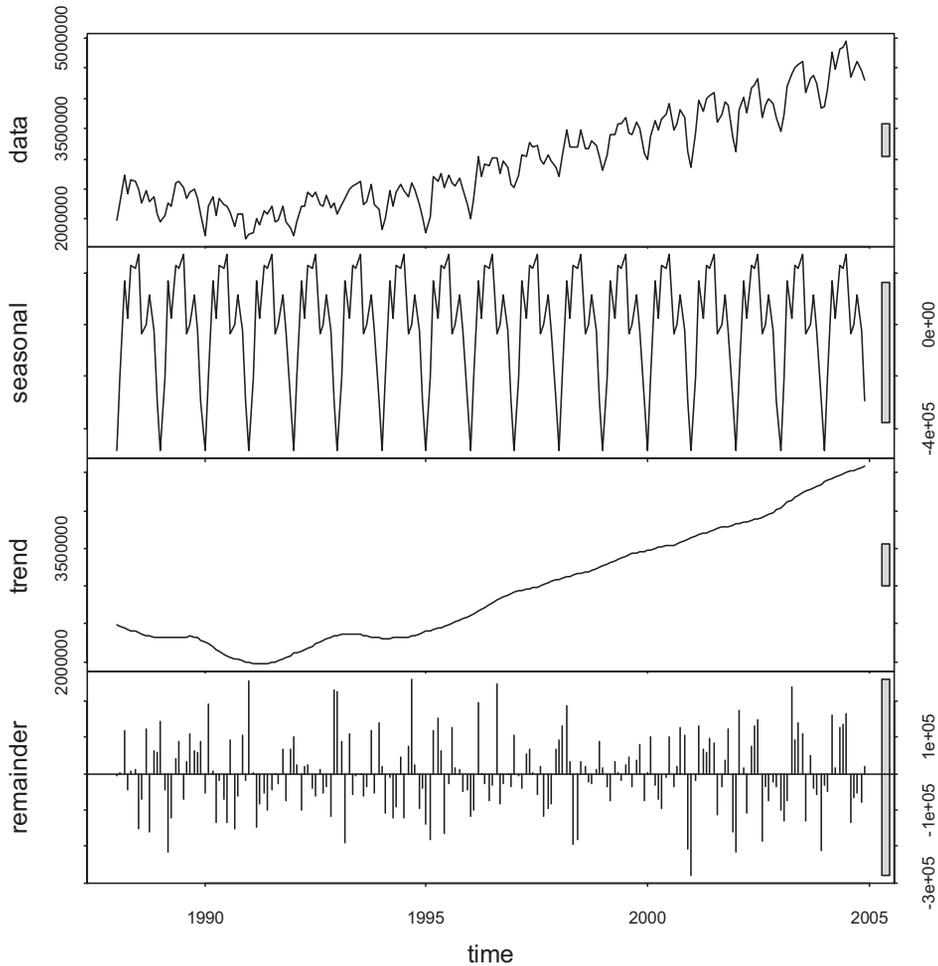


Figura 1.4. Descomposición de la serie "cemento" en sus componentes tendencia, estacionalidad y aleatoria

Ejemplo 1.3. Descomposición de la serie $Z_t = \text{"azulejos"}$ correspondiente a la exportación mensual en m^2 de baldosas cerámicas en España entre 1990 y 2006. Se ha obtenido la descomposición de la serie en sus componentes tendencia, estacionalidad y aleatoria.

En la Figura 1.5, se observa una fuerte tendencia al alza entre 1993 y 2003 correspondiente al crecimiento exponencial del sector inmobiliario en esa década seguida de un estancamiento. No se puede apreciar una clara componente estacional.

En la Figura 1.6, vamos a poder observar cada componente por separado y podremos determinar que existe un el patrón de comportamiento estacional muy marcado. En este caso podemos observar claramente como la tenencia de las exportaciones de baldosas cerámicas había llegado a un nivel máximo estabilizado a partir del 2004.

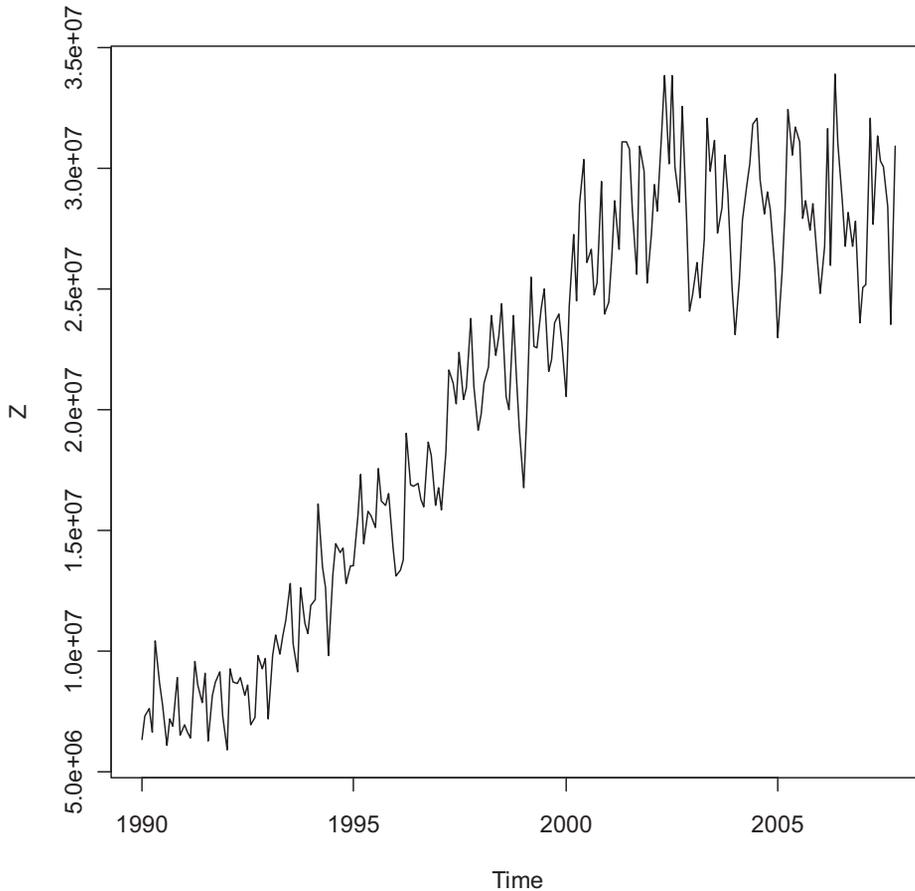


Figura 1.5. Serie $Z_t = CEEsem$ del consumo en MWh de energía eléctrica horaria en España de una semana laborable de lunes a viernes

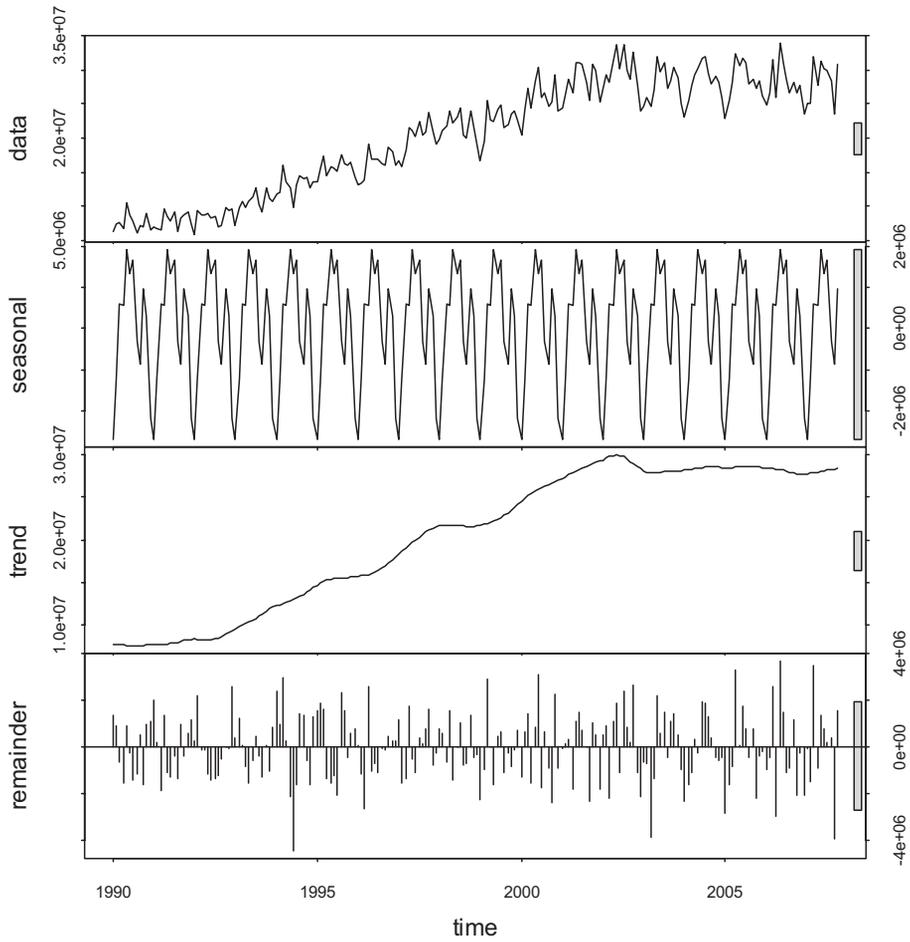


Figura 1.6. Descomposición de la serie "azulejos" en sus componentes tendencia, estacionalidad y aleatoria

Estas componentes pueden estar actuando según un esquema aditivo, multiplicativo o mixto para dar el comportamiento global observable de la serie temporal.

Según el **esquema aditivo** la serie es el resultado de la suma de las cuatro componentes, lo cual significa que pueden tratarse de forma independiente.

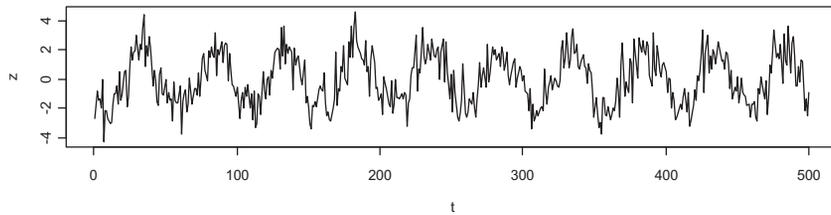
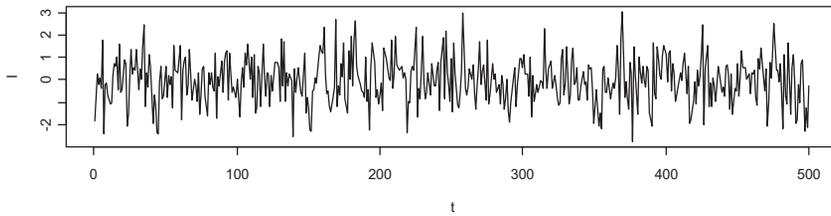
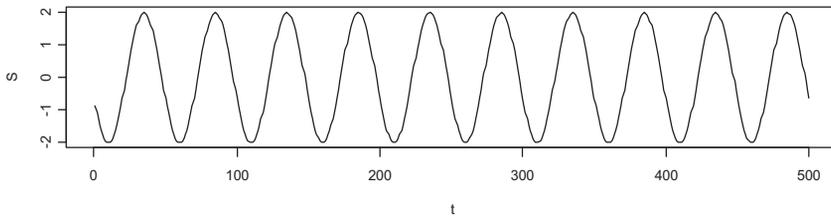
$$Z_t = T_t + C_t + S_t + I_t \quad (3.1)$$

Ejemplo 1.4. Obtener una serie temporal $Z_t = S_t + I_t$ mediante la combinación lineal de una serie determinista senoidal (S_t) y una serie aleatoria (I_t) representándolas gráficamente.

$$S_t = 2\cos(2\pi t/50 + 0.6\pi)$$

$$I_t = \text{Normal}(0,1)$$

$$Z_t = S_t + I_t$$



Mientras que en el caso de que exista alguna relación entre componentes y no se puedan tratar como independientes estaríamos ante un **esquema de tipo multiplicativo**:

$$Z_t = T_t \cdot C_t \cdot S_t \cdot I_t \tag{3.2}$$

Para seguir leyendo haga click aquí